ELSEVIER

Journal of Chromatography A, 804 (1998) 37–43

JOURNAL OF CHROMATOGRAPHY A

Statistical quality control applied to ion chromatography calibrations

Frederick S. Stover*, Robert V. Brill

Technology and Commercial Development, Solutia Inc., 10300 Olive Boulevard, P.O. Box 66760, St. Louis, MO 63166-6760, USA

Abstract

Multivariate statistical quality control principles, including control charting of calibration parameters and control samples, are applied to multilevel ion chromatography (IC) calibrations to determine instrument response stability and minimum calibration frequency. For phosphate species quantitated by suppressed-conductivity IC with NaOH gradient elution, ortho- and pyrophosphate exhibit stable responses over the 10-week course of the study, while tripolyphosphate shows a response that decreases with increasing eluent age. Analyses of control samples are used to compare the effect of different calibration protocols on long-term method precision and accuracy. Cases are found where using a single, averaged calibration curve for extended periods can give either better or worse precision than a method which employs daily instrument calibration. Optimum calibration frequencies determined for phosphate analyses were weekly for tripolyphosphate, and ≥10 weeks for ortho- and pyrophosphate. Methods for detecting when a calibration has shown a statistically significant change are discussed. Statistical calculations are presented in a simple algebraic form amenable to the use of spreadsheets for data analysis.

Keywords: Statistical quality control; Quality control; Calibration methods; Precision; Principal component analysis; Hotelling's T^2 value; Ion chromatography; Chemometrics; Phosphates

1. Introduction

It has been our experience that analysts using chromatographs to quantitate chemical species often calibrate instruments by running standards daily, or prior to each use. These practices are in contrast to those used with process analyzers, where recalibration of an on-line instrument may require a process upset or analyzer removal. For process analysis or simpler instrumental methods, control sample or standard analysis results alone are often used to verify proper instrument operation and to minimize calibration frequency [1–3]. Control limits placed on the latter methods can be rigorously derived from statistical process control principles applied to evaluation of the measurement process, or statistical quality control (SQC) [4].

Use of SQC in the chromatography laboratory usually involves implementing control charts for QC standards to determine and verify long-term method performance [5–7]. Less common is the use of the more powerful aspects of SQC for minimizing calibration frequency and avoiding ''over-calibration'' errors [8]. Multivariate SQC is best utilized for this purpose. Examples of using SQC for optimizing recalibrations and minimizing precision and bias errors can be found for X-ray fluorescence [9], atomic emission [10], vapor pressure [11], photometry [12] and immunoassay [13]. For chromatographic methods, SQC examples mainly deal with process analyzer applications [2,8] where the benefits of reducing calibration frequency while maintaining accuracy and precision are well recognized. SQC

*Corresponding author.

0021-9673/98/$19.00
PII S0021-9673(98)00094-6

appears under-utilized as a tool in the chromatographic laboratory for studying calibrations and their effect on overall method performance.

The present study was undertaken to apply modern multivariate SQC practices to ion chromatography (IC) calibrations in an effort to determine the proper calibration frequency for a typical IC system with suppressed conductivity detection. It was suspected that an IC is stable enough (drift-free) to reduce the calibration frequency to well below the typical once/day. For a stable system, indiscriminate use of daily calibration curves actually can be deleterious to method performance. If the error in an analytical result obtained by an external standard calibration curve procedure is due to random error in both calibrant and sample measurements ($\sigma^2_{\text{total}}=\sigma^2_{\text{cal}}+\sigma^2_{\text{sample}}$), it is reasonable to assume that if the calibration is not changing with time, better method performance can be obtained by using a fixed, predetermined curve until a statistically significant change in the instrument response occurs.

We present here the results of 44 daily, four-level IC calibrations for ortho-, pyro- and tripolyphosphate, and concurrent results from the analyses of two control samples. IC response stability, impact of daily recalibrations, optimum calibration frequencies, long term precision and control sample and eluent stability are presented for a typical IC method for the three phosphate anions.

2. Experimental

2.1. Chemicals

Phosphate anion standard solutions were prepared from the salts tetrasodium pyrophosphate decahydrate (TSPPD, $Na_4P_2O_7\cdot 10H_2O$) and pentasodium tripolyphosphate hexahydrate (STPH, $Na_5P_3O_{10}\cdot 6H_2O$), obtained from Sigma–Aldrich (St. Louis, MO, USA), and potassium phosphate monobasic (MKP, KH_2PO_4, certified) from Fisher Scientific (Pittsburgh, PA, USA). Sodium tripolyphosphate anhydrous (STP, $Na_5P_3O_{10}$) and disodium phosphate dihydrate (DSPD, $Na_2HPO_4\cdot 2H_2O$) were obtained from Solutia (St. Louis, MO, USA). 200 m*M* NaOH was prepared by diluting 10.5 ml reagent grade 50% NaOH to 1 l with deionized water degassed with He. All water was deionized and obtained from a Millipore Milli-Q 4-Bowl Plus analytical purification system (Bedford, MA, USA). Eluents were used and stored under ~3 p.s.i.g. He in glass bottles (1 p.s.i.=6894.76 Pa).

2.2. Instrumentation

IC was performed on Dionex (Sunnyvale, CA, USA) equipment consisting of a Model ASM autosampler, Model AGP gradient pump with microbore heads, a 50×2 mm ATC trap column, a CSI column stand/injector with a Rheodyne 9126 microbore valve and 25-μl sample loop, a 2 mm anion self-regenerating suppressor (ASRS) and controller module used in the recycle mode, and a CDM-2 conductivity detector and cell at 30 μS full scale. A 50×2 mm AG11 guard column and a 250×2 mm AS11 column were used in series. Data acquisition and instrument control were accomplished using a Dionex AI-450 data system.

2.3. Methods

A mixed stock standard containing 100 μg/ml PO_4^{3-}, $P_2O_7^{4-}$ and $P_3O_{10}^{5-}$ was prepared by diluting 71.5 mg MKP, 128.2 mg TSPPD and 94.1 mg STPH to 500 ml. Calibrations were performed by injecting mixed 0.5, 2.0, 5.0 and 10.0 μg/ml standards prepared by diluting 50, 200, 500 and 1000 μl, respectively, of the stock standard to 10 ml. Control sample stock solutions were prepared by diluting 66.1 mg DSPD or 144.1 mg STP to 100 ml. Control samples were prepared by diluting 100 μl of the stock solutions separately to 10 ml. A Rainen (Woburn, MA, USA) EDP2 electronic digital pipet was used for all dilutions. Stock standards and control samples were stored refrigerated in polyethylene bottles at 4°C, and were allowed to come to ambient temperature for ~1 h just prior to dilution. Standards and control samples were analyzed within 5 h after dilution. IC separations were performed using a linear NaOH gradient from 20 to 70 m*M* NaOH over 20 min and holding for 2 min. Integration of IC peak areas was performed manually using the AI-450 data system optimize screen.

3. Results and discussion

3.1. Multivariate methods for control charting calibration parameters

Phosphate calibrations and control sample results from this study were generated to test long-term calibration stability and to optimize calibration frequencies for a typical IC system. Each four-level calibration set was subjected to linear least-squares regression to obtain slope and intercept calibration parameters. While the use of individual slope and intercept control charts can indicate general calibration trends, it is not advisable to base recalibration decisions on such plots. For example, it is possible for a point to appear ''in-control'' when viewed this way when, in reality, that particular combination of slope and intercept is not typical. Two procedures for making a statistically sound judgment on whether a calibration has changed are presented here.

The standard procedure to determine whether a multivariate response is out-of-control is to use Hotelling's T^2 [14], which is a measure of the ''statistical distance'' between a particular multivariate response point and the center of the points used to estimate the population of in-control values. For the case of a linear regression $y=ax+b$, Hotelling's T^2 is given by

$$T_i^2 = [s_b(a_i - a_0)^2 + s_a(b_i - b_0)^2 - 2s_{ab}(a_i - a_0)(b_i - b_0)]/[s_a s_b - (s_{ab})^2] \quad (1)$$

where a_i and b_i are the slope and intercept of the ith observation, a_0 and b_0 are the average slope and intercept, and variances s are defined in Table 1. Critical values for T^2 can be calculated and used as upper control limits in a T^2 control chart. T^2 has the distribution

$$T^2 \sim [p(n+1)(n-1)F(p, n-p, \alpha)]/[n(n-p)] \quad (2)$$

where $F(p, n-p)$ is the f distribution with p numerator and $n-p$ denominator degrees of freedom. p is the number of components in each multivariate observation (two in this case) and n is the number of observations used to estimate the population of in-control values. For example, the first 20 in-control orthophosphate calibrations were used to set the control limits for that chart, $n=20$. Critical

Table 1
Example calculations and summary statistics for Hotelling's T^2 and first principal component (PC1) control charts for orthophosphate calibrations based on runs 1, 3–10 and 12–22

Run (i)	Slope (a_i)	Intercept (b_i)	Normalized		T^2	PC1
			Slope (A_i)	Intercept (B_i)		
1	790 952	−33 236	−2.288	−0.241	5.979	−1.447
2	768 709	202 925	−4.540	8.504	78.167	−9.224
3	805 615	32 353	−0.803	2.118	4.847	−2.115
⋮	⋮	⋮	⋮	⋮	⋮	⋮
44	793 217	−47 536	−2.058	−0.770	6.060	−0.911

Summary statistics for orthophosphate calibrations

Description	Symbol	Value
Average intercept	b_0	−26 735
Average slope	a_0	813 541
Intercept variance	s_b	729 259 476
Slope variance	s_a	97 502 113
Covariance	s_{ab}	−69 170 174
Correlation coefficient	r	−0.27305
No. of in-control points	n	20
No. of components	p	2
F critical value	$F(2, 18, 0.0027)$	8.3639
T^2 critical value for $\alpha=0.0027$	–	18.54

values for T^2 are calculated by choosing a type-I error ($\alpha = 0.05$ or 0.0027 corresponding to the 2σ and 3σ limits for univariate control charts) and looking up the corresponding f value (usually a spreadsheet function) for use in Eq. (2).

A second procedure for detecting out-of-control calibrations is to use principal component analysis (PCA) [15]. When the components of a multivariate response are correlated, such as the slope and intercept from a linear regression of calibration data, PCA can be used to reduce the dimensionality of the response space. This is done by creating a reduced set of variables, t_i which (1) are linear combinations of the original components, (2) are orthogonal to each other and (3) contain most of the information in the response space and can be used as an efficient way to represent that space. The first principal component is often sufficient to contain most of the information from the slope and intercept values.

Before calculating principal component scores, the slope and intercept data are each centered and scaled to a mean of 0 and standard deviation of 1 by taking the individual observations, subtracting the average, and dividing by the standard deviation (for the n in-control values). The first principal component for the ith observation is given by

$$t_{i,1} = 0.707A_i \pm 0.707B_i \quad (3)$$

where A_i and B_i are the normalized slope and intercept, and the sign is the same as that of the correlation coefficient between slope and intercept. Since we are using only one principal component, a normal (univariate) individuals control chart can be used.

Example calculations for Hotelling's T^2 and the first principle component from the orthophosphate calibration data are given in Table 1.

3.2. Phosphate species calibration stability

Inspection of both control sample and calibration data generated during this study indicates that a generally stable response can be obtained over several months for ortho- and pyrophosphate. Conversely, for tripolyphosphate, a decreasing response is observed during times between eluent preparation. These cases have distinct implications for calibration frequency and control chart strategies. A key feature of the results is whether the use of daily calibration curves vs. a single, average calibration gives better long-term precision (σ_{DAILY} vs. σ_{AVE}) as measured by control sample results.

For pyrophosphate, a stable response ($\sigma_{\mathrm{DAILY}} > \sigma_{\mathrm{AVE}}$) is seen from the analysis of the STP control sample containing ~1 μg/ml $P_2O_7^{4-}$. Fig. 1a plots the pyrophosphate concentration obtained for the control sample vs. time using (1) daily calibration curves for each measurement and (2) using a single calibration based on the average slope and intercept from the first 22 curves. The time frame of these control charts covers 10 weeks at 4–5 runs/week, and control limits are calculated from the first 22

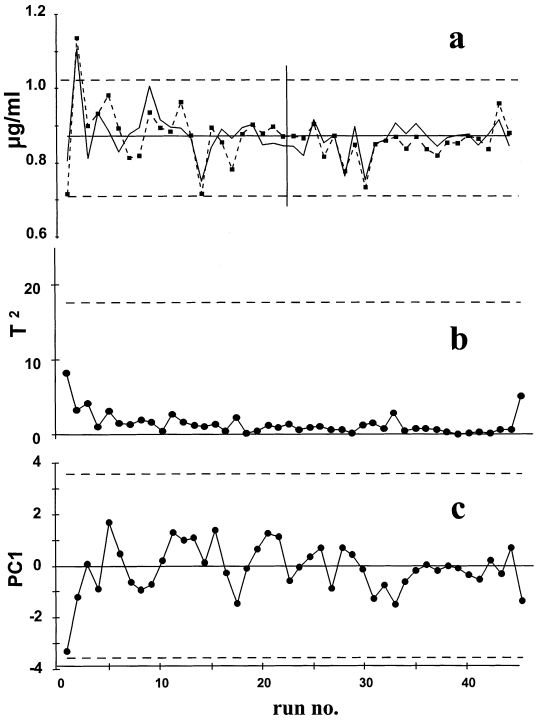

Fig. 1. Control charts of (a) STP control sample results, and (b) Hotelling's T^2 and (c) first principal component (PC1) of calibration parameters for pyrophosphate. Dashed lines are $\pm 3\sigma$ control limits, or the equivalent, based on in-control results from the first 22 runs (points before vertical line). (a) – ■ –, results using daily calibration; —, results using average in-control calibration parameters from the first 22 runs. Control limits based on average calibration results. Average μg/ml±S.D. (for all points except No. 2) = 0.861±0.057 for daily calibration and 0.866±0.047 for average calibration.

in-control values using the average calibration. Results from all 44 runs show that use of a single calibration instead of daily calibrations reduces the relative standard deviation (R.S.D.) of the analysis from 6.6 to 5.4% without introducing bias.

Fig. 1b and c plot the Hotelling's T^2 (T^2) and first principal component (PC1) control charts for the calibrations run concurrently with the control samples. As expected, both plots show that the pyrophosphate calibration was in-control throughout the 10-week course of the study. This stable response for pyrophosphate means that use of daily calibration curves causes a degradation in method performance vs. use of the stable, average calibration. This is a classic example of ''over-calibration'' error, where increased calibrations simply add to the overall method error since there is no calibration drift which needs corrected. Significant time savings can be realized by using control sample analyses and charting to replace frequent instrument calibration for such in-control analyses.

For tripolyphosphate, a drifting response ($\sigma_{DAILY} < \sigma_{AVE}$) is seen both from the PC1 calibration parameter plot and from the analysis of the STP control sample containing 9.3 μg/ml $P_3O_{10}^{5-}$, as shown in Fig. 2. The cyclic nature of the results is readily apparent, and can be tied directly to fresh eluent preparation. From a separate plot of calibration slope vs. time, the response to tripolyphosphate was seen to steadily decreases by about 10% during the two weeks between eluent preparation. The effect of this drift on control sample results using daily vs. average calibration can be seen in Fig. 2a. Ignoring the out-of-control results in runs 14 and 28, daily calibrations (which correct for the response drift) give an improvement in R.S.D. from 4.6 to 1.8% vs. an average calibration. In this case, it is clear that frequent calibration is needed to compensate for the changing response. Nevertheless, close examination of the data shows that a reduction in calibration frequency to once/week (within four days after eluent preparation) is possible while maintaining good precision (<2% R.S.D.).

Changes in chromatographic performance accompanied the response shifts seen in Fig. 2. Separations using aged NaOH eluent showed shifted retention times and increased tailing of tripolyphosphate peaks. The origin of this effect was investigated by

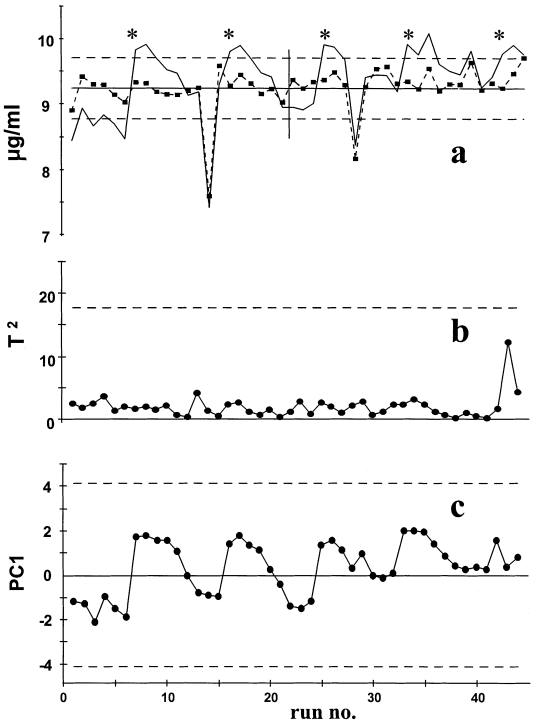

Fig. 2. Control charts of (a) STP control sample results, and (b) T^2 and (c) PC1 of calibration parameters for tripolyphosphate. Asterisks denote days when fresh eluent was prepared. Other notation as in Fig. 1, except control limits in (a) based on daily calibration results. (a) Average μg/ml±S.D. (for all points except Nos. 14 and 28)=9.31±0.17 for daily calibration and 9.42±0.43 for average calibration.

pH titration of eluents with acid to determine the CO_3^{2-} content. No carbonate could be detected (<10 μg/ml) in either freshly prepared or 11-day-old NaOH, although the base strength did decrease from 199 to 197 m*M* on aging. Elemental analysis of fresh vs. aged eluent showed an increase in Si content from <1 to 12 μg/ml, as expected for NaOH stored in glass bottles. Whether this change in silicate content is sufficient to cause the drifts observed in the chromatography is being explored (by e.g., changing to plastic eluent reservoirs).

For orthophosphate, an intermediate case of ''moderate'' response stability ($\sigma_{DAILY} \sim \sigma_{AVE}$) is observed. Fig. 3a plots the orthophosphate concentration found in the DSPD control sample using daily calibrations vs. those found with an average calibration from the first 20 in-control curves. Results

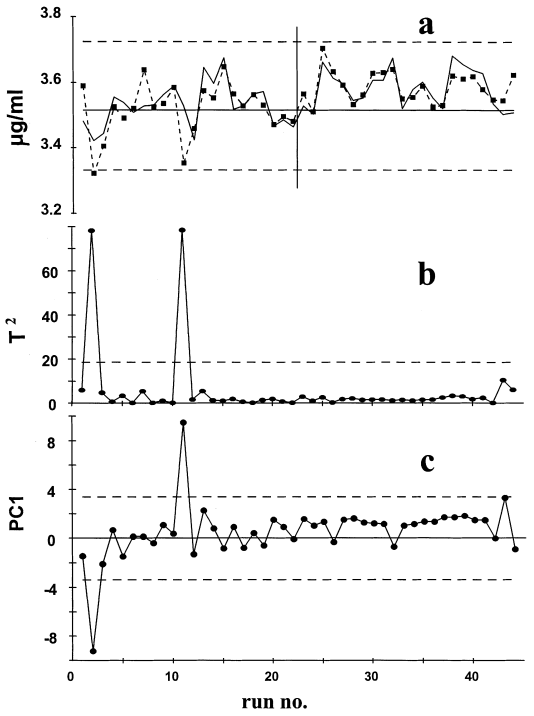

Fig. 3. Control charts of (a) DSPD control sample results, and (b) T^2 and (c) PC1 of calibration parameters for orthophosphate. Control limits and other notation as in Fig. 1. (a) Average µg/ml±S.D. (for all points except Nos. 2 and 11)=3.562±0.060 for daily calibration and 3.560±0.067 for average calibration. Theoretical=3.537 µg/ml.

for all 44 runs fell within the 3σ control limits when the average calibration was used, indicating no major change in instrument response. However, daily calibrations gave an out-of-control PO_4^{3-} result for run 2 and a warning for run 11. Also, the T^2 and PC1 charts in Fig. 3b and c indicate out-of-control calibrations for runs 2 and 11. If the control sample results from an average calibration showed a reasonably stable response over all 44 runs, why did the calibration control charts indicate a changed response for two runs?

Inspection of the areas for individual standards showed that the out-of-control calibration results were due to a single gross error in one standard for each of these runs, not a general change in all standards typical of an instrument response shift. A real shift in instrument response should have resulted in *both* the control sample (average curve) and calibration parameter charts going out of control. The utility of charting control samples and calibration parameters is clearly demonstrated since, without their use, these gross errors in a single standard might not have been detected.

For the orthophosphate control sample, the precision found from the daily calibration procedure (neglecting out-of-control calibrations) was slightly better than that found using the average curve. Examination of the calibration data showed a slight upward shift in slope for the later runs. This small response drift is responsible for the interesting intermediate result for orthophosphate ($\sigma_{\mathrm{DAILY}} \sim \sigma_{\mathrm{AVE}}$), where the normal over-calibration error is balanced by slightly better tracking of a small response change. However, these small changes were insufficient to require recalibration during the course of the study, since orthophosphate control sample analysis at 3.5 µg/ml shows similar (<2% R.S.D.) precision using either a single (average) or daily calibrations, with no bias between the two approaches.

4. Conclusions

Based on the results of this study, we can make the following recommendations for performance monitoring of IC, other chromatography, and multilevel external standard calibration methods in general. (1) Routine monitoring of instrument performance can be economically performed by control charting ''control samples'' or other surrogates. (2) When the ''control sample'' control chart indicates a change in instrument response, a new calibration curve can be generated. (3) The new calibration parameters should be plotted on a multivariate control chart. (4) If the calibration control chart indicates that the instrument response has changed in a statistically significant manner, consistent with the change observed in the control sample, the new calibration should be implemented and the control sample reanalyzed. (5) If the calibration control chart indicates that the instrument response has not changed, or does not explain the control sample result, possible errors in control sample or standard preparation, stability, chromatography, etc., should be investigated.

While simple control sample monitoring can drastically reduce calibration frequency and improve overall IC precision, some periodic runs of calibration standards should be performed for maintaining the multivariate control charts. For example, weekly calibration runs are useful for monitoring the long term precision of the calibration parameters. The key point is that the new calibration parameters should be merely recorded and used to update control charts until a statistically sound indication of a change in the instrument response is detected. Based on our experience with this data, multivariate control charts using Hotelling's T^2 and, particularly, the first principal component perform well for this purpose.

Quantitation of phosphates using suppressed-conductivity, gradient IC has been shown to give stable responses (calibrations) for ortho- and pyrophosphate, while tripolyphosphate response tends to decrease with increasing age of the NaOH. For cases where the response is stable, use of a single calibration curve over several months can give equal or superior method precision vs. daily instrument calibration. The optimum calibration frequency for a particular analysis can be inferred from Table 2. We have found in our laboratory that proper calibration frequencies are $\geq$10 weeks for ortho- and pyrophosphate and weekly for tripolyphosphate.

Table 2
Relative standard deviations observed as a function of calibration frequency

Control sample analysis	R.S.D. (%) calculated for a given recalibration frequency		
	Daily	Weekly	$\geq$10 Weeks
3.5 μg/ml PO_4^{3-} (DSPD)	1.7	–	1.9
0.86 μg/ml $P_2O_7^{4-}$ (STP)	6.6	–	5.4
9.3 μg/ml $P_3O_{10}^{3-}$ (STP)	1.8	2.1	4.6

These results should be useful for other laboratories contemplating a reduction in IC calibration frequency. In addition, implementing a general SQC program, including control charting of calibration parameters and control samples, can lead to a greater understanding of the sources of error in external standard IC methods. One such discovery in the current study was the dependence of tripolyphosphate response on NaOH eluent age, the origins of which are being further investigated.

References

[1] J.C. Miller and J.N. Miller, Statistics for Analytical Chemistry, Ellis Horwood, Chichester, 1984.
[2] L.F. Peveto, J.G. Converse, ISA Trans. 30 (1991) 79.
[3] W.L. Paul, N.S. Barrett, Lab. Autom. Inf. Manage. 31 (1995) 141.
[4] D.J. Wheeler and R.W. Lyday, Evaluating the Measurement Process, Statistical Process Controls, Knoxville, TN, 1984.
[5] A.R. Fernandez-Alba, A. Valverde, A. Aguera, M. Contreras, S. Chiron, J. Chromatogr. A 721 (1996) 97.
[6] M. Hiemstra, J.A. Joostem, A. de Kok, J. AOAC Int. 78 (1995) 1267.
[7] D.H. McMahon, R.G. Thurman, E.E. Balder, J. Chromatogr. Sci. 19 (1981) 349.
[8] G.D. Nichols, Adv. Instrum. Control 45 (1990) 349.
[9] A. Mueller, G. Staats, V. Troebs, Fresenius J. Anal. Chem. 348 (1994) 615.
[10] P. Fallavier, G.H. Llimous, M. Breysse, P. Cazevieille, Commun. Soil. Sci. Plant. Anal. 27 (1996) 441.
[11] H. Schneider, J.M. Pruett, L. Magee, ISA Trans. 32 (1993) 355.
[12] O. Mestek, J. Pavlik, M. Suchanek, Fresenius J. Anal. Chem. 350 (1994) 344.
[13] E. Jervis, D.G. Kilburn, Biotechnol. Prog. 7 (1991) 28.
[14] S.J. Wierda, Statistical Neerlandica 48 (1994) 147.
[15] A. Block, A.K. Smilde, K.A. Duineveld, P.M. Coenegracht, J. Chemometrics 10 (1996) 351.