ANALYTICA CHIMICA ACTA 642 (2009) 222–227

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/aca

Quality control of packed raw materials in pharmaceutical industry

O.Ye. Rodionova[a,*], *Ya.V. Sokovikov*[b], *A.L. Pomerantsev*[a]

[a] *Semenov Institute of Chemical Physics, Kosygin 4, 119991 Moscow, Russia*
[b] *SchelTec AG, Kosygin 19, 119334 Moscow, Russia*

ARTICLE INFO

Article history:
Received 25 June 2008
Received in revised form
19 July 2008
Accepted 8 August 2008
Published on line 19 August 2008

Keywords:
NIR measurements
PAT
Fiber optic probe
Measurements through PE packaging
Classification limits
SIMCA

ABSTRACT

The possibility of routine testing of pharmaceutical substances directly in warehouses is of great importance for manufactures, especially taking into account the demands of PAT. The application of NIR instruments with remote fiber optic probe makes these measurements simple and rapid. On the other hand carrying out measurements through closed polyethylene bags is a real challenge. To make the whole procedure reliable we propose the special trichotomy classification procedure. The approach is illustrated by a real-world example.

1. Introduction

Rapid, non-destructive analysis of chemical and physical properties of pharmaceutical substances is achievable using NIR measurements. Fiber optic probe can measure remote samples in diffuse reflectance or transmission modes [1]. As a result, testing can be performed not only in the lab but also directly in a warehouse. Raw materials are quickly tested for identity and quality conformance. Once a model for a substance is developed, routine testing takes place in a few seconds, making it possible to test every unit of incoming ingredient to verify the identity. This procedure is an essential part of Process Analytical Technology (PAT) [2,3], which is widely applied in pharmaceutical industry nowadays.

At the same time the application of fiber probe diffuse reflectance NIR spectroscopy is an especially challenging problem when measurements are carried out through closed polyethylene (PE) bags. This could produce spectral artifacts that are comparable with the substance physical or chemical fingerprints. Additional problems arise due to changes in the position of the probe held by an operator.

In this paper we consider a two-stage approach for classification that allows an operator to recognize raw substances of satisfactory quality reliably in spite of these adverse circumstances. It is worthy of mentioning that each confident classification implies some errors that manifest themselves in two misclassification events. The first one is the case when a sample of good quality is rejected, i.e. classified as a forgery

* *Corresponding author*. Tel.: +7 495 9397483; fax: +7 495 9397483.
E-mail address: rcs@chph.ras.ru (O.Ye. Rodionova).
0003-2670/$ – see front matter
doi:10.1016/j.aca.2008.08.004

or another substance (Type I error). The second event is the recognition of a fake sample as a genuine one (Type II error) [1,4].

Common approach to the classification problem is to construct an acceptance area using the training (or calibration) set of samples of the required quality. This area is further used in a routine recognition in such a way that a new (or a test) sample is accepted if its spectrum hits the area. The area size depends on a given Type I error value that should be small enough, say, $\gamma = 0.01$, or 1%. However, the smaller the γ, the wider the area, and at $\gamma = 0$ the acceptance area covers all possible samples. Simultaneously, the Type II error is growing. Therefore, the problem is to find the smallest (effective) acceptance area for a given γ value. This is a very complicated mathematical task, which is fully solved only for some simple cases involving the normal distributions. In this study we use a special procedure [5] for building an effective acceptance area within the SIMCA approach [6,7].

Typically, a pattern recognition result is a dichotomous decision when a new sample is either accepted, or rejected. We suggest applying a trichotomy recognition: accepted, rejected, or extra measurements are required. The latter decision is made if a new spectrum bears the signs of the abovementioned artifacts. Such an approach is in line with a method of the sequential test of statistical hypotheses that was introduced by Wald [8] as far back as in the 40s.

The advocated approach includes the following main steps:

- the appropriate splitting of the initial calibration objects into two classes employing a global PCA model;
- the construction of two separate PCA models with an apt number of principal components (PCs);
- the application of Soft Independent Modeling by Class Analogy (SIMCA) with the ad hoc method [5] for the acceptance area calculation.

The approach is illustrated by a real-world example.

2. Theory

2.1. Principal component analysis (PCA)

Data sets with many variables can be simplified through variable reduction and thereby be more easily interpreted. Principal component analysis (PCA) [9] is a well-known variable reduction technique, in which spectral matrix $\mathbf{X}$ is decomposed as

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{t} + \mathbf{E}, \tag{1}$$

Here $\mathbf{X}$ is the $I \times J$ data matrix, $\mathbf{T}$ the $I \times A$ matrix of score vectors, $\mathbf{P}$ the $J \times A$ matrix of loading vectors, $\mathbf{E}$ the $I \times J$ residual matrix, I the number of objects, J the number of variables (which in our case is the number of wave numbers), and A is the number of components calculated (i.e., principal components, PCs). Matrix $\mathbf{X}$ is preprocessed by mean-centering.

The $(A \times A)$ matrix $\mathbf{\Lambda}$

$$\mathbf{\Lambda} = \mathbf{T}^{t}\mathbf{T} = \mathrm{diag}(\lambda_1, \ldots, \lambda_A) \tag{2}$$

is diagonal with the elements

$$\lambda_a = \sum_{i=1}^{I} t_{ia}^2 \tag{3}$$

They are the first A eigenvalues of matrix $\mathbf{X}^{t}\mathbf{X}$ ordered descendingly.

Two important characteristics of the PCA model with respect to each calibration object can be defined. They are the score distance (SD) and the orthogonal distance (OD). For a given number of principal components, A, the SD is defined as [9]

$$h_i = \mathbf{t}_i^{t}(T_A^{t}T_A)^{-1}\mathbf{t}_i = \sum_{a=1}^{A} \frac{t_{ia}^2}{\lambda_a}, \quad i = 1, \ldots, I. \tag{4}$$

It is equal to the squared Mahalanobis distance from the model center to sample i within the score subspace. The average SD is calculated as [5]

$$h_0 = \frac{1}{I}\sum_{i=1}^{I} h_i \equiv \frac{A}{I}. \tag{5}$$

The OD, v_i, is calculated as the sum of the squared residuals presented in matrix $\mathbf{E} = \{e_{ij}\}$

$$v_i = \sum_{j=1}^{J} e_{ij}^2. \tag{6}$$

The OD is the squared Euclidian distance from object i to the model subspace. The average OD value is calculated as

$$v_0 = \frac{1}{I}\sum_{i=1}^{I} v_i. \tag{7}$$

2.2. Soft independent modeling of class analogy (SIMCA)

SIMCA is a supervised pattern recognition method [10]. The idea behind this method is that each group of objects is independently subjected to PCA with its own complexity (the number of PCs) and the acceptance area for each model is defined. In this study we apply a modified SIMCA approach that was earlier presented by one of the authors [5]. This approach is based on the following principles.

(1) It is supposed that both the SDs and ODs values are chi-square distributed, namely

$$\frac{h}{h_0} \propto \frac{1}{N_h}\chi^2(N_h) \quad \frac{v}{v_0} \propto \frac{1}{N_v}\chi^2(N_v), \tag{8}$$

notation $z \propto G$ means that variable z follows the G distribution.

(2) These chi-squared distributions depend on parameters N_h and N_v that are the numbers of degrees of freedom (DoF). Their values are not derived from theoretical considera-

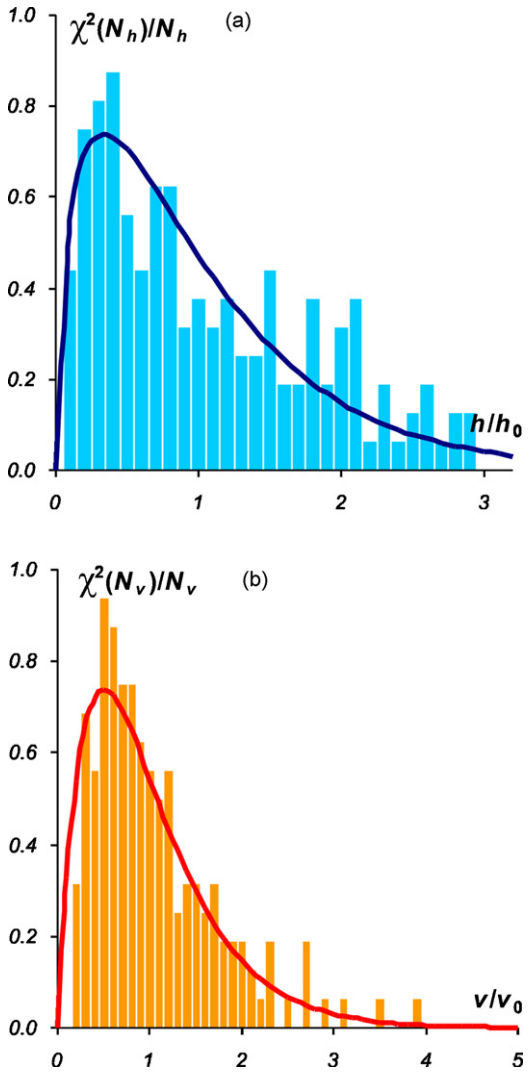

Fig. 1 – Sampling (bars) and estimated (curves) distributions for the SDs (a) and ODs (b).

tions, but estimated from the calibration set, employing values h_i and v_i, $i = 1, \ldots, I$.

These two principles are illustrated in Fig. 1. It represents two histograms (bars) and the corresponding χ^2 densities (curves). The plots are constructed using the calibration data set of Group 1 which is described below. It consists of 160 objects and the estimated DoFs are $N_h = 3$, $N_v = 4$.

(3) For the construction of the acceptance area for a given Type I error γ, the following approach is applied. Firstly, the χ^2 distributed values $N_h h/h_0$ and $N_v v/v_0$ are converted with the Wilson–Hilferty transformation [11]

$$\frac{(\chi^2/N)^{1/3} - (1 - s^2)}{s} \propto N(0, 1), \quad s^2 = \frac{2}{9N}, \tag{9}$$

that gives two independent standard normal variates. Then the acceptance area in the transformed coordinates is easily calculated for a given γ. At last, the area is transformed back into initial coordinates h/h_0 and v/v_0.

3. Materials and methods

The substance under investigation is Taurine, a non-essential sulfur-containing amino acid. The NIR spectra were recorded on the Spectrum 100N FT-NIR spectrometer (PerkinElmer UK) fitted with a hand held diffuse reflectance fiber optic probe. The spectra were measured through closed polyethylene bags in the 4000–10,000 cm^{-1} region with a 2 cm^{-1} spectral resolution. Prior it is known that all samples present the substance of satisfactory quality. All spectra were pre-treated by the SNV procedure.

For detailed data processing, model analysis, and general approach elaboration various spectra were measured: those of the substance in closed bags, substance without packaging, and empty polyethylene bags. 246 spectra (each bag was measured 3 times in different places) were recorded for 82 drums (substance in the closed bags). They comprise data set A. Three bags were opened and NIR spectra of the substance (5 replicates per drum) were recorded by means of the same fiber optic probe. These 15 objects were collected in the S data set. The spectra of 18 empty polyethylene bags folded several times were measured by the same equipment. The 18 spectra comprise data set P.

For testifying purposes spectra of another substance, Caffeine, used at the same pharmaceutical factory were also measured. Pure Caffeine is a plant-based alkaloid which is applied to enhance the heart function in the way similar to Taurine. Spectra of 5 samples of Caffeine comprise data set C.

4. Results and discussion

4.1. *Explorative analysis*

The explorative PCA analysis of set A shows an essential difference between the samples. More than 60 objects (out of 246) may be treated as doubtless outliers. The source of such variations was found after comparing the spectra of the substance in a bag (spectra A1, A2 in Fig. 2), the spectra of unpacked substance (S1), and the spectra of empty polyethylene bags (P). The addition of the PE spectrum (P) to the substance spectrum (S1) may cause the distortion of the main peaks. This is the case for spectrum A1 but not for spectrum A2. The PE peak in the central part of the NIR region (first overtone around 5770 cm^{-1}) significantly shifts the substance peak to the left (in MIR direction). Second PE peak (combination bands around 4300 cm^{-1}) amplifies the corresponding substance peak. Due to varying thickness of PE bags caused by folds, the influence of the PE spectra on the routine measurements results in different distortions of the main substance peaks (compare spectra A1 and A2). A preliminary PCA analysis, namely the first loading vector plot, confirms this conclusion. The influence of packaging can partly be decreased by a proper choice of spectral region. In the present study we use the region of 4400–7400 cm^{-1}. It is worthy of mentioning that application of the averaged spectra instead of three repeated measurements of one drum does not improve the situation appreciably. The contrivance that helps to solve the problem is the construction of two

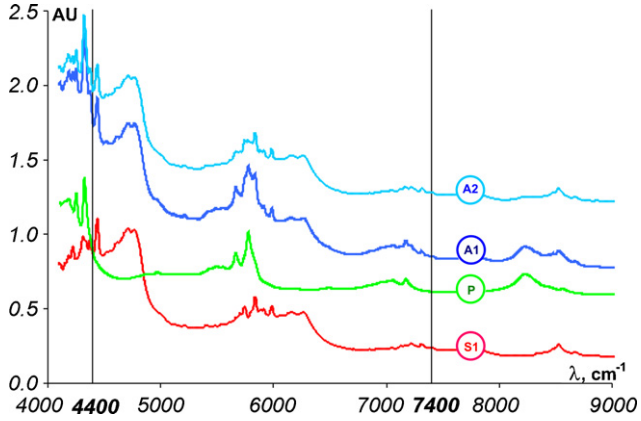

Fig. 2 – Spectrum S1 obtained from sample 1 without PE bag (substance), P is a spectrum of PE bag, A1 is a spectrum of sample 1 in PE bag, A2 is a spectrum of sample 2 in PE bag. Spectral region between vertical lines is used for further data processing.

PCA models and the application of sequential discrimination.

4.2. *Models for pattern recognition*

It is known that all measurements have been performed on perfect samples; therefore, the main idea is to keep the outliers and to treat them as a special calibration set. The 'outliers set' presents the PCA model that describes objects highly deviated from average objects but still presenting the spectra of satisfactory quality substances. In our case, all regular spectra of 184 samples comprise Group 1. The corresponding PCA (Model 1) employs three PCs to explain 98% of variance. The main variation is referred to substance peaks directly. The measurements attributed to outliers are collected in Group 2 (62 spectra). This group spans over rather a wide area in the scores plots.

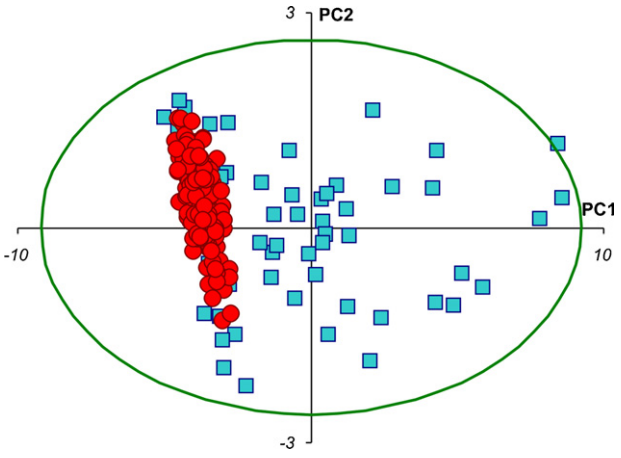

Fig. 3 – Score plot (PC1 vs. PC2) for Model 2. Squares are calibration samples (Group 2) and dots are test samples (Group 1).

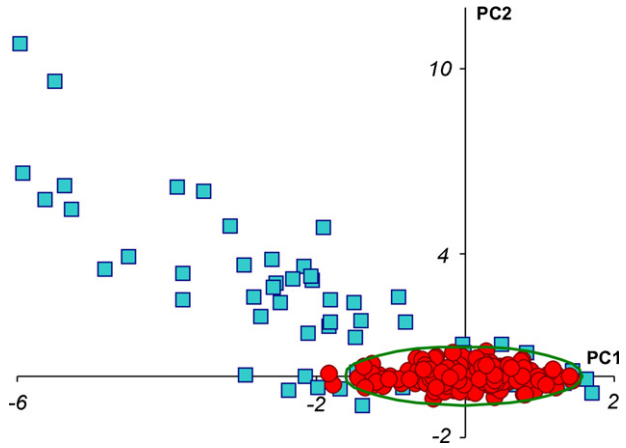

Fig. 4 – Score plot (PC1 vs. PC2) for Model 1. Dots are calibration samples (Group 1) and squares are test samples (Group 2).

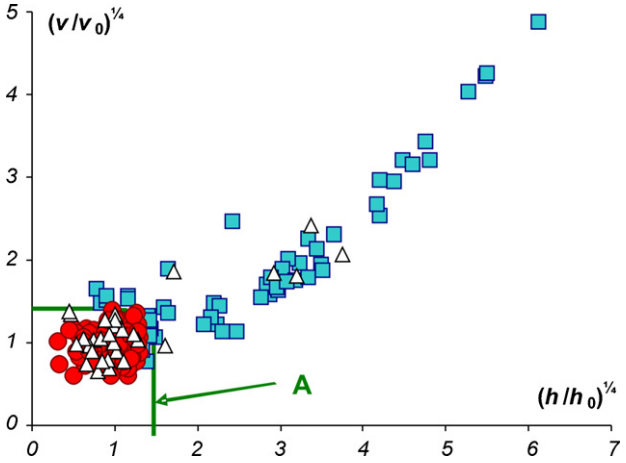

Fig. 5 – Classification plots for Model 1. Axes represent normalized score (*h*) and orthogonal (*v*) distances. Dots are calibration samples from Set 1 (Group 1), squares are samples from Set 2 (Group 2), triangles are validation Set 3 samples (Groups 1 and 2). Acceptance area is limited with curve A.

The corresponding PCA Model 2 with two PCs explains 99% of total variance. The main variation in Model 2 is referred to the impact of PE peaks. Being projected onto the PC hyperplaine of Model 2, the samples from Group 1 are located in a compact cluster inside Hotteling T^2 Ellipse (Fig. 3). Vice versa, if Group 2 is projected onto the Model 1 hyperspace, many samples from Group 2 are located beyond the Hotteling T^2 area (Fig. 4).

4.3. *Classification*

For classification and validation purposes the data are split into three sets. Set 1 (160 samples from Group 1) is used for the construction of the model for perfect measurements (Model 1). Set 2 (56 samples from Group 2) is used to build the 'outliers model' (Model 2). Set 3 is used for validation and consists

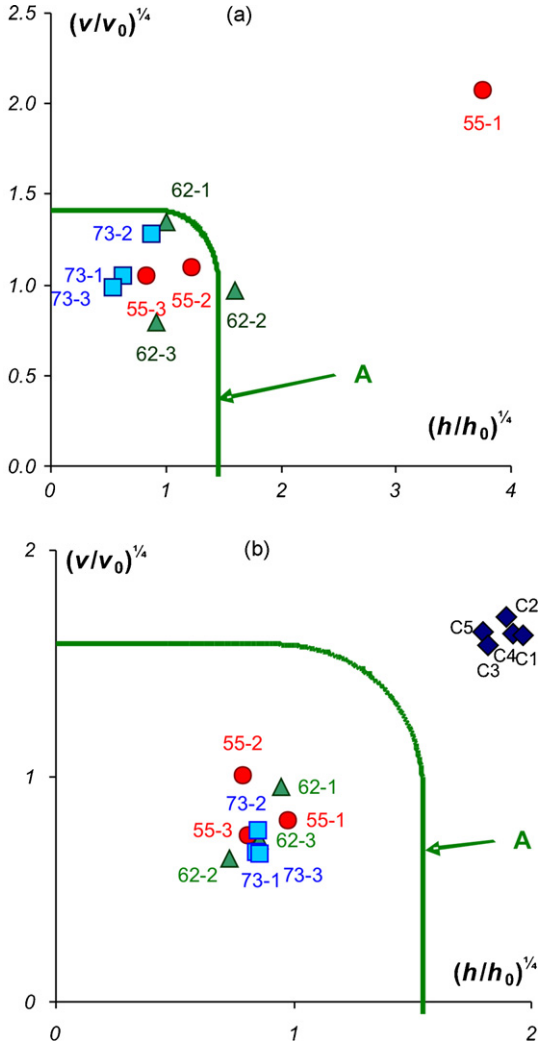

Fig. 6 – Classification of samples from the three drums with three replicate measurements. Axes represent the normalized score and orthogonal distances. Acceptance area is limited by curve A (a) Model 1 and (b) Model 2.

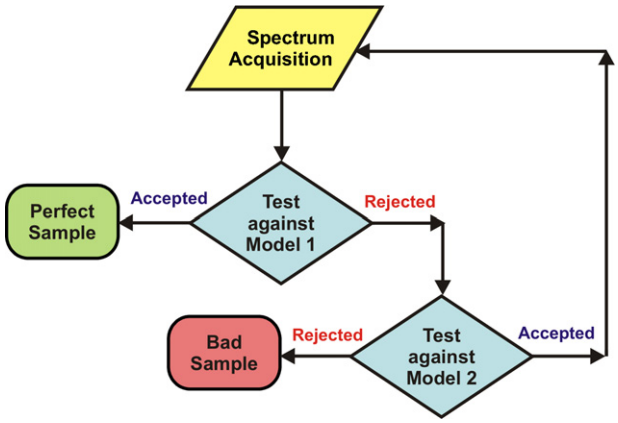

Fig. 7 – The flowchart of the sample routine testing.

of 24 samples from Group 1, and 6 samples from Group 2. The results of testing of Set 2 and Set 3 against Model 1 are presented in Fig. 5, which is a typical SIMCA plot adjusted in accordance with the above-described method. The axes represent the scores, h (aka, leverage) and orthogonal, v (aka, sample residual variance) distances. They are normalized to the corresponding mean values (h_0 and v_0). To make the plot more visual the fourth root is applied to both axes. The acceptance area is limited by curve A. It is constructed for Type I error $\gamma = 0.01$. As expected, all samples from Group 2 (Set 2 and 6 samples from Set 3) are located beyond the classification limits. To avoid the rejection of such measurements the classification against Model 2 is employed. The results of the two-stage classification for three different drums are presented in Fig. 6 (the first two digits of point label correspond to a drum number; the third digit corresponds to a replica number). There are no problems with the classification of drum 73. All three measurements are recognized by Model 1. For drum 55, first sample (55-1) is rejected by Model 1 (Fig. 6a), but recognized by Model 2 (Fig. 6b), resulting in the decision "extra measurements are required". The repeated measurements 55-2 and 55-3 identify the substance in drum 55 as of satisfactory quality. The same holds for drum 62.

It is also important to verify that Model 2 is not "too wide", i.e. that the samples of other substances are classified by Model 2 as aliens. In Fig. 6b samples of Caffeine (C1–C5) are marked by diamonds. One can see that they are located beyond the acceptance area. Both high OD and SD determine the location of these samples.

To summarize, the routine testing is conducted as follows:

- if a sample belongs to Class 1, this is a sample of a satisfactory quality (decision "accepted");
- if a sample belongs to Class 2, measurement should be repeated (no decision);
- if a sample does not belong to Class 2, such a sample is an alien (decision "rejected").

The flowchart of the routine testing is shown in Fig. 7.

It is important to be sure that such a sequential procedure will end with a definite decision (accepted or rejected) after a reasonable number of iterations. This totally depends on the share of the outliers (Group 2) in the whole population of objects. In our case this ration is $p = 62/184 \approx 0.25$. Therefore, the average number of repetitions is $1/p - 1 \approx 3$. The same could be formulated differently: a chance that more than 5 measurements of the same sample will be necessary is lower than 0.001.

5. Conclusions

Ordinary outliers are excluded from calibration. Nevertheless there are situations when outliers of a special type can dramatically improve the data analysis. In case we know that all measurements come from samples with satisfactory quality, a special model for outliers helps to make substance classification more reliable and decrease the errors caused by the measurements through closed bags and application of fiber optic sampling.

REFERENCES

[1] D.A. Burns, E.W. Ciurczak, Handbook of Near-Infrared Analysis, 3rd ed., CRC Press, 2001.
[2] K. Pöllänen, A. Häkkinen, M. Huhtanen, S.-P. Reinikainen, M. Karjalainen, J. Rantanen, M. Louhi-Kultanen, L. Nyström, Anal. Chim. Acta 544 (2005) 108.
[3] M. Kim, H. Chung, Y. Woo, M.S. Kemper, A new non-invasive, quantitative Raman technique for the determination of an active ingredient in pharmaceutical liquids by direct measurement through a plastic bottle, Anal. Chim. Acta 587 (2007) 200.
[4] A. Pulido, I. Ruisánchez, R. Boqué, F.X. Rius, Anal. Chim. Acta 455 (2002) 267.
[5] A.L. Pomerantsev, Acceptance areas for multivariate classification derived by projection methods, J. Chem., doi:10.1002/cem.1147, in press.
[6] S. Wold, Pattern Recognit. 8 (1976) 127.
[7] G.R. Flåten, B. Grung, O.M. Kvalheim, Chemom. Intell. Lab. Syst. 72 (2004) 101.
[8] A. Wald, Ann. Math. Stat. 16 (1945) 117.
[9] A. Höskuldsson, Prediction Methods in Science and Technology, vol. 1, Thor Publishing, Copenhagen, Denmark, 1996.
[10] T. Næs, T. Isaksson, T. Fearn, T. Davies, Multivariate Calibration and Classification, NIR Publications, 2002.
[11] G.R. Terrell, Biometrika 90 (2003) 445.