MULTIVARIATE CALIBRATION OF THERMOCHROMATOGRAPHIC DATA

Arkadi EBBER, Kaire IVASK, and Mihkel KOEL

Eesti Teaduste Akadeemia Keemia Instituut (Institute of Chemistry, Estonian Academy of Sciences), Akadeemia tee 15, EE-0026 Tallinn, Eesti (Estonia)

Abstract. The thermochromatograms of blends composed of polychloroprene rubber and polyacrylonitrile-butadiene rubber were measured. The three-dimensional thermochromatographic data were unfolded and multivariate calibration methods (PCR, PLSR) were applied to develop models for the prediction of ingredient content in polymeric composition.

Key words: multivariate calibration, thermochromatography.

Multivariate calibration was initially designed for calibrating some property of specimens with a multitude of known or easily measurable properties [1]. Very often the calibrated property is the amount of an analyt in the sample, which is calibrated with e.g. spectral or chromatographic data, i.e. with a vector of data.

With the application of "hyphenated" methods, a matrix of data can be obtained for each specimen. These data have to be correlated with the properties of interest, which should be predicted in the future [2]. Thermochromato-graphy belongs to the second-order methods that give a two-dimensional response for each sample. In thermochromatography a small amount of the sample is heated at a linearly rising temperature. After every given time interval the gas phase, containing substances evolving from the sample, is rapidly GC-analysed. The data collected from multiple chromatograms form a thermochromatogram that contains much information on the substances present in the sample and processes occurring in the course of heating.

The purpose of this report is to show that thermochromatographic data may be used for the calibration and prediction of the contents of ingredients in polymeric composition.

EXPERIMENTAL

Experiments were carried out in the home-built thermochromato-graphic system, described in [3]. A sample of 0.5–1 mg was heated at a rate 10°C/min. Starting from 90°C fast chromatograms of the gas phase were taken in every 57 s. So, the last 37th chromatogram was run at a sample temperature of 432°C and the experiment lasted 34.2 min. GC analysis was performed in a 15 m long capillary column with Carbowax 20M at 100°C and pure nitrogen (oxygen should be avoided!) was used as the carrier gas. For each chromatogram 256 digital points were collected. The sampling device, heating, and data acquisition were controlled by a computer. The specimens taken as the calibration set were polychloroprene rubber (CR), polyacrylonitrile-butadiene rubber (NR), and their blends with various ratios of CR to NR. All specimens contained also 32.5% of additives such as carbon black, ZnO, stearic acid, and some others, but throughout this paper we mean the relative percentage of CR and NR (i.e., CR percentage plus NR percentage is equal to 100%). As an example, Fig. 1 shows a thermochromatogram of a polymeric blend consisting of 80% of CR and 20% of NR. The reproducibility of thermochromatographic data was between 5–10%.
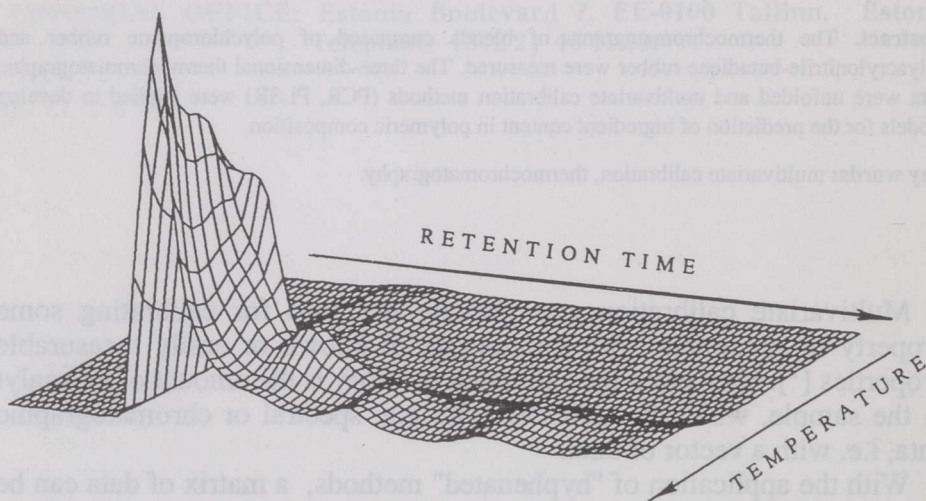

Fig. 1. Thermochromatogram of polymeric composition composed of polychloroprene (80%) and polyacrylonitrile-butadiene (20%).

CALIBRATION

A collection of stacked thermochromatograms forms a third-order tensor and, in principle, trilinear decomposition of data should be applied [4]. However, as our task was not spectra resolution, we could apply conventional multivariate calibration techniques: principal component regression (PCR) and partial least squares regression (PLSR). To reduce the dimensionality, the data were unfolded, i.e. instead of each two-

dimensional thermochromatogram we got a long vector. We did not use the whole thermochromatogram, but cut out the area where notable changes occurred: temperature range 180.5–432°C (27 points) and retention time interval 5.7–17.7 s (56 readings). So, we had for calibration a 12×1 column vector **Y** of the CR percentage in the samples and a 12×1008 matrix **X** of unfolded thermochromatographic data that were mean-centred prior to processing.

In order to model data by PCR, 10 principal components were calculated from matrix **X** using NIPALS algorithm [5] and then the linear models relating **Y** to various numbers of principal components were compared. Since the reproducibility of our experimental data was not better than 5%, there was no sense in taking too many factors and describe the noise. A model with four factors was sufficient to describe data without substantial overfitting and with acceptable accuracy.

In the case of PCR, maximum variation is extracted from **X** by each principal component. In PLSR, each factor takes into consideration from **X** only relevant information covariating linearly with **Y**. So, PLSR should model the data with fewer factors than PCR. The results of the description of **Y**-data by PCR and PLSR with two, three, and four factors are presented in Table 1. It may be concluded that at least three factors should be taken for proper calibration. This is in compliance with the results of Lochmüller et al. [6], who studied the blends and mechanical mixtures of CR and NR by thermogravimetry. They found that, unlike mechanical mixtures, in the case of blends CR and NR form a complex and three factors have to be used for the description of the blends.

Table 1

Percentage of polychloroprene in calibration samples and the values calculated by multivariate models

Polychloroprene in polymeric composition	Calculated with PCR		Calculated with PLSR	
	3 factors	4 factors	3 factors	4 factors
100.0	98.6	103.0	99.8	102.6
100.0	98.4	96.5	98.0	98.6
80.0	84.9	79.6	81.0	79.4
80.0	87.7	82.0	86.0	80.6
60.0	50.0	51.0	52.2	53.4
60.0	53.9	60.2	57.7	62.3
40.0	35.1	43.0	38.5	43.9
40.0	36.3	40.3	37.5	38.1
20.0	20.1	22.8	20.8	22.3
20.0	21.2	22.7	21.3	21.7
0.0	16.9	5.6	12.1	5.2
0.0	-3.3	-6.8	-5.9	-6.2
RMSE*	6.8	4.2	5.0	3.7

* root mean squared error.

3

In order to assess the ability of our models to predict the CR content in the future specimens, the cross validation procedure was used. In the cross validation, samples were left out one at a time and from the remaining data models with various numbers of factors were built up. Then the CR percentage was predicted for the left out sample and compared with the known figure. One experimental point seemed to be somewhat suspicious and we repeated the cross validation cycle leaving it out. As the result, the prediction error decreased as can be seen from Table 2. A plot of the **Y** values predicted in the cross validation versus CR percentage in the samples is given in Fig. 2.

Table 2

Root mean squared error of the prediction of polychloroprene percentage in the cross validation procedure

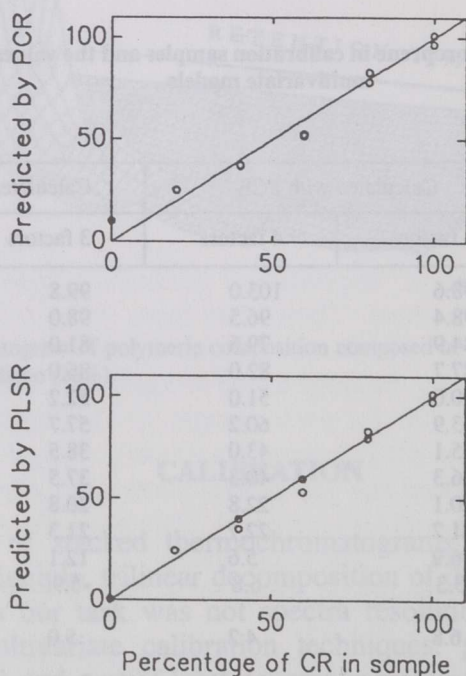| Number of factors in the model	Prediction by PCR		Prediction by PLSR	
	Number of samples			
	12	11	12	11
3	10.7	6.1	8.8	5.2
4	8.8	5.4	8.8	3.4
5	8.2	3.9	10.9	3.5
6	9.4	4.7	12.2	3.3

Fig. 2. Values predicted by cross validation with a four-factor model versus percentage of CR in 11 samples.

All the programs were written and the calculations performed in MATLAB (the MathWorks, Inc.) environment.

CONCLUSIONS

Unfolded thermochromatographic data may be used for constructing multivariate models to predict the ingredients content in polymeric composition. In the case of polychloroprene–polyacrilonitrile–butadiene blends the percentage of one component can be predicted with an accuracy around 5% in the range of 0–100% and at least three factors should be taken into account. Both PCR and PLSR methods can properly perform the calibration of thermochromatographic data, but priority should be given to PLSR. The experiments were carried out in a region of parameters not quite optimal: the temperature of samples should be increased more and the high temperature chromatograms should be run until the thermal degradation is finished. This would improve the calibration data and more sophisticated processing methods could be applied.

ACKNOWLEDGEMENT

REFERENCES

1. Martens, H., Naes, T. Multivariate Calibration. John Wiley & Sons, Chichester, 1991.
2. Geladi, P. Analysis of multi-way (multi-mode) data. – Chemometrics and Intelligent Laboratory Systems, 1989, **7**, 11–30.
3. Koel, M., Urov, K. Investigation of oil shales by thermochromatography. – Oil Shale, 1993, **10**, 4, 261–269.
4. Sanchez, E., Kowalski, B. Tensorial resolution: A direct trilinear decomposition. – J. Chemometrics, 1990, **4**, 29–45.
5. Wold, S. Principal component analysis. – Chemometrics and Intelligent Laboratory Systems, 1987, **2**, 37–52.
6. Lochmüller, C., Breiner, S., Koel, M., Eloma, M. Factor analysis of the thermogravimetry of rubber blends by singular value decomposition and variance minimization. – Anal. Chem., 1991, **63**, 1685–1691.

TERMOKROMATOGRAAFILISTE ANDMETE MITMEMÕÕTMELINE KALIIBRIMINE

Arkadi EBBER, Kaire IVASK, Mihkel KOEL

Mõõdeti mitmes vahekorras võetud kahest polümeerist (polü-kloropreenist ja polüakrüülnitriil-butadieenist) koosnevate polümeerse

kompositsiooni proovide termokromatogrammid ning kasutades mitmemõõtmelise kaliibrimise meetodeid (peakomponentide regressiooni ja partsiaalsete vähimruutude regressiooni) koostati lineaarsed mudelid, mis seovad ühe komponendi sisalduse proovis termokromatogrammis sisalduva infoga ja võimaldavad prognoosida ingrediendi sisaldust viieprotsendise täpsusega.

МНОГОМЕРНОЕ КАЛИБРОВАНИЕ ТЕРМОХРОМАТОГРАФИЧЕСКИХ ДАННЫХ

Аркадий ЭББЕР, Кайре ИВАСК, Михкель КОЭЛЬ

Измерены термохроматограммы полихлоропрена, полиакрилнитрилбутадиена и ряда их полимерных композиций с различными соотношениями составляющих компонентов. Трехмерная матрица термохроматографических данных послойно развернута в двухмерную матрицу, и содержание компонента в полимерной композиции скоррелировано с этими данными методами многомерного калибрования – PCR и PLSR. Полученные модели с четырьмя факторами характеризуют содержание полихлоропрена и полиакрилнитрилбутадиена с точностью около 5%, что соответствует воспроизводимости термохроматограмм.