Periodica Polytechnica
Chemical Engineering

59(4), pp. 277-282, 2015
DOI: 10.3311/PPch.7725

Comparison of Different Types of Near Infrared (NIR) Instruments in Ability to Measure Alkaloids in Capsule of Poppy *(P. somniferum)*

Szilveszter Gergely[1], Eszter Izsó[1*], András Salgó[1]

RESEARCH ARTICLE

Received 24 September 2014; accepted 20 December 2014

Abstract

The poppy capsule is one of the most important raw materials for the pharmaceuticals industry containing in about 25 different alkaloids. Among these there are important substances, such as morphine, with analgesic and anaesthetic properties, anti-tussive codeine, and noscapine with anti-tumor activity. My first objective was comparing dispersive and Fourier transformed (FT) near-infrared (NIR) instruments in order to measure alkaloids using the same sample population via mathematical pre-treatments (i.e. gap-segment derivatives) of the spectra and partial least squares (PLS) calibrations. The best PLS calibration using cross validation (CV) was for morphine based on derivative spectra of dispersive NIR with R-square (R^2) 0.924 and root mean square error (RMSE) 1.198 mg (g dry matter)$^{-1}$ in 1.24–20.44 mg (g dry matter)$^{-1}$ range. Comparison of the changes in the output of the PLS parameter for the two instruments revealed a variance depending on the derivative and the segment size.

Keywords

poppy capsule, alkaloids, NIR, PCA

[1] Department of Applied Biotechnology and Food Science, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, H-1111 Budapest, Szent Gellért tér 4, Hungary

[*] Corresponding author, e-mail: izso@mail.bme.hu

1 Introduction

Poppy is the one of our oldest cultivated plants. One of the most popular drugs is prepared from its milk sap, but due to its high opium content, coming from its milk sap too, poppy is mostly regarded as a drug and not as food. Chemically the most important part of poppy is poppy capsule, because it contains 25 different kinds of alkaloids. The main alkaloids are morphine, codeine, thebaine, papaverine and noscapine. Today these compounds are used for medical purposes (*e.g.* cough mixture, painkiller) [1]. Alkaloids are commercially extracted using a process invented by Kabay in 1925 [2,3]. Poppy straw, which is the dried head and stalk of *P. somniferum* is the principal source of both morphine and thebaine [3]. The morphine content is between 0.4–1.5 %; for noscapine it is 0.3–1.0 %; papaverine ranges between 0.3–1.0 %; while concentration for codeine and thebaine varies between 0.3–0.5 % [4]. Alkaloids are produced by the pharmaceutical industry where the active substance concentration level of the incoming raw materials is a key factor due to quality compliance. Quantification of alkaloid content of the starting plant material (poppy capsule) is carried out by high-performance liquid chromatography (HPLC) [5,6,7,8], but gas chromatography (GC) based techniques have shorter analysis time and better cost/effectiveness factor, according to a comparative study by Acevska et al. [9] Non-chromatographic methods such as vibrational spectroscopic techniques to determine the main important alkaloids are becoming more and more popular instead of the time-consuming and expensive chromatographic methods [10,11]. Mid-infrared and Raman spectroscopic methods for the simultaneous prediction of morphine, codeine, papaverine, thebaine and noscapine in poppy capsules, poppy milk as well as aqueous-ethanolic extract were developed by Schulz et al. [10] The authors highlighted the advantages of spectroscopic techniques (less than five minute measurement time compared to an hour for HPLC analyses; evaluation of poppy breeding material for quality control purposes in the food and pharmaceutical industry). Not only mid-infrared and Raman but also near-infrared spectroscopy was used for the determination of all alkaloids in line with the "green trend".

The square of correlation coefficients (R^2) between the second derivatives of NIR spectra and the reference (HPLC) data

were more than 0.93 in case of morphine, codeine, papaverine, thebaine and noscapine [11].

The aim of the present study was to develop calibrations for morphine, codeine, thebaine and noscapine based on spectra of dispersive and FT NIR spectrophotometers, and to compare the statistics data of calibration results coming from two different types of NIR instruments.

2 Experimental

2.1 Samples

75 grinded poppy capsule samples were provided by Alkaloida Chemical Company, Tiszavasvári, Hungary. 21,10 and 44 samples originated from the 2008, 2009 and 2010 crop years, respectively covering the genetic and environmental effects.

2.2 Reference measurements

The isolation of alkaloids from the poppy capsules was performed by applying ultrasonic extraction for 30 minutes using 1.0 g sample and 50 mL ammonium/methanol (1:25). HPLC separation of the alkaloids was carried out on an HP HPLC 3D ChemStation with the HP 1100 liquid chromatograph (Hewlett-Packard Co., Waldbronn, Germany) using Chromolith Performance RP-18 endcapped 100-4.6 reversed phase column (4.6×100 mm, Merck KGaA, Darmstadt, Germany). The mobile phase consisted of water [A: H_2O:trichloroacetic acid (99.8 %:0.2 %)], acetonitrile [B] and methanol [C]. Flow rate was 1 mL min^{-1}; UV-detection was set at 282 nm. Morphine sulphate, codeine base, papaverine hydrochloride, thebaine base and noscapine hydrochloride (Sigma-Aldrich Co., St. Louis, USA) were used as standard substances. For the quantification of the individual alkaloids the external standard method was applied.

2.3 Spectroscopic measurements

The three sample sets were scanned using two different instruments parallel to collecting the raw spectra. Three and two independent scans (*i.e.* two scans by 2008 and three scans by 2009 and 2010 were recorded for each sample and the means of these replicates were used in subsequent calculations) were recorded from each spectral sample from the 2008 and 2009, 2010 crop year, respectively. The dispersive NIR instrument, NIRSystems 6500 monochromator system (Foss-NIRSystems, Silver Spring, MD, USA) fitted with a Sample Transport Module (STM) and standard sample cups equipped with threaded back. Samples were scanned (32 scans co-added) from 1100 nm to 2498 nm in reflectance mode (R mode: PbS detector). Data were collected every 2 nm (700 data points per spectrum). The FT-NIR instrument, Spectrum 400 FT-IR/FT-NIR spectrometer (PerkinElmer, Waltham, MA, USA) fitted with a Near Infrared Reflectance Accessory (NIRA) and the same sample cups as described above. Samples were scanned (32 scans co-added) from 1000 nm to 2500 nm (from 10000 cm^{-1} to 4000 cm^{-1}) in reflectance mode (R mode: InGaAS detector). Data were collected every 1 cm^{-1} (3001 points per spectrum). The same loading of sample was consecutively scanned using dispersive NIR and FT-NIR instrument avoiding the pitfalls of different sample handling.

2.4 Data processing

Spectral and reference data were processed by using Vision 3.20 (Foss NIRSystems Inc., Silver Spring, MD, USA), Spectrum 10.00 (PerkinElmer, Waltham, MA, USA), Statistica 9.1 (StatSoft, Inc., Tulsa, OK, USA) and Unscrambler 10.0 (CAMO Software AS, Oslo, Norway) software packages.

2.5 Derivatives

Gap-segment and the Savitzky-Golay method were used for calculating the first and second order derivatives (D1OD and D2OD) of NIR and FT-NIR spectra, respectively. The first derivative is commonly used to eliminate baseline offset, while the second derivative eliminates both offset and slope within a set of spectra. The parameters of the gap-segment algorithm are a gap factor and a smoothing factor that are determined by the segment size and gap size chosen by the user [12,13]. If too large a segment is defined, one may decrease the resolution of the peaks, but too narrow a segment may generate noise in the derivative data. Raw NIR spectra were transformed into D1OD and D2OD using 2/0 nm (1/1 point), 4/0 nm (3/1 point), 8/0 nm (5/1 point), 12/0 nm (7/1 point), 16/0 nm (9/1 point) and 20/0 nm (11/1 point) segment and gap size, respectively by Vision 3.20 software.

The Savitzky-Golay algorithm is based on performing a least squares linear regression fit of a polynomial around each point in the spectrum to smoothen the data [14]. The derivative is then the derivative of the fitted polynomial at each point. The algorithm includes a smoothing factor that determines how many adjacent variables will be used to estimate the polynomial approximation of the curve segment. Raw FT-NIR spectra were transformed into D1OD and D2OD using 5, 9, 13, 25, 37, 49 and 149 point smoothing factor, respectively by Spectrum 10.00 software.

Derivatives were calculated by software controlling spectrometers assuming the manufacturers built the most fitting algorithms for pre-processing the raw spectra in their software. Comparing the number of points using for smoothing between derivatives applied for NIR and FT-NIR spectra (*e.g.* 1, 3, 5, 7, 9, 11 and 5, 9, 13, 25, 37, 49, 149, respectively) shows differences, because the applied dispersive NIR and FT-NIR spectrometers have a distinct wavelength selection technology resulting in different number of absorbance data vs. nm and cm^{-1} scaling, respectively. The ratio of the number of points per spectrum projecting to 1000–2500 nm in case of NIR vs. FT-NIR spectra is 751:3001, nearly 1:4. Distance between data points in the case of FT-IR spectra [calculating with reciprocal function between wavelength (nm) and wavenumber (cm^{-1})] increases from 0.20 nm to 1.25 nm along the spectra from 1000 nm to 2500 nm.

2.6 Principal component analyses (PCA)

PCA is a projection method that provides an interpretable overview of the main information contained in a multidimensional table. It takes the information carried by the original variables and projects them onto a smaller number of latent variables called principal components (PC) [15,16]. By plotting PCs important sample and variable interrelationships can be revealed, leading to the interpretation of certain sample groupings, similarities or differences.

2.7 Non-linear iterative partial least squares (NIPALS)

Partial least squares (PLS) model both the spectroscopic (X) and reference (Y) matrices simultaneously to find the latent (or hidden) variables in X that will best predict the latent variables in Y. These PLS components are similar to principal components, but will be referred to as factors. One of the most common algorithm used in PLS is non-linear iterative partial least squares (NIPALS). It is useful when there are missing values and when only the first few factors of a large data set need to be calculated [17]. The methods provided for the validation of PLS models were full cross validations, also known as leave-one-out cross validation. This produces as many calibration submodels as many samples there are in the data set. The maximum number of PLS factors was 25 and the optimal number of factors in the model was determined by the software. R-square (R^2), root mean square error (RMSE) and residual predictive deviation (RPD) were calculated for calibration (C) and cross validation (CV).

3 Results and discussion

Basic statistics of reference data obtained of dried homogenised poppy capsules are shown in Table 1. Checking the crop year effect (including genetic [G] and environmental [E] effects with G×E interaction) a PCA was developed for raw NIR and FT-NIR instruments. The first three principal components (PCs) described 99.98 % and 99.95 % of total variance of spectroscopic data in the case of NIR and FT-NIR spectrometers, respectively. Comparing the scatter plots of score values (Fig. 1) shows distinct groups defined by crop year. Using ellipses contouring the groups (that the length of its horizontal and vertical projection onto the x- and y-axis, respectively is equal to the mean ± range, where the mean and range refer to the X or Y variable) results fully separated ellipsis in case of PC 2 vs. PC 3 of FT-NIR instrument [Fig. 1(f)]. In order to separate the overlapping of peaks in the raw spectra transformation procedures were used. First and second derivatives of spectra with different segment size were compared to optimize the PLS output parameters. The advantage of derivation is that the signals of low intensity peaks are also emphasized however the noise peaks can be intensified, too. Mathematical treatment of the spectra should be checked that some low segment values do not result noisy derivative spectra reducing the performance indicator of calibrations. In the case of too high segment values loss of information can be possible because of the "oversmoothing" effect. The number of independent samples are limiting the number of factors which were required in the calibration model. Below the optimal factor number, the models are not precise enough while the overfitting can cause a deficiency in the robustness of the model.

Table 1 Basic statistics of reference data performed at dried homogenised poppy capsules [* results are expressed inmg (g dry matter)$^{-1}$; n.d. = no data]

Morphine					
Crop Year	N	Minimum*	Maximum*	Mean*	Std. Dev.*
2008	21	1.24	12.28	7.28	3.35
2009	10	6.16	9.43	7.88	1.12
2010	44	1.40	20.44	9.97	4.86
Total	75	1.24	20.44	8.94	4.30

Codeine					
Crop Year	N	Minimum*	Maximum*	Mean*	Std. Dev.*
2008	21	0.36	2.12	1.17	0.42
2009	10	0.74	1.23	1.00	0.19
2010	44	0.15	3.34	1.19	0.63
Total	75	0.15	3.34	1.16	0.54

Thebaine					
Crop Year	N	Minimum*	Maximum*	Mean*	Std. Dev.*
2008	21	0.25	1.52	0.83	0.33
2009	10	0.60	1.18	0.83	0.17
2010	44	0.36	6.24	1.37	1.27
Total	75	0.25	6.24	1.15	1.02

Papaverine					
Crop Year	N	Minimum*	Maximum*	Mean*	Std. Dev.*
2008	n.d.	n.d.	n.d.	n.d.	n.d.
2009	n.d.	n.d.	n.d.	n.d.	n.d.
2010	24	0.07	2.23	0.54	0.45
Total	24	0.07	2.23	0.54	0.45

Noscapine					
Crop Year	N	Minimum*	Maximum*	Mean*	Std. Dev.*
2008	14	0.14	28.00	4.72	8.54
2009	n.d.	n.d.	n.d.	n.d.	n.d.
2010	29	0.03	18.16	1.74	4.50
Total	43	0.03	28.00	2.71	6.17

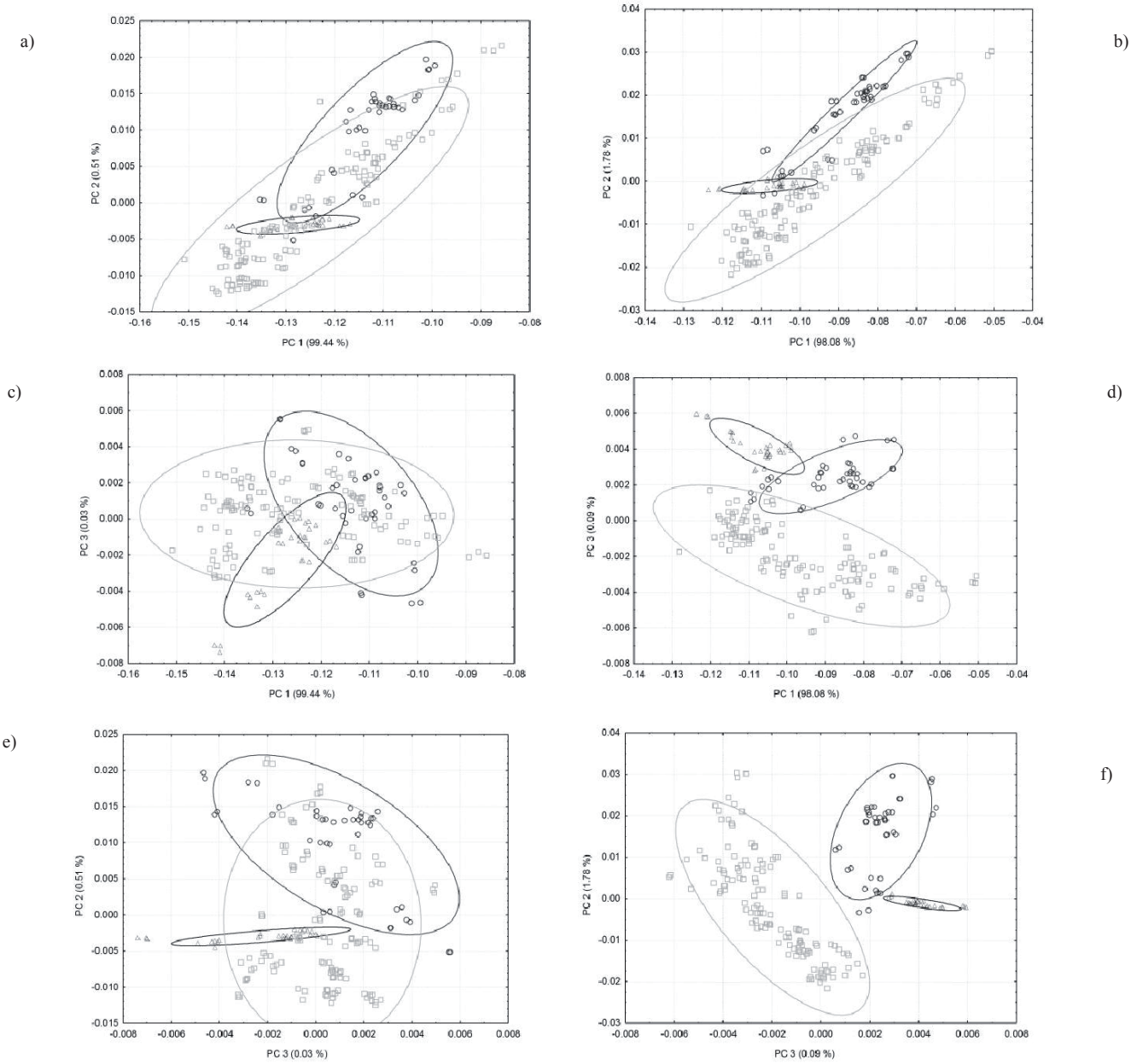

Fig. 1 Scatter plots of PCA scores. (a, c, e) dispersive NIR spectrometer (b, d, f) FT-NIR spectrometer ○ = samples from 2008 crop year, △ = samples from 2009 crop year, □ = samples from 2010 crop year

Figure 2(a) shows that the R^2 values were only light or no influenced by the size of derivation segment using dispersive NIR instrument. The optimized factor values chosen from second derivatives were lowered compared to first derivatives and the extent of RMSE(CV) values using second derivatives [Fig. 2(c)] were smaller compared to errors predicted from first derivatives. In the case of FT-NIR method from the second derivative calculated R^2 values (using factor number 8-9) [Fig. 2(b)] were smaller compared to first derivatives, while the RMSE(CV) values [Fig. 2(d)] were slightly higher (using similar factor number) compared to data taken with dispersive NIR instrument.

It was pointed out that spectra taken with dispersive NIR spectrometer were less sensitive to different mathematical transformations and provided slightly better results concerning preciseness and robustness compared to FT-NIR. If the comparison of the instruments had been carried out by using raw spectra [first columns of Fig. 2(a-d)], there were no significant differences between dispersive NIR and FT-NIR spectrometer in accuracy and R^2 values. Both order of derivative (*i.e.* first or second derivatives) and the segment size had different effect to PLS output parameters. The observed trend in PLS outputs (*i.e.* R^2, RMSE) could explain with different signal-to-noise ratio of dispersive NIR and FT-NIR raw spectra. It could attribute to 1) distinct wavelength selection technology (*i.e.* dispersive vs. Fourier transform), 2) material and so sensitivity of detectors (PbS detectors vs. InGaAs detector), 3) location of detectors (at angles of 45° to sample surface vs. integrated sphere).

The best performance indicators of PLS calibrations by alkaloids were summarized in Table 2. It is obvious that calibrations

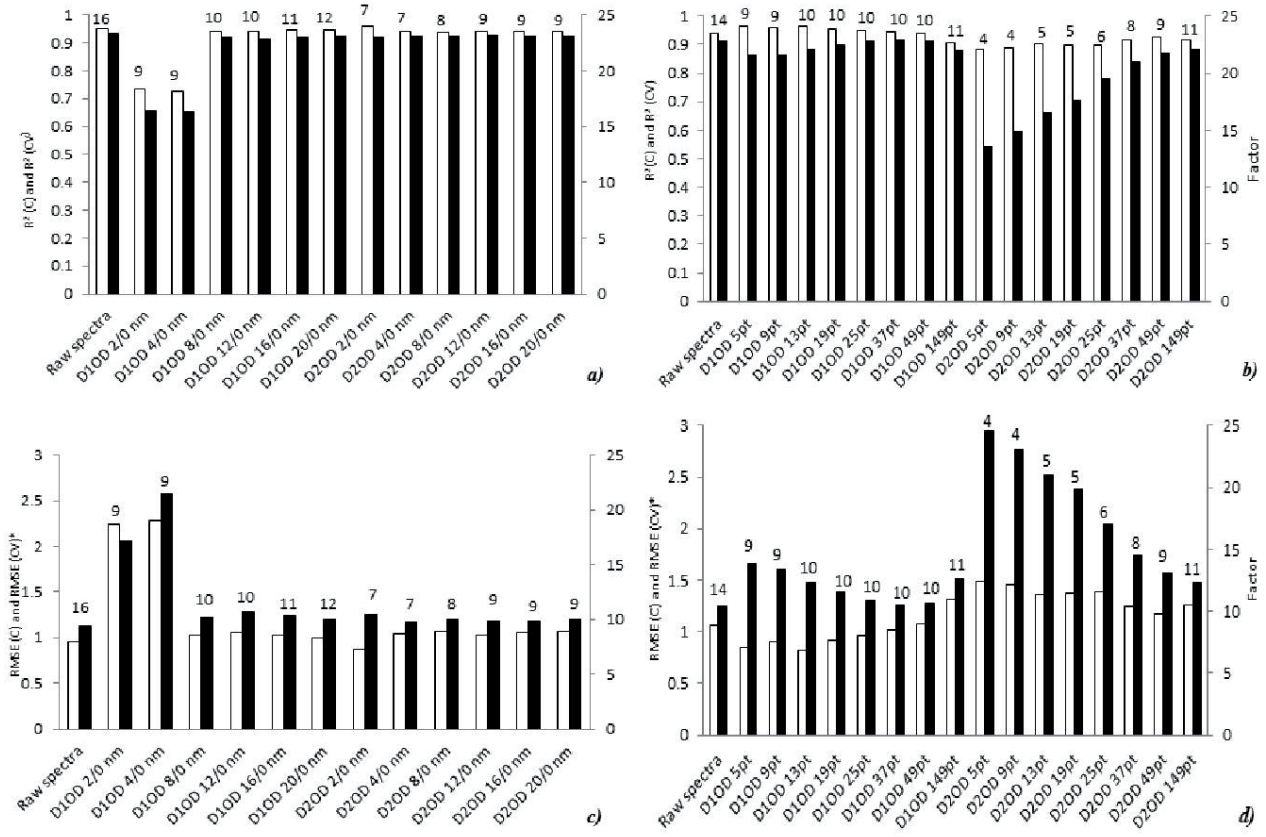

Fig. 2 PLS output parameters vs. mathematical pre-treatments in case of calibration of morphine. (a, c) dispersive NIR spectrometer, (b, d) FT-NIR spectrometer (a, b) □=R2(C), ■=R2(CV), on the top of the columns are showed the optimal number of factors (c, d) □=RMSE(C), ■=RMSE(CV), on the top of the columns are showed the optimal number of factors [* results are expressed in mg (g dry matter)–1]

Table 2 Values of performance indicators in case of dispersive NIR and FT-NIR spectrometers [* results are expressed in mg (g dry matter)–1]

Dispersive NIR				
Alkaloids	Mathematical pre-treatment	R^2(CV)	RMSE(CV)*	RPD(CV)
Morphine	D2OD 8/0 nm	0.924	1.198	3.621
Codeine	D2OD 8/0 nm	0.857	0.208	2.628
Thebaine	D1OD 4/0 nm	0.545	0.737	1.437
Papaverine	D1OD 4/0 nm	0.299	4.865	1.185
Noscapine	D2OD 16/0 nm	0.688	0.249	1.803

FT- NIR				
Alkaloids	Mathematical pre-treatment	R^2(CV)	RMSE(CV)*	RPD(CV)
Morphine	D1OD 149 pt	0.879	1.513	2.866
Codeine	D1OD 49 pt	0.741	0.279	1.965
Thebaine	D2OD 19 pt	0.366	0.849	1.249
Papaverine	D2OD 25 pt	0.871	2.081	2.771
Noscapine	D2OD 37 pt	0.468	0.327	1.373

for morphine show good results owing to the wide range of morphine content. PLS models for codeine and for papaverine (in case of FT-NIR only) are also acceptable, but we have noticed that the Pearson's correlation coefficients between morphine vs. codeine and morphine vs. papaverine are 0.734 and 0.773, respectively.

4 Conclusion

The most important advantage of the quick and non-destructive NIR technique is that the sample is measured with no or little pre-treatment (*e.g.* grinding), to facilitate the transfer. The main aim was not to replace the HPLC with NIR but was to develop a fast (quasi approximate) method to gain useful information on the incoming raw material.

Comparison of the dispersive NIR and the FT-NIR spectrometer using outputs of PLS calibrations were based on raw spectra, there were no significant differences but the effects of derivatives on these outputs were prominent. The cause of the differences can be explained by the variance in the signal-to-noise ratio of dispersive NIR and FT-NIR raw spectra and it could be an important fact for future projects tending to develop and/or transfer calibrations.

PLS calibrations based on the derivative spectra of dispersive NIR for morphine were the best with R^2(CV) = 0.924 and RMSE(CV) = 1.198 mg (g dry matter)$^{-1}$ in 1.24–20.44 mg

(g dry matter)$^{-1}$ range compared to the pervious study [11], where the R^2(CV) = 0.986 and RMSE(CV) = 2.429 mg (g dry matter)$^{-1}$ were in 0.01–11.08 mg (g dry matter)$^{-1}$ range concluded that the NIR spectroscopy is suitable for measuring of main alkaloid of the poppy seed capsule. Comparing the RMSE(CV) in our study, calculated with a significant lower value in spite of the whole sample population, covered 3 crop years.

References

[1] Pushpangadan, P., Singh, S. P. "*Handbook of Herbs and Spices*." Woodhead Publishing Ltd, Sawston, Cambridge, UK. 2001.

[2] Kabay, J. "Die Erzeugung des Morphins nach ungarischen Verfahren." *Pharmazeutische Monatshefte.* 11, p. 73. 1930. (in German)

[3] Bayer, I. "Manufacture of alkaloids from the poppy plant in Hungary." *Bulletin on Narcotics.* 13 (1). pp. 21-28. 1961. https://www.unodc.org/unodc/en/data-and-analysis/bulletin/bulletin_1961-01-01_1_page005.html

[4] Földesi, D. "The Poppy." *Journal of Oil Soap Cosmetics.* 46 (3). p. 93. 1997.

[5] Bosch, M. E., Sánchez, A. R., Rojas, F. S., Ojeda, C. B. "Morphine and its metabolites: Analytical methodologies for its determination." *Journal of Pharmaceutical and Biomedical Analysis.* 43 (3). pp. 799-815. 2007. DOI: 10.1016/j.jpba.2006.12.005

[6] Krenn, L., Boros, B., Ohmacht, R., Jelinek, L. "HPLC separation of opium alkaloids on porous and non-porous stationary phases." *Chromatographia.* 51 (1). pp. S175-S178. 2000. DOI: 10.1007/BF02492801

[7] Szabó, B., Lakatos, Á., Kőszegi, T., Botz, L. "HPTLC and HPLC determination of alkaloids in poppies subjected to stress." *JPC - Journal of Planar Chromatography - Modern TLC.* 16 (4). pp. 293-297. 2003. DOI: 10.1556/JPC.16.2003.4.9

[8] Acevska, J., Dimitrovska, A., Stefkov, G., Brezovska, K., Karapandzova, M., Kulevanova, S. "Development and Validation of a Reversed-Phase HPLC Method for Determination of Alkaloids from *Papaver somniferum* L. (Papaveraceae)." *Journal of AOAC International.* 95 (2). pp. 399-405. 2012. DOI: 10.5740/jaoacint.11-102

[9] Acevska, J., Stefkov, G., Petkovska, R., Kulevanova, S., Dimitrovska, A. "Chemometric approach for development, optimization, and validation of different chromatographic methods for separation of opium alkaloids." *Analytical and Bioanalytical Chemistry.* 403 (4). p. 1117-1129. 2012. DOI: 10.1007/s00216-012-5716-1

[10] Schulz, H., Baranska, M., Quilitzsch, R., Schütze, W. "Determination of alkaloids in capsules, milk and ethanolic extracts of poppy (*Papaver somniferum* L.) by ATR-FT-IR and FT-Raman spectroscopy." *Analyst.* 129 (10). pp. 917-920. 2004. DOI: 10.1039/B408930H

[11] Schulz, H., Pfeffer, S., Straka, P., Nothnagel, T. "Determination of valuable alkaloids in poppy capsules by near infrared reflectance spectroscopy." In: Near Infrared Spectroscopy: Proceedings of the 11th International Conference, Córdoba, Spain, 2003. pp. 883-886. ISBN: 0952866641

[12] Norris, K. H. "Extracting information from spectrophotometric curves. Predicting chemical composition from visible and near-infrared spectra." In: *Food Research and Data Analysis.* pp. 95-113. Applied Science Publishers Ltd, London, UK. 1983.

[13] Hopkins, D. W. "What is a Norris derivative?." *NIR News.* 12 (3). pp. 3-5. 2001. DOI: 10.1255/nirn.611

[14] Savitzky, A., Golay, M. J. E. "Smoothing and differentiation of data by simplified least squares procedures." *Analitical Chemistry.* 36 (8). pp. 1627-1639. 1964. DOI: 10.1021/ac60214a047

[15] Wold, S., Esbensen, K., Geladi, P. "Principal component analysis." *Chemometrics and Intelligent Laboratory Systems.* 2 (1–3). pp. 37-52. 1987. DOI: 10.169-7439/87/$03.50

[16] Martens, H., Næs, T. "Multivariate calibration." John Wiley & Sons Ltd, New Jersey, USA. 1991.

[17] Geladi, P., Kowalski, B. R. "Partial least-squares regression: a tutorial." *Analytica Chimica Acta.* 185. pp. 1-17. 1986. DOI: 10.1016/0003-2670(86)80028-9