Research Article

Received: 18 March 2013 Revised: 26 August 2013 Accepted article published: 1 October 2013 Published online in Wiley Online Library: 11 November 2013

(wileyonlinelibrary.com) DOI 10.1002/jsfa.6427

Prediction of essential oil content of oregano by hand-held and Fourier transform NIR spectroscopy

Cédric Camps,[a*] Marianne Gérard,[a,b] Mélanie Quennoz,[b] Cécile Brabant,[c] Carine Oberson[c] and Xavier Simonnet[b]

Abstract

BACKGROUND: In the framework of a breeding programme, the analysis of hundreds of oregano samples to determine their essential oil content (EOC) is time-consuming and expensive in terms of labour. Therefore developing a new method that is rapid, accurate and less expensive to use would be an asset to breeders. The aim of the present study was to develop a method based on near-inrared (NIR) spectroscopy to determine the EOC of oregano dried powder. Two spectroscopic approaches were compared, the first using a hand-held NIR device and the second a Fourier transform (FT) NIR spectrometer.

RESULTS: Hand-held NIR (1000–1800 nm) measurements and partial least squares regression allowed the determination of EOC with R^2 and SEP values of 0.58 and 0.81 mL per 100 g dry matter (DM) respectively. Measurements with FT-NIR (1000–2500 nm) allowed the determination of EOC with R^2 and SEP values of 0.91 and 0.68 mL per 100 g DM respectively. RPD, RER and RPIQ values for the model implemented with FT-NIR data were satisfactory for screening application, while those obtained with hand-held NIR data were below the level required to consider the model as enough accurate for screening application.

CONCLUSION: The FT-NIR approach allowed the development of an accurate model for EOC prediction. Although the hand-held NIR approach is promising, it needs additional development before it can be used in practice.

Keywords: hand-held NIR; FT-NIR; PLS; essential oil content (EOC); oregano

INTRODUCTION

Owing to its high demand, especially in the food industry, as dry grass or in the form of essential oil,[1] oregano has been the subject of numerous studies, including breeding programmes.[2,3] Although the chemical composition differs depending on the species and variety,[4–6] the trade name 'oregano' includes species that are rich in monoterpenoid phenols, mainly carvacrol and occasionally thymol.[5,6] It has been shown that the essential oil, which is rich in these molecules, has antimicrobial[7] and antioxidant[4,7] properties that can be used not only for the benefit of human health but also in the farming and food industries.[7]

Currently, qualitative and quantitative analyses of the components of oregano or its essential oil by conventional methods (i.e. hydrodistillation, gas chromatography, high-performance liquid chromatography) are time-consuming and expensive. In terms of speed of analysis, such methods are difficult to use when a series of hundreds of samples has to be analyzed, as in the case of a breeding programme.[8] In the last 20 years a predictive, rapid and low-cost method based on near-infrared spectroscopy (NIRS) has been developed for determining the quality of various agricultural and food products. Several studies have already been carried out successfully in the field of aromatic and medicinal plants.[9,10] Studies have been reported on cumin,[11,12] fennel,[11,13] coriander,[11] green tea leaves,[14] sage,[15] thyme[9] and rosemary.[16]

The aim of the present study was to develop a method to quantify the contents of oregano essential oil by NIRS. The method developed must be fully usable in the context of a breeding programme. Two technologies have been tested, a hand-held NIR device and a Fourier transform (FT) NIR spectrometer, both adapted to the needs of a breeding programme.

MATERIALS AND METHODS

Plant material

The oregano samples used in this study comprised species and varieties grown in the experimental fields of Agroscope Research Station (Conthey, Switzerland). Samples were gathered from two harvest years (2009 and 2010) and stored at room temperature in

* *Correspondence To: Cedric Camps, Agroscope Research Station, Research Department of Production and Plant Protection of Crops in Alpine Areas/Greenhouse Crops, Route des Vergers 18, CH-1964 Conthey, Switzerland. E-mail: cedric.camps@acw.admin.ch*

a *Agroscope Research Station, Research Department of Production and Plant Protection of Crops in Alpine Areas/Greenhouse Crops, Route des Vergers 18, CH-1964, Conthey, Switzerland*

b *Mediplant, Route des Vergers 18, CH-1964, Conthey, Switzerland*

c *Agroscope Research Station, Research Department of Arable Crop Plant Breeding and Genetic Resources, Route de Duillier 50, CP 1012, CH-1260, Nyon, Switzerland*

the dark. Several species were studied in order to have the widest range of essential oil content, namely *O. vulgare*, *O. minutiflorum*, *O. syriacum* (ssp. *syriacum* and ssp. *bevanii*), *O. vulgare* ssp. *hirtum* (variety of seed supplier Bolier), *O. vulgare* (var. Carva: *O. vulgare* ssp. *viridulum* × *O. vulgare* ssp. *hirtum*), *O. vulgare* ssp. *hirtum* (var. Origalia). A total of 101 samples were used in the present study.

Determination of essential oil content

Essential oils were obtained by hydrodistillation of samples of dried leaves according to the standard method.[17] All samples were distilled 1 week before analysis by NIRS, providing reference data necessary for the calibration step of the model. The dry matter (DM) content of samples was measured by drying at 105 °C for 12 h. The essential oil content (EOC) was expressed in mL per 100 g DM.

Hand-held NIR approach

Fractions of oregano dry samples were pulverized in a laboratory mill (10 000 rpm, 0.5 mm grid; Variable-speed Rotor Mill PULVERISETTE 14, Fritsch GmbH, Idar-Oberstein, Germany). The powder was carefully placed in closed vials and stored at room temperature in the dark. A first set of 74 samples (calibration set) and a second set of 27 samples (test set) were used for the calibration and validation steps respectively.

Spectra were acquired in reflectance mode using a MEMS-based PHAZIR (NIR PHAZIR 1018, Anatec, Eke, Belgium). Samples of oregano dry powder were placed in adapted vials closed with a plastic cap (vials for PHAZIR PCX-ACC-4 solids adapter, 15 mm i.d.). Spectral acquisition was carried out in direct contact analysis mode by placing the vials in the specific PHAZIR adapter situated at the end of the NIR pistol. Absorbance spectra (average of 30 scans) were recorded at a resolution of 8 nm from 1000 to 1800 nm. Before analyzing the set of samples, a white reference scan was carried out using a piece of Spectralon®.

Within this framework, NIR measurements were performed three times by rotating the vial a few degrees between each measurement. A total of 222 (3 measurements × 74 samples) spectra were collected to calibrate the model for the prediction of EOC, and 81 (3 measurements × 27 samples) spectra were collected to constitute the validation set. In order to compensate the effects of uncontrolled baseline and intensity variations, spectra were pretreated using a second-derivative method.[10]

FT-NIR approach

Spectra were acquired in reflectance mode using an FT-NIR spectrometer (NIRFlex Solids, Büchi, Flawil, Switzerland). Powder of dried leaves was presented to the instrument in a rotating glass Petri dish, and NIR spectra were collected from 1000 to 2500 nm at a resolution of 12 cm^{-1}.

NIR measurements were performed six times by rotating the Petri dish between each measurement. A total of 444 (6 measurements × 74 samples) spectra were collected to calibrate the model for the prediction of EOC, and 27 (1 measurement × 27 samples) spectra were collected to constitute the validation set. Spectra were pretreated by the standard normal variate (SNV) method[18] and detrending.

Data analysis

Partial least squares regression (PLS) was carried out to produce linear models of prediction between spectral data and reference values (EOC). The models were built in three steps, i.e. (1) calibration, (2) cross-validation and (3) validation. Cross-validation was performed using a leave-*k*-out procedure, were *k* is the number of spectral acquisitions per sample.[19] The optimal number of latent variables (LV) introduced in the models corresponded to a compromise that allowed us to obtain a model presenting on the one hand the relatively lowest RMSECV value and on the other hand the relatively highest R^2 value.[20]

The accuracy of the predictions is discussed according to the coefficient of determination of calibration (R^2) and the standard errors of calibration (SEC), cross-validation (SECV) and validation (SEP), while other calculated parameters allowed us to attempt an evaluation of model quality according to the range of reference data and the eventual bias measured when the external validation was performed:

$$R^2\,(\mathrm{C/CV/P}) = 1 - (\mathrm{PRESS/TSS})$$

$$\mathrm{SE\,(C/CV/P)} = \left[\sum_{i=1}^{n} (y_i - \hat{y}_i)^2/n\right]^{1/2}$$

$$\mathrm{bias} = \sum_{i=1}^{n} (\hat{y}_i/n) - \sum_{i\ =1}^{n} (y_i/n) = \bar{\hat{y}} - \bar{y}$$

$$\mathrm{SE\,(C/CV/P)_c} = \left[\sum_{i=1}^{n} (\hat{y}_i - \mathrm{bias} - y_i)^2/n\right]^{1/2}$$

$$\mathrm{RSE\,(C/CV/P)_c\,(\%)} = (100/\bar{y}) \left[\sum_{i\ =1}^{n} (\hat{y}_i - \mathrm{bias} - y_i)^2/n\right]^{1/2}$$

where $\hat{y}_i$ is the predicted value, y_i the mean value and y_i the actual value of EOC in the PLS model, n is the number of samples in the PLS model, PRESS is the prediction residual error of the sum of squares, TSS is the total sum of squares and the subscript 'c' indicates that the parameters (SE(C/CV/P) and RSE(C/CV/P)) have been corrected for bias.

The accuracy and robustness of the PLS models are discussed according to the following parameters, all corrected for bias value:

$$\text{coefficient of variation, } \mathrm{CV_c\,(\%)} = \mathrm{SEP_c/mean}$$

$$\text{ratio of performance to deviation, } \mathrm{RPD_c} = \mathrm{SD/SEP_c}$$

where SD is the standard deviation;

$$\text{ratio of } \mathrm{SEP_c} \text{ to reference data range, } \mathrm{RER_c} = (y_{\max} - y_{\min})/\mathrm{SEP_c}$$

where $y_{\max}$ and $y_{\min}$ are the maximum and minimum reference values of EOC respectively;

$$\text{ratio of } \mathrm{SEP_c} \text{to interquartile,}^{21}\ \mathrm{RPIQ_c} = (Q_3 - Q_1)/\mathrm{SEP_c}$$

where Q_3 and Q_1 are the values of the third and first quartiles of reference data respectively.

RESULTS

Reference data values

The EOC reference values obtained from hydrodistillation of the essential oils ranged from 0.23 to 10.1 mL per 100 g DM. The histogram of EOC values with superimposed normal density curve in

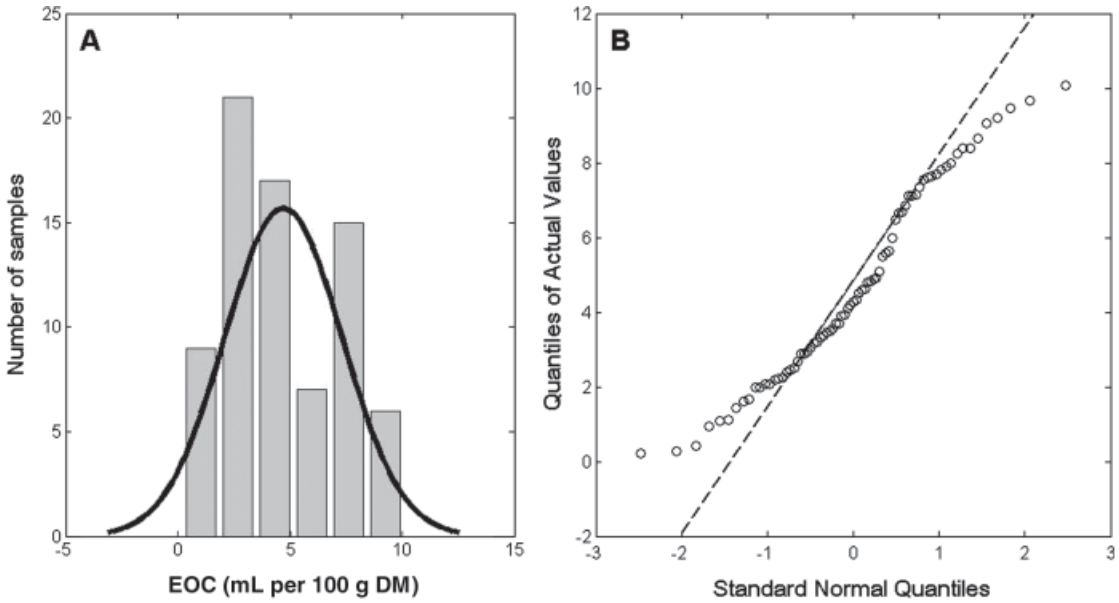

Figure 1. (A) Histogram of EOC values with superimposed normal density curve. (B) Quantile–quantile plot.

Table 1. PLS data of EOC determination (hand-held NIR and FT-NIR)

		PHAZIR 1018			FT-NIR		
PLS data	Unit	Calibration	Cross-validation	Validation	Calibration	Cross-validation	Validation
N	—	74	74	27	74	74	27
EOC range	mL per 100 g	0.23–10.1	0.23–10.1	0.4–8.1	0.23–10.1	0.23–10.1	0.4–8.1
EOC mean value	mL per 100 g	4.8	4.8	4.89	4.8	4.8	4.89
EOC SD	mL per 100 g	2.61	2.61	2.34	2.61	2.61	2.34
λ range	nm	939–1797	939–1797	939–1797	1000–2500	1000–2500	1000–2500
LV	—	3	3	3	6	6	6
R^2	—	0.92	0.92	0.58	0.93	0.94	0.91
SE(C/CV/P)	mL per 100 g	0.75	0.77	2.20	0.7	0.68	0.69
Bias	mL per 100 g	1.40×10^{-2}	1.55×10^{-2}	−2.04	4.5×10^{-7}	1.08×10^{-2}	8×10^{-2}
$SE(C/CV/P)_c$	mL per 100 g	0.75	0.77	0.81	0.7	0.68	0.68
$RSE(C/CV/P)_c$	Relative %	15	15	18	15	15	14
CV_c	Relative %	15	15	18	15	15	14
RPD_c	—	3.54	3.44	3.51 (1.30)[a]	3.7	3.82	3.24
$RPIQ_c$	—	6.19	6.01	5.03 (1.87)[a]	6.6	6.8	4.55
RER_c	—	13.17	12.78	9.45 (3.50)[a]	14.04	14.51	11.31
Spectral treatment		Golay second derivative (step 3)			SNV + detrending		

N, number of samples; EOC, essential oil content; SD, standard deviation; λ range, wavelength range of PLS model; LV, number of latent variables; R^2, determination coefficient; SE, standard error; RSE, relative standard error of prediction; CV, coefficient of variation; RPD, ratio of performance to deviation; RPIQ, ratio of performance to interquartile; RER, ratio of error to range; subscript 'c' (SE_c, RSE_c, CV_c, RPD_c, $RPIQ_c$ and RER_c), parameters calculated after bias correction; SNV, standard normal variate.
[a] Values in parentheses are RPD, RPIQ and RER before correction for bias.

Fig. 1A and the quantile–quantile plot in Fig. 1B illustrate the distribution of the EOC data set. The pattern of the quantile–quantile plot suggests a non-normal distribution of data, mainly at the extremities. Data analysis using Kolmogorov–Smirnov ($P = 1.03 \times 10^{-049}$) and Shapiro–Wilk ($P = 0.017$) tests at a threshold of 5% confirmed the non-normal distribution.

PLS model using hand-held NIRS

The PLS model data obtained using hand-held NIRS are reported in Table 1. Several parameters were calculated to evaluate the accuracy of the model and to measure the fit of the predicted values to the reference data. Figure 2 shows the actual *versus* predicted values for calibration (Fig. 2A) and validation (Fig. 2B). The calibration step was evaluated according to R^2 and SEC, for which values of 0.92 and 0.75 mL per 100 g DM respectively were obtained. The number of LV used was three, which is relatively small, thus avoiding potential overfitting of the model.[22] The cross-validation procedure presented a very small bias close to zero, while R^2 and SECV values were close to those obtained in calibration. Three parameters allowing us to evaluate the accuracy of the model according to the range of reference data were calculated, RPD_c, RER_c and $RPIQ_c$, whose values were 3.54, 13.17

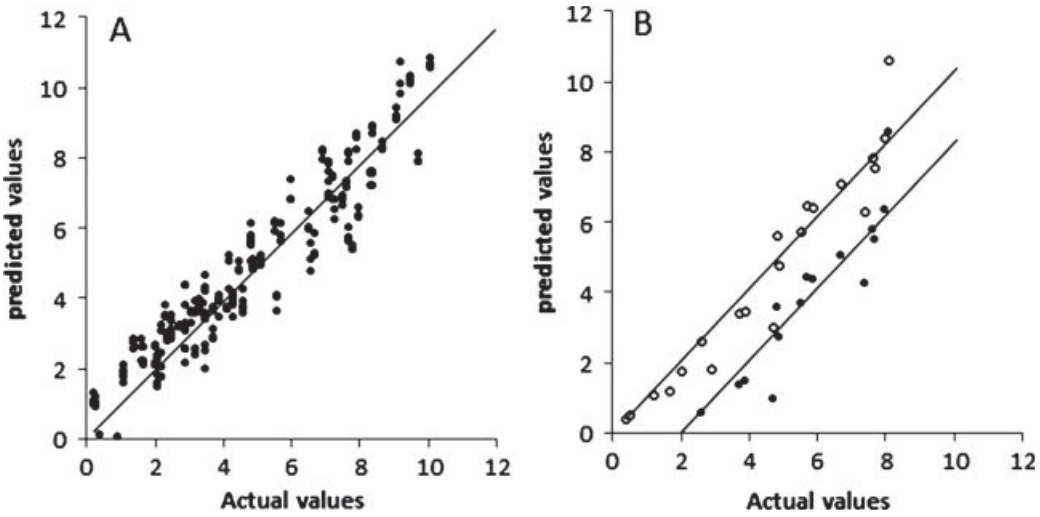

Figure 2. Actual *versus* predicted values of EOC using hand-held NIR device PHAZIR 1018: A, calibration; B, validation; ●, data without bias correction; ○, data corrected for bias value.

and 6.19 respectively for calibration. In cross-validation, RPD_c, RER_c and $RPIQ_c$ were 3.44, 12.78 and 6.01 respectively. For both calibration and cross-validation the RPD_c and RER_c values are greater than 3 and 10 respectively. This means that the model could be suitable for quantitative analysis.[23] RPIQ is used in the case of a non-normal distribution of the reference data set to standardize the SE value. RPIQ uses the inter-quartile ($IQ = Q_3 - Q_1$) parameter instead of SD to standardize the SE value. In the case of a non-normal distribution, IQ could be a better indicator of the data spread around the median. In the present study the RPIQ value indicates that the accuracy of the model is more than five times lower than the interquartile distance.

A validation step was performed with samples not used in the cross-validation model. The R^2 value obtained during validation ($R^2 = 0.58$) was lower than that obtained in cross-validation ($R^2 = 0.92$). An SEP value of 2.20 mL per 100 g DM was calculated, including a bias value of −2.04 mL per 100 g DM. Thus, after correcting for bias, the SEP_c value decreased to 0.81 mL per 100 g DM. Differently to R^2, the SEP_c value remained close to the SEC value obtained in the cross-validation step (0.77 mL per 100 g DM). RPD_c was greater than 3 (3.51), but RER_c decreased to 9.45. $RPIQ_c$ remained at a high level with a value of 5.03.

PLS model using FT-NIRS

The PLS model data obtained using FT-NIRS are reported in Table 1. Figure 3 shows the actual *versus* predicted values for calibration (Fig. 3A) and validation (Fig. 3B). The same parameters as for hand-held NIRS were calculated to evaluate the accuracy of the model and to measure the fit of the predicted values to the reference data. R^2 and SEC values of 0.93 and 0.7 mL per 100 g DM were obtained in the calibration step. Similar R^2 and SE values were obtained in cross-validation. The number of LV remained relatively small at six and the bias was negligible. RDP_c values of 3.7 and 3.82 and RER_c values of 14.04 and 14.51 were obtained in calibration and cross-validation respectively. In the validation step, R^2 and SEP values were at least equal to or better than those calculated in calibration. RPIQ values were 6.6 and 6.8 for calibration and cross-validation respectively and 4.55 in the validation step. No bias was measured during validation (0.08 mL per 100 g DM), in contrast to the model validation with the hand-held NIR device.

DISCUSSION

The aim of the present study was to evaluate the ability of a hand-held NIR device for determining the EOC of oregano dry powder and to compare its performance with that of an FT-NIR spectrometer commonly used in the laboratory. The advantages of a hand-held device are that it can be moved between several breeding sites and is cheaper than laboratory NIR equipment.

The results showed that the determination of EOC was possible with the FT-NIR spectrometer and promising with the hand-held NIR device. In terms of performance, chemometric analysis of the FT-NIR data allowed us to determine the EOC with an accuracy of about 0.70 mL per 100 g DM (cross-validation and validation). The R^2 value higher than 0.9 (calibration and validation) showed the good fit between reference and predicted data. Furthermore, no bias was measured in the model using FT-NIR data.

The model implemented using hand-held NIR data showed promising results but would not be usable in practice in its present state of development. Indeed, in terms of performance the model allowed a measurement accuracy of 0.77 mL per 100 g DM (calibration) and 0.81 mL per 100 g DM (validation), slightly lower than the FT-NIR results. The R^2 values, particularly in validation ($R^2 = 0.58$), showed that the reference and predicted values did not fit as well as with FT-NIR data. Moreover, a bias was measured in the validation step between reference and predicted values. This bias showed that the predicted values underestimated the reference values by about 2 mL per 100 g DM. The presence of such a bias is difficult to understand.

The bias is a systematic error that can have different origins in the chain of steps leading to the construction of a prediction model using NIRS data.

- The first two potential sources of error are a lack of reproducibility of the reference measurement and a lack of reproducibility of the sample preparation procedure between calibration and validation. In the present work the procedures were exactly the same when analysing samples of the calibration and validation steps.
- Another potential source of systematic error is a significant change in environmental conditions during spectral acquisition. In particular, variations in temperature and relative humidity could affect the quality of NIR spectra collected. In this study,

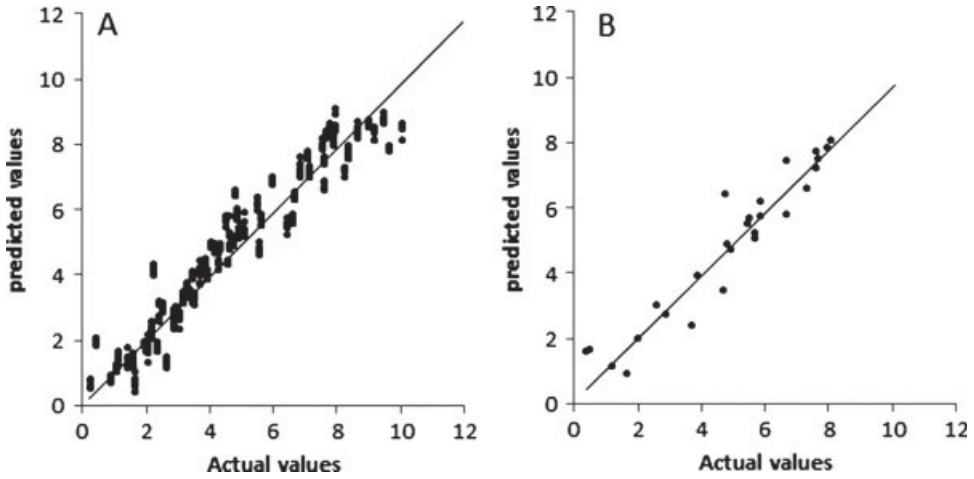

Figure 3. Actual *versus* predicted values of EOC using FT-NIR spectrometer: A, calibration; B, validation.

measures of calibration and validation were performed in the same laboratory under controlled environmental conditions.

- Also, a high genetic variability of plant samples could induce a systematic error between calibration and validation steps. In the present experiment, calibration was performed with samples of various genetic backgrounds (hybrids) to avoid or limit this effect. Storage of plant material for too long in unsuitable conditions could lead to chemical or physical deterioration, which may also introduce a bias. In the present study the spectral acquisitions of calibration and validation samples were carried out at different times, so it is possible that the samples for validation were altered. However, to confirm this hypothesis, a bias should be found in the model using the FT-NIR spectra, but this was not the case.
- A last potential source of bias is related to the manipulator itself. Indeed, a portable spectrometer requires working with the utmost rigour and perfect reproducibility. This parameter was not considered in this study and will be given special attention in future steps for the development of methods using a portable spectrometer.

Concerning the other calculated parameters, RPD, CV and RER are commonly used by NIR spectroscopists to evaluate their models and, more precisely, the error of the models as a function of the range values of the reference measure. To consider a model as 'correct' for 'plant screening', RPD and RER values have to be equal to or greater than 3 and 10 respectively.[23–25] RPD values of the model obtained with FT-NIR spectral data were higher than those of the model obtained with hand-held NIR spectral data. The RER value of the FT-NIR model reached 11.31 in validation, confirming the possible usability of such a device for screening samples. Concerning the model using hand-held NIR data, RPD and RER values were only satisfactory after correction for bias, which increased RPD from 1.30 to 3.51 and RER from 3.50 to 9.45. The relatively low RER value obtained after correction for bias (<10) confirms the relatively low R^2 value (0.58) and thus the non-usability of this model at its present stage of development.

RPIQ is a parameter allowing one to evaluate the spread of predicted *versus* reference data around the median in cases where the reference data distribution is non-normal (skewed distribution). Measurements of EOC in the present study followed a non-normal distribution as confirmed by Kolmogorov–Smirnov and Shapiro–Wilk tests. Thus RPIQ could be a more useful parameter than RPD to describe the obtained model of predictions. As stated above, RPIQ was higher for the model implemented with FT-NIR data (4.55) than for that implemented with hand-held NIR data before correction for bias (1.87). In the model using data from the hand-held NIR device, the RPIQ value of 1.87 means that the model error is less than two times smaller than the interquartile range of reference data, which is far from sufficient for good model performance. In contrast, the RPIQ value of 4.55 obtained with FT-NIR data means that the error of the model is less than four times smaller than the interquartile range of reference data. In this last case the performance of the model can be considered as good. Since the RPIQ parameter is a relatively recent index, no scale value has been published yet (contrary to the RPD parameter) allowing one to evaluate the prediction models.

In the present state of the model, with a 2 mL per 100 g DM bias measured during validation, the hand-held NIR device is not usable in practice. Additional samples of oregano allowing one to increase the variability of samples (various EOCs, geographical origins, cultivation practices, etc.) have to be collected to enrich the model and thus try to diminish the bias value.

CONCLUSION

The aim of this study was to investigate the potential of NIRS to facilitate the screening of hundreds or thousands samples of oregano, with particular emphasis on their EOC, in the context of a breeding programme.

Two approaches in terms of technology/device were attempted: hand-held NIRS and FT-NIRS. The approach using FT-NIR allowed the correct prediction of oregano EOC with an accuracy of 0.68 mL per 100 g DM. All parameters used to evaluate the performance of the model reached expected levels, indicating that the FT-NIR approach is suitable for good screening. Although the hand-held NIR approach is promising, the obtained results are not suitable for use in practice. The performance of the model ($SEP_c = 0.81$ mL per 100 g DM) is inferior to that obtained with the FT-NIR approach. Moreover, a bias correction of about 2 mL per 100 g DM had to be made to achieve an accuracy of 0.81 mL per 100 g DM. However, the RPIQ value calculated after bias correction is promising for future development of hand-held NIRS. In a next step the calibration data set has to be enriched with additional samples from different origins and different levels of EOC in order to minimize the bias value. The development of measurement

methods using portable tools must take into account the effect of the manipulator in order to minimize the systematic error between calibration and validation measurements.

REFERENCES

1 Olivier GW, The world market of oregano, in *Oregano. Proceedings of the IPGRI International Workshop on Oregano, 8–12 May 1996, CIHEAM, Valenzano, Bari, Italy*, ed. by Padulosi S. IPGRI, Rome, pp. 141–145 (1997).
2 Rey C, Carron CA, Bruttin B and Cottagnoud A, La variété d'origan 'Carva'. *Rev Suisse Vitic Arboric Hortic* **34**(2):I–VIII (2002).
3 Van Der Mheen H, Selection and production of oregano rich in essential oil and carvacrol. *Acta Hort* **709**:95–99 (2006).
4 Bernáth J, Some scientific and practical aspects of production and utilization of oregano in central Europe, in *Oregano. Proceedings of the IPGRI International Workshop on Oregano, 8–12 May 1996, CIHEAM, Valenzano, Bari, Italy*, ed. by Padulosi S. IPGRI, Rome, pp. 75–92 (1997).
5 Skoula M and Harborne JB, The taxonomy and chemistry of origanum, in Oregano. The Genera *Origanum* and *Lippia*, ed. by Kintzios SE. CRC Press, Boca Raton, FL, pp. 67–108 (2002).
6 Economou G, Panagopoulos G, Tarantilis P, Kalivas D, Kotoulas V, Travlos IS, *et al*, Variability in essential oil content and composition of *Origanum hirtum* L., *Origanum onites* L., *Coridothymus capitatus* (L.) and *Satureja thymbra* L. populations from the Greek island Ikaria. *Ind Crops Prod* **33**:236–241 (2011).
7 Zupancic A and Baricevic D, *Biological Activity of Oregano (Origanum vulgare ssp. vulgare)*. Slovensko Agronomsko Drustvo (SAD), Ljubljana (2002).
8 Carlen C, Breeding and cultivation of medicinal plants, in *Herbal Medicines. Development and Validation of Plant-derived Medicines for Human Health*, ed. by Bagetta G, Cosentino M, Corasaniti MT and Sakurada S. CRC Press, Boca Raton, FL, pp. 79–91 (2012).
9 Schulz H, Quilitzsch R and Kruger H, Rapid evaluation and quantitative analysis of thyme, oregano and chamomile essential oils by ATR-IR and NIR spectroscopy. *J Mol Struct* **661**:299–306 (2003).
10 Camps C, Toussirot M, Quennoz M and Simonnet X, Determination of artemisinin and moisture contents of *Artemisia annua* L. dry powder using hand-held near-infrared spectroscopy. *J Near Infrared Spectrosc* **19**:191–198 (2011).
11 Schulz H, Drews HH, Quilitzsch R and Krüger H, Application of near infrared spectroscopy for the quantification of quality parameters in selected vegetables and essential oil plants. *J Near Infrared Spectrosc* **6**:A125–A130 (1998).
12 Toxopeus H and Bouwmeester HJ, Improvement of caraway essential oil and carvone production in the Netherlands. *Ind Crops Prod* **1**:295–301 (1992).
13 Steuer B and Schulz H, Near-infrared analysis of fennel (*Foeniculum vulgare* Miller) on different spectrometers – basic considerations for a reliable network. *Phytochem Anal* **14**:285–289 (2003).
14 Schulz H, Engelhardt UH, Wegent A, Drews HH and Lapczynski S, Application of near-infrared reflectance spectroscopy to the simultaneous prediction of alkaloids and phenolic substances in green tea leaves. *J Agric Food Chem* **47**:5064–5067 (1999).
15 Elementi S, D'Antuono LF, Schulz H, Krüger H, Schütze W and Steuer B, *Salvia officinalis* L. essential oil and carnosic acid analysis by means of NIR spectroscopy. *Acta Hort* **723**:234–247 (2006).
16 Schulz H, Steuer B, Kruger H, Schutze W, Junghanns W and Weinreich B, Rapid determination of quality parameters in rosemary leaves (*Rosmarinus officinalis* L.) using near infrared spectroscopy. *Z Arznei Gewurzpflanzen* **6**:79–84 (2001).
17 EDQM, *Pharmacopée Européenne* (6.0 edn). EDQM, Strasbourg, pp. 269–270 (2008).
18 Barnes RJ, Dhanoa MS and Lister SJ, Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl Spectrosc* **43**:772–777 (1989).
19 Mouazen AM, De Baerdemaeker J and Ramon H, Effect of wavelength range on the measurement accuracy of some selected soil constituents using visual–near infrared spectroscopy. *J Near Infrared Spectrosc* **14**:189–199 (2006).
20 Camps C, Robic R, Bruneau M and Laurens F, Rapid determination of soluble solids content and acidity of black currant (*Ribes nigrum* L.) juice by mid-infrared spectroscopy performed in series. *LWT – Food Sci Technol* **43**:1164–1167 (2010).
21 Bellon-Maurel V, Fernandez-Ahumada E, Palagos B, Roger JM and McBratney A, Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *Trends Anal Chem* **29**:1073–1081 (2010).
22 Gowen AA, Downey G, Esquerre C and O'Donnell CP, Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients. *J Chemometrics* **25**:375–381 (2011).
23 Williams P and Sobering D, How do we do it: a brief summary of the methods we use in developing near infrared calibrations, in *Spectroscopy: the Future Waves*, ed. by Davis AMC and Williams P. NIR Publications, Chichester, pp. 185–188 (1996).
24 Malley DF, McClure C, Martin PD, Buckley K and McCaughey WP, Compositional analysis of cattle manure during composting using a field-portable near-infrared spectrometer. *Commun Soil Sci Plant Anal* **36**:455–475 (2005).
25 Williams P, Variables affecting near-infrared reflectance spectroscopic analysis, in *Near-infrared Technology in the Agricultural and Food Industries*, ed. by Williams P and Norris K. American Association of Cereal Chemists, St Paul, MN, pp. 143–167 (1987).

