



## Correction of moisture effects on near infrared calibration for the analysis of phenol content in eucalyptus wood extracts

T. Giordanengo, J.P. Charpentier, J.M. Roger, S. Roussel, L. Brancheriau, G. Chaix, Baillères H Nom\_exemple

### ► To cite this version:

T. Giordanengo, J.P. Charpentier, J.M. Roger, S. Roussel, L. Brancheriau, et al.. Correction of moisture effects on near infrared calibration for the analysis of phenol content in eucalyptus wood extracts. *Annals of Forest Science*, 2008, 65 (8), p. 1 - p. 8. 10.1051/forest:2008065 . hal-00454498

**HAL Id: hal-00454498**

**<https://hal.science/hal-00454498>**

Submitted on 8 Feb 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Correction of moisture effects on near infrared calibration for the analysis of phenol content in eucalyptus wood extracts

Thomas GIORDANENGO<sup>1</sup>, Jean-Paul CHARPENTIER<sup>2</sup>, Jean-Michel ROGER<sup>3</sup>, Sylvie ROUSSEL<sup>4</sup>,  
Loïc BRANCHERIAU<sup>5</sup>, Gilles CHAIX<sup>6\*</sup>, Henri BAILLÈRES<sup>5</sup>

<sup>1</sup> R&D Tonnellerie Radoux, 10 Avenue Faidherbe, 17503 Jonzac, France

<sup>2</sup> INRA Orléans, Unité d'Amélioration, Génétique et Physiologie Forestières, 2163 Avenue de la Pomme de Pin, 45075 Orléans Cedex 2, France

<sup>3</sup> CEMAGREF, UMR ITAP, BP 5095, 34196 Montpellier Cedex 1, France

<sup>4</sup> ONDALYS, 385 Avenue des Baronnes, 34 730 Prades le Lez, France

<sup>5</sup> CIRAD, PERSYST Department, Production and Processing of Tropical Woods Research Unit, TA B40/16 BP 5035,  
34398 Montpellier Cedex 5, France

<sup>6</sup> CIRAD, BIOS Department - TA A-39, Genetic diversity and breeding of forest species unit research BP 5035, 34398 Montpellier Cedex 5, France

(Received 13 March 2008; accepted 30 August 2008)

## Keywords:

EPO /  
eucalyptus /  
phenols /  
moisture effect correction /  
NIRS

## Abstract

- Methods based on near infrared spectroscopy used to assess wood properties are susceptible to variations in physical parameters (temperature, grain size, etc.). As wood is a hygroscopically sensitive material, we studied the effects of moisture on near infrared absorbance and calibration to accurately determine the application potential of this technique under routine.
- A collection of *Eucalyptus urophylla* × *E. grandis* hybrid wood pieces were analysed to obtain reference calibration of polyphenol contents in wood extracts via NIR spectra acquired under constant moisture conditions. Other specimens from the same source were assessed to obtain spectra for eight moisture contents spanning a broad variation range. The effects of moisture on absorption and on estimates based on a reference model were analysed.
- An increase in moisture content prompted a rise in near infrared absorption over the entire spectrum and for water O-H absorption bands. The polyphenol content estimates obtained by assessing specimens against the reference calibration at variable moisture contents revealed prediction bias. Five correction methods were then tested to enhance the robustness relative to moisture.
- In-depth calibration and external parameter orthogonalization (EPO) were found to be the most efficient methods for offsetting this factor.

## Mots-clés :

EPO /  
eucalyptus /  
phénols /  
correction effets humidité /  
SPIR

## Résumé – Correction de l'influence de l'humidité sur l'étalonnage proche infrarouge de la teneur en phénols dans les extraits de bois d'*Eucalyptus*.

- Les méthodes basées sur la spectroscopie proche infrarouge pour estimer les propriétés du bois sont sensibles aux variations des paramètres physiques (température, granulométrie...). Le bois étant un matériau hygroscopique sensible, l'influence de l'humidité sur l'absorbance et l'étalonnage proche infrarouge a été étudiée afin de mieux considérer les possibilités d'applications dans des conditions réelles.
- Un étalonnage de référence de la quantité de polyphénols présents dans les extraits a été établi à partir de spectres d'une collection de bois d'hybrides d'*Eucalyptus urophylla* × *E. grandis* à humidité constante fixée. D'autres spectres ont été obtenus sur des échantillons de même provenance mais à 8 teneurs en eau couvrant une large plage de variation. L'influence de l'humidité sur l'absorbance proche infrarouge puis sur l'estimation par le modèle de référence a été analysée.

\* Corresponding author: [gilles.chaix@cirad.fr](mailto:gilles.chaix@cirad.fr)

- Une augmentation de l'humidité élève la ligne de base du spectre d'absorbance et également les régions d'absorbance caractéristiques de la liaison O-H de l'eau. Les estimations de la quantité de polyphénols à partir de l'étalonnage de référence sur les échantillons à humidité variable révèlent un biais sur les prédictions.
- Parmi les cinq méthodes de correction testées pour rendre l'étalonnage robuste vis-à-vis de l'humidité, l'étalonnage exhaustif et l'EPO (External Parameter Orthogonalisation) s'avèrent être les plus efficaces et permettent de s'affranchir de ce facteur.

## 1. INTRODUCTION

Near infrared spectroscopy (NIRS) is a rapid nondestructive analysis method that was developed for use in the agri-food sector in the 1960s (Ben-Gera and Norris, 1968; Hart et al., 1962). This technique can be implemented to assess a broad range of chemical and physical properties of materials when combined with multivariate analysis methods (Burns and Ciurczak, 2001; Bertrand and Dufour, 2000; Naes et al., 2001). Mathematical models based on statistical analysis have been developed to assess interactions between the near infrared absorption spectrum and the property evaluated to calibrate the apparatus. Near infrared spectroscopy is currently used in many other fields, including chemical and pharmaceutical industries.

The first research studies on NIRS of woods were conducted in the late 1980s, with conclusive applications in the paper industry (Birkett and Gambino, 1988, 1989; Easty et al., 1990). Many other studies have been carried out in this field since then. Not only the chemical aspects have been studied, calibrations were carried out for physical and mechanical features (Tsuchikawa, 2007) and NIRS has been proven efficient for assessing many wood properties.

It should be noted that all of these studies were conducted under perfectly controlled laboratory conditions. It has been shown that some physical parameters can upset calibration. In the agri-food sector, some studies have highlighted that temperature variations degrade the quality of the assessment model (Hansen et al., 2000; Kawano et al., 1995). Thygesen and Lundqvist (2000a; 2000b) reported their results on the impact of temperature on near infrared absorption in wood. This parameter was found to alter the calibration performances. Büning-Pfaue (2003) analysed the detrimental effects of moisture in various foods on calibrations used in the agri-food industry.

Factors that could upset near infrared calibration in wood are not necessarily the same as those affecting NIRS analysis of foods. Wood is a hygroscopic material, i.e. its properties vary according to the moisture content. Although processing can be closely monitored, the moisture content in the material used may vary markedly, thus inducing negative effects on near infrared calibration.

Moisture is a chemical component of wood, but it can also be considered as a physical parameter since it influences the wood characteristics. The aim of the present study was to study the effects of moisture on near infrared calibration by analysing the impact of moisture on near infrared absorption and calibration, while making near infrared calibration more robust with respect to this parameter.

A collection of *Eucalyptus urophylla* × *E. grandis* hybrid test pieces was used to directly calibrate polyphenol contents in wood extracts. Polyphenols are the main wood extracts and they have an essential role in the wood properties as colour, stability and natural durability. Milled samples of these specimens with different moisture contents were measured directly by NIRS. The effects of moisture variations on near infrared absorption and on calibration were then assessed. Different correction methods were tested to determine the impact of this factor.

## 2. MATERIALS AND METHODS

Test pieces corresponding to 200 full sibs were obtained from a Congolese *Eucalyptus urophylla* × *E. grandis* breeding stand. The trees had been planted at Pointe Noire (Congo) in 1992 and were felled in 1998 at the age of 59 months. A 5 cm thick wooden disk was sampled from each tree at breast height. Clear parts of these disks were milled (through a 0.5 mm mesh screen).

Three sample batches (S0, S1 and S2) were prepared to study the impact of moisture on the calibration of polyphenol contents in wood by NIRS. S0 was the calibration batch, S1 was used to study the moisture effects, and S2 was the independent validation batch. S0, S1 and S2 samples were randomly selected by polyphenol content class from the initial batch S (175 available samples) so that they would best represent the polyphenol content variability. S0, S1 and S2 consisted of 100, 45 and 30 samples, respectively.

### 2.1. Experimental procedure

Part of the ground wood meal was sampled for chemical analysis of wood extract polyphenol contents. These contents were quantified by colorimetric assay based on oxidation-reduction reactions with Folin-Ciocalteu reagent. In alkaline medium, phenols reacted with a phosphotungstic and phosphomolybdic acid mixture to form a blue complex that was visible at 630 nm. The staining intensity was then proportional to the polyphenol content in the extracts (two successive 80% methanol extractions were performed with 50 mg of eucalyptus wood meal).

Samples were placed in a closed atmosphere to avoid moisture loss during handling. Glass flat bottomed 1 mm thick × 22 mm diameter pillboxes were used. The glass composition ensured efficient near infrared radiation transmission, especially in the 12 500–4 000 cm<sup>-1</sup> range, where it was over 90%. The pillboxes were sealed with a silicone plug. For each sample, around 1.5 g of wood meal was placed in a pillbox. Wood meal from each wooden disk was homogenized.

All of the S1 and S2 samples were stabilized at eight moisture levels (0%, 3%, 6%, 9%, 12%, 15%, 18% and 21%) by humidification in a temperature and humidity controlled chamber. S1 and S2

wood meal sample batches were first scanned at  $h = 9\%$  under stable conditions to bring the moisture content down to  $3\%$ . The moisture levels of these samples were then gradually increased by  $3\%$  stages until reaching  $21\%$ . Five control samples (not considered in the NIRS study) were used to check for mass variations (every 4 h after at least 48 h of stabilisation) in order to determine whether the moisture level was at a steady state. Following the study at the eight moisture levels ( $3\%$ ,  $6\%$ ,  $9\%$ ,  $12\%$ ,  $15\%$ ,  $18\%$ ,  $21\%$ ), we dried the samples at  $103^\circ\text{C}$  using an anhydrous process and measured the absorbance at  $0\%$  moisture.

The samples were regularly homogenized to avoid an excessively high moisture gradient in the pillboxes. Once the target moisture levels were reached (checked by weighing), the pillboxes were resealed and spectral analysis was performed 48 h later. Moisture levels were monitored with standard deviations lower than or equal to  $1.1\%$  for S1 and S2 samples. The maximum moisture measurement error calculated for all samples was  $\pm 0.1\%$  for S1 and S2.

NIRS was performed under diffuse reflection directly in the pillboxes containing wood meal using a Bruker Vector 22/Ni spectrometer. This Fourier transform spectrometer is designed for reflection analysis of solids with an integrating sphere. This instrument is especially suitable for the analysis of heterogeneous samples. Spectral analysis was performed within the  $12\,500\text{--}3\,500\text{ cm}^{-1}$  ( $800\text{--}2\,850\text{ nm}$ ) range at  $8\text{ cm}^{-1}$  resolution (each spectrum consisted of 1 167 absorption values). A sintered gold standard was used to determine the near infrared absorption. This gold standard was measured during each new series of measurements. One hundred scans were performed for each sample, and means were calculated and compared to the standard in order to obtain the absorption spectrum of the sample. Note, finally, that the temperature remained constant ( $20^\circ\text{C}$ ) throughout the experiment, so this parameter did not bias the study results.

A reference calibration of the polyphenol content was first performed on the S0 samples. The partial least squares regression (PLS) method was used to correlate the near infrared spectrum findings with the reference chemical analysis data. The number of latent variables of the model, or the rank, was determined by leave-one-out cross validation. The regression coefficients of the reference model  $b$  and  $b_0$  of the following equation were determined.

$$\hat{y} = \mathbf{X} \mathbf{b} + b_0$$

where  $\mathbf{X}$  is the spectral data matrix,  $\mathbf{y}$  is the vector of the extract polyphenol content, and  $\hat{y}$  is the vector of the extract polyphenol content estimated by PLS model.

To test this model, the polyphenol content was estimated for the S2 validation samples, with spectra obtained at the same moisture level as that of the reference samples ( $h = 9\%$ ).

Finally, the reference calibration was used to estimate the polyphenol contents of the S2 validation samples at eight different moisture levels, which enabled us to assess the impact of moisture on the near infrared calibration.

The indicators used were the standard error of prediction (SEP), the determination coefficient ( $R^2$ ) between the reference and predicted values, the ratio of performance to deviation (RPD) and the bias linking the reference and predicted values.

$$\text{SEP} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

where  $\hat{y}$  is the estimated polyphenol content,  $y$  is the polyphenol content measured by the reference chemical analysis method, and  $n$  is the number of samples in the batch.

Moreover, the ratio of performance to deviation ( $\text{RPD} = \text{SD}/\text{SEP}$ ), or the ratio of the standard deviation of reference chemical analysis values to the SEP is used. The advantage of this indicator is that it standardizes SEPs and provides a basis for comparison between calibrations of different properties (Williams et al. 1993).

The bias ( $\text{bias} = \bar{\hat{y}} - \bar{y}$ ) represents the mean deviation of reference value estimations, where:  $\bar{\hat{y}}$  is the mean polyphenol content, and  $\bar{y}$  is the mean polyphenol content measured by the reference method.

By varying the moisture of the S2 sample batch, a bias could be calculated for each of the eight different moisture levels. The root mean square bias (RMbias) was calculated to account for this information with a single indicator.

$\text{RMbias} = \sqrt{\frac{\sum_{i=1}^k \text{bias}_i^2}{k}}$  with  $\text{bias}_i$ : bias obtained with samples scanned at moisture  $i$ .

In this study, no samples were considered as outliers and thus eliminated from the analysis.

## 2.2. Moisture correction methods

Several methods for correcting the impact of the moisture parameter have been tested (Chauchard et al., 2004).

### 2.2.1. Correction with known moisture levels in new samples

- A priori correction – This method involves transforming the acquired spectrum  $\mathbf{x}$  into  $\mathbf{x}^*$  through its moisture  $h$ :  $\mathbf{x}^* = f(\mathbf{x}, h)$ . A simple method involves forming a concatenation of  $\mathbf{h}$  (moisture column vector) as a supplementary variable of  $\mathbf{X}$ . Moisture is thus directly accounted for in the regression and the created model takes moisture variability into account. Hence,  $\mathbf{y}$  is estimated by:  $\hat{\mathbf{y}} = \mathbf{X}^* \mathbf{b} + b_0$ . The calibration was obtained with S1 batch samples scanned at the eight moisture levels. The concatenation method was used, where the moisture of each S1 sample was added to the absorbance matrix in a supplementary column. The spectral matrix was centred-reduced prior to performing the PLS regression.
- Model correction – It is also possible to correct  $\mathbf{b}$  according to the moisture:  $\mathbf{b}^* = f(\mathbf{b}, h)$ . The model is thus obtained by:  $\hat{\mathbf{y}} = \mathbf{X} \mathbf{b}^* + b_0$ . A model is developed at different moisture levels.  $\mathbf{y}$  is then estimated by these models. When the moisture  $h$  of a sample is known,  $\hat{\mathbf{y}}$  is obtained by interpolation. A calibration model was calculated for each moisture level on the basis of the S1 samples. The polyphenol content was thus estimated for each mean moisture level. Since the moisture of each sample was known, an interpolation (linear or 3rd order polynomial) was then used to determine the polyphenol content estimated by this correction method.
- A posteriori correction – This correction is performed by adding a moisture function  $h$  and  $b_0$  to the model.  $b_0^*$  is thus a correction term, such that:  $b_0^* = f(h, b_0)$  and  $\hat{\mathbf{y}} = \mathbf{X} \mathbf{b} + b_0^*$ . The polyphenol content of the S1 samples was estimated by calibration of the S0 reference batch. The determined bias was then adjusted by a moisture function. A 4th order polynomial function as well as a function containing the square root term were tested. The reference calibration was then validated (S2 batch) and the estimated values were corrected via these functions.

We considered that the moisture level was constant for these three corrections, i.e. measurable. If the moisture is not measurable during the application, it could likely be estimated by NIRS. We attempted to obtain a calibration of the moisture content in the S1 batch samples. The moisture contents of the S1 samples were thus predicted and subsequently used in the three methods described above.

### 2.2.2. Correction with unknown moisture levels in new samples

- (d). Exhaustive model – The developed model is based on a batch of samples spanning the variation range of the disturbance parameter. The moisture variation is thus accounted for by the regression method. A calibration encompassing all of the S0 and S1 samples was calculated and validated with the S2 batch.
- (e). External parameter orthogonalization (EPO) method – The EPO method involves orthogonalization (Roger et al., 2003). It is based on the hypothesis that the spectral domain is divided into three subdomains, where:  $\vec{C}$  represents the domain containing chemical information independent of the external parameter,  $\vec{G}$  is generated by disturbances caused by this parameter, and  $\vec{R}$  contains residual spectral information. Only  $\vec{C}$  contains information that is essential for the model. In projecting  $\mathbf{X}$  on the subdomain containing this information, the developed model no longer uses drop-in data. A projection operator  $\mathbf{Q}$  on  $\vec{G}$  is used, and  $\mathbf{X}$  then resolves into:  $\mathbf{X} = \mathbf{XQ} + \mathbf{X}^*$ , where  $\mathbf{X}^*$  is the corrected spectral matrix of the disturbance information due to the moisture parameter.  $\mathbf{Q}$  is estimated by performing a PCA on a batch of independent samples whose spectra are recorded at different moisture levels.

In this study, a PCA was performed on the S1 samples, which enabled us to determine a representative base for the variability induced by the moisture parameter. A new calibration was then established using the S0 absorbance matrix corrected for the moisture effect.

## 3. RESULTS

### 3.1. Reference chemical analysis

Table I gives the distribution of polyphenol contents in extracts from S0, S1 and S2 samples. Note that the means for the three groups were very close (4.65; 4.72; 4.62 mg gallic acid equivalent (GAE)/g dry matter for S0, S1 and S2, respectively). Moreover, S0 and S1 spanned a relatively broad and comparable variation range ( $CV \approx 40\%$ ). It should also be noted that S2 had a smaller polyphenol content range than S0 and S1. As S2 was the validation sample batch, no extrapolation was performed with the estimation models developed for S0 and S1. The chemical analysis measurement error calculated by a repeatability test, while taking the extraction and measurement into account, was 0.38 mg GAE/g dry matter.

### 3.2. Moisture impact

Absorption generally increased with moisture. Figure 1 presents the mean standardized spectra for S1 (standard normal variate method). This figure shows that two major absorption peaks attributed to water (1 455 nm and 1 929 nm) sharply

**Table I.** Distribution of polyphenol contents in wood extracts for S0, S1 and S2 and descriptive statistics.

Polyphenol contents in extracts <sup>1</sup>					
	Sample number	Mean	Standard error of deviation	Range	Coefficient of variation (%)
S0	100	4.65	1.87	1.30–10.85	40.3
S1	45	4.72	1.90	1.75–10.17	40.3
S2	30	4.62	1.51	2.11–7.78	32.6

<sup>1</sup> mg GAE/g dry matter.

increased according to the moisture. They consist of the first overtone of the O-H stretching band ( $2\nu_{1,3}$ ) at 1 455 nm and the combination of the O-H stretching band and the O-H bending band ( $2\nu_{1,3} + \nu_2$ ) at 1 929 nm. Several spectral regions were noted relative to moisture variations. Moisture seemed to have very little effect in the 1 150–1 400 nm and 1 550–1 850 nm regions. However, the impact was greater within the 800–1 150 nm and 2 050–2 550 nm ranges. The first absorption range included two important water absorption maxima at 970 nm for the second overtone of the O-H stretching band ( $3\nu_{1,3}$ ) and at 1 190 nm for the combination of the first overtone of the O-H stretching and the O-H bending band ( $2\nu_{1,3} + \nu_2$ ). The second absorption range could be explained by the fact that water associates strongly with organic polymers such as cellulose by hydrogen bonds. Cellulose is highly absorbent in the 2 000–2 500 nm range<sup>1</sup>, so cellulose absorption bands in the near infrared spectrum are strongly influenced by the interaction with moisture.

The same absorption ranges were also noted by Thygesen and Lundqvist (2000a) in a study on calibration of moisture contents in green common spruce wood beyond the fiber saturation point (FSP  $\approx 30\%$ ). The absorption patterns did not seem to be the same at 1 455 nm and 1 929 nm. Wood is a hygroscopic material — most of its physicochemical properties are modified by moisture variations. It is thus likely that the absorption patterns at 1 929 nm were associated with moisture-wood interaction phenomena.

### 3.3. Correction methods

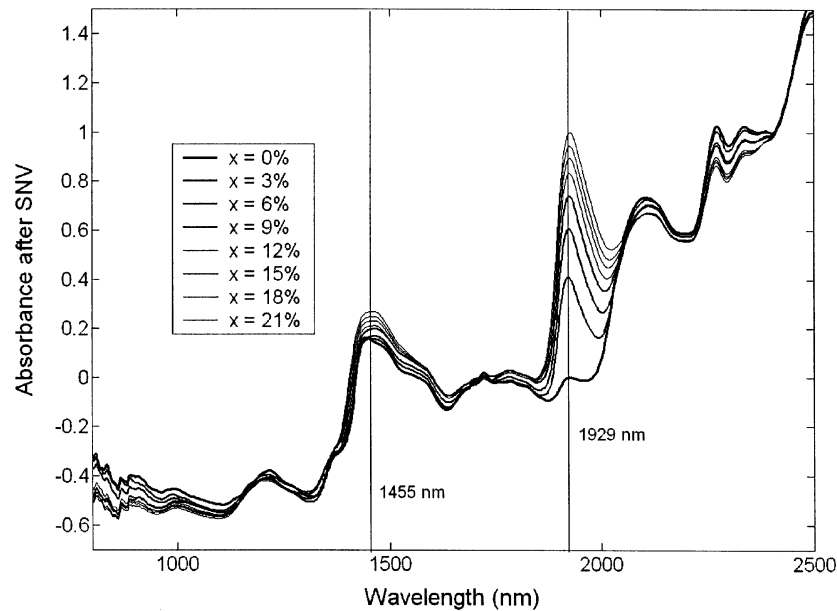
All of the results are summarized in Tables II and III.

#### 3.3.1. Correction methods with known moisture contents

- (a). A priori correction – The results obtained with the concatenation method indicated that it is highly efficient ( $R^2 = 0.74$  and SEP = 0.77 mg GAE/g dry matter). The SEP was even lower than that obtained during validation of S2 samples at  $h = 9\%$  on the reference calibration (SEP = 0.87 mg

<sup>1</sup> 2 336 nm corresponds to the C-H stretch and the C-H deformation of the cellulose molecule. The second wavelength (2100 nm) represents the O-H deformation and a C-O stretch of the NIR spectrum and is the designated region of the spectrum where starch absorbs.





**Figure 1.** Mean near infrared absorption spectra for 45 samples of the S1 batch acquired at eight different moisture contents (from 0% to 21%) after SNV normalization (line widths decrease inversely with the humidity).

**Table II.** Model validation results without correction and models corrected by methods requiring knowledge of the moisture content.

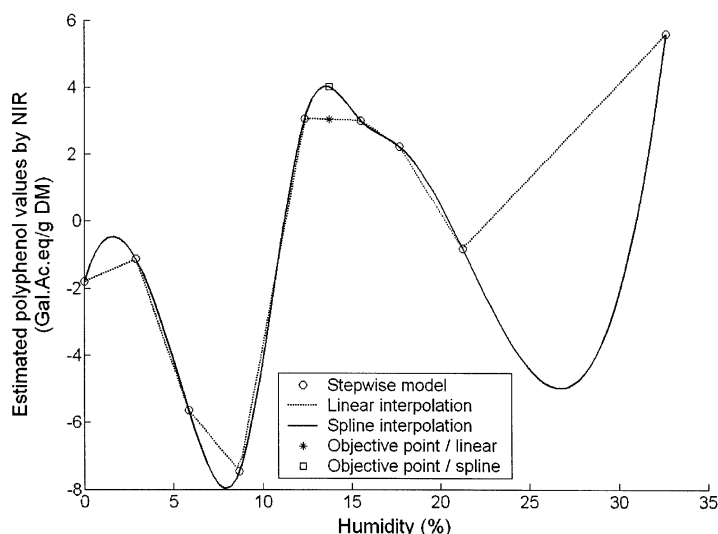
	Laboratory conditions		Known humidity				
	Initial models		A priori correction	Model correction by interpolation		A posteriori correction	
	Moisture fixed	S2 at moisture variable		Linear	3rd order polynomial	4th order polynomial adjustment	Root function adjustment
Calibration set	S0 centred	S0 centred	(S1, y) centred - reduced	S1 centred	S1 centred	S0 centred	S0 centred
Validation set	S2 at h = 9%	S2	S2	S2	S2	S2	S2
Rank	8	8	11	6–10	6–10	8	8
SEP <sup>1</sup>	0.87	15.11	0.77	0.89	0.88	1.23	1.41
RPD	1.7	0.1	2.0	1.7	1.7	1.2	1.1
RM Bias <sup>1</sup>	0.47	5.31	0.04	0.08	0.05	0.17	0.30
R <sup>2</sup>	0.77	0.02	0.74	0.66	0.67	0.48	0.41

<sup>1</sup> mg GAE/g dry matter.

**Table III.** Model validation results using an *h* estimate by near infrared spectroscopy and models with unknown moisture content.

	Spectrometric estimation of humidity				Unknown humidity		
	A priori	Model correction		A posteriori	EPO: npc = 3		
	correction	by interpolation		correction	Selectionwavelength		
	Concatenation	Linear	3rd order polynomial	4th order adjustment polynomial adjustment	Exhaustive model for S0 + S1	Same range as the initial model	1270–2150 nm wavelength range
Calibration set	(S0, y) centred-reduced	S4 centred	S1 centred	S0 centred	(S0; S1) centred	S0 centred	S0 centred
Validation set	S2	S2	S2	S2	S2	S2	S2
Rank	11	6–10		8	11	7	9
SEP <sup>1</sup>	0.76	0.95	1.93	1.26	0.73	1.04	0.78
RPD	2.0	1.6	0.8	1.2	2.1	1.5	1.9
RM Bias <sup>1</sup>	0.03	0.10	0.28	0.23	0.05	0.23	0.08
R <sup>2</sup>	0.74	0.62	0.22	0.51	0.76	0.72	0.75

<sup>1</sup> mg GAE/g dry matter.



**Figure 2.** A model correction for sample number 7 of the S2 batch at 15% moisture content (actual moisture: 13.9%). Comparison of linear interpolation and 3rd order polynomial interpolation results on estimated polyphenol contents (Gal.Ac.eq/g DM: gallic acid equivalent per gram of dry matter).

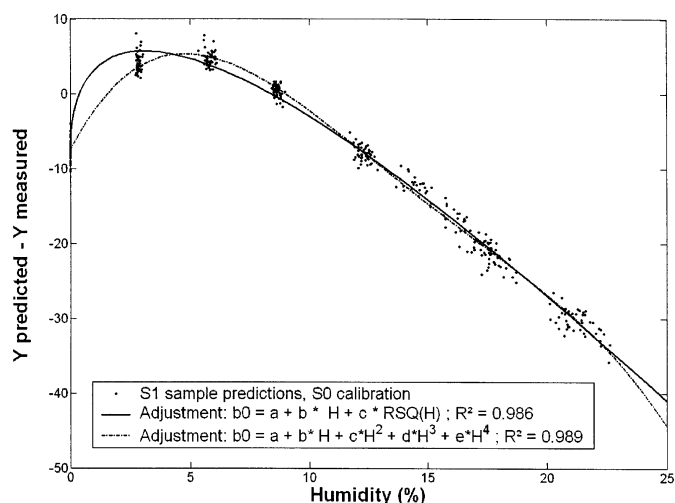
GAE/g dry matter). The RMbias was also almost nil (RMbias = 0.04 mg GAE/g dry matter). Note also that three additional latent variables were involved in the reference calibration on S0. New latent variables had to be taken into account on the basis of information derived from the moisture variable added to the spectral matrix.

- (b). Model correction – Figure 2 presents the polyphenol content in sample No. 17 of the S2 batch estimated by linear and polynomial (3rd order) interpolation. A model was developed for each moisture level with the S1 sample batch. The polyphenol content in the validation sample was estimated for each moisture level. As the sample moisture content was known, a linear or polynomial interpolation was performed to obtain the estimated polyphenol content. This figure highlighted the difference between the linear and polynomial interpolation results. The model correction was not very effective, i.e. validation of this method led to relatively unsatisfactory results with either linear or polynomial interpolation (respectively  $R^2 = 0.66$  and  $0.67$ ; SEP = 0.89 and 0.88 mg GAE/g dry matter). The RMbias was still less than 0.1 mg GAE/g dry matter, indicating that the method efficiently corrected the bias.

- (c). A posteriori correction – Two different functions were adjusted (4th order polynomial and root function) with respect to the variations in estimation bias for S1 by the reference calibration (Fig. 3). In both cases, this method did not efficiently correct the moisture effect, i.e. the validation gave poor results ( $R^2 < 0.5$  and SEP > 1.2).

### 3.3.2. Correction methods with unknown moisture contents

- (d). Exhaustive model – The exhaustive model developed on the basis of all the S0 and S1 samples gave interesting



**Figure 3.** Evolution along humidity of S1 estimated values (polyphenol contents) and S1 reference values bias by two adjusted functions (H: humidity).

results ( $R^2 = 0.76$  and SEP = 0.73 mg GAE/g dry matter). The RMbias was also almost nil (RMbias = 0.05 mg GAE/g dry matter). However, a calibration carried out only with S1 samples gave equivalent results ( $R^2 = 0.75$  and SEP = 0.75). Information on the S0 calibration set was therefore unnecessary, and the S1 training set was sufficient to perform the calibration.

- (e). External parameter orthogonalization (EPO) – The EPO method had one major advantage — all of the spectral information of the S0 calibration set was used after the impact of the disturbance factor had been overcome. Following PCA of the mean-centred S1 spectra, information of the first three components was eliminated from the S0 set.

The validation results revealed that the method is relatively efficient ( $R^2 = 0.72$ ; SEP = 1.04 mg GAE/g dry matter, RMbias = 0.23 mg GAE/g dry matter for seven latent variables). As explained, EPO is designed to eliminate part of the spectral basis. However, the wavelength selection can be broadened to improve part of the information remaining in the set. A tradeoff therefore has to be found because if the wavelength range is too broad, the added information will increase the noise. The EPO method was then tested with a 540 wavelength range (1 270–2 150 nm), which gave much better results ( $R^2 = 0.75$ ; SEP = 0.78 mg GAE/g dry matter; RMbias = 0.08 mg GAE/g dry matter for nine latent variables).

## 4. DISCUSSION

### 4.1. Impact of moisture on absorption

As shown, moisture variations had two different impacts on near infrared absorption. An increase in moisture content first affected the entire spectrum by shifting the baseline (additive effect). Radiation scattering could explain this type of phenomenon (Bertrand and Dufour, 2000). A moisture variation leads to a modification in the wood meal structure, i.e. the granulometry is modified as a result of shrinkage and swelling mechanisms. It is also likely that the addition of water molecules modifies the refraction indices. These two effects have a direct impact on reflection, refraction and diffraction phenomena. Near infrared waves are thus deviated from their incidence direction a high number of times and scattered in all directions, which is called radiation scattering. Statistically, the waves reflect off the material a high number of times and the intensity of the reflected radiation is offset by the scatter phenomenon. This increases the overall spectrum absorption. Moreover, an increase in the moisture content has a local impact on the near infrared absorption spectra. Indeed, water associates strongly with organic polymers by hydrogen bonds, so water absorption bands in the near infrared spectrum are influenced by effects of solutes in water. There are absorption shifts to lower or higher wavelengths and these shifts seem to be related to the hydration potential of the respective solutes (Bertrand and Dufour, 2000; Büning-Pfaue, 2003; Williams and Norris, 2001). This corroborates the fact that the near infrared spectrum is readily affected by conditions surrounding water molecules. The hydrogen bonds are relatively weak, and therefore the light absorbance energy is very sensitive to the ion strength, gelatinisation, swelling, shrinkage and the presence of other ions and molecules.

It should be noted that the scattering phenomenon had a local multiplicative effect as well as an overall additive effect (baseline shift). According to the Beer-Lambert law, there is then a linear increase in the absorption of these bonds.

### 4.2. Comparison of correction methods

The most effective correction methods were found to be the a priori method, with concatenation of the moisture contents

to the absorption matrix, the exhaustive model and EPO. Note that the a priori method was efficient in our case since the results were close to those obtained with the exhaustive method. All the S1 samples showed high moisture variability, and this was not necessarily information generated by concatenation, which would enable such a performance. The proof was obtained by validation of an exhaustive calibration based only on S1, which generated results equivalent to those obtained with the a priori correction method ( $R^2 = 0.75$  and SEP = 0.75 mg GAE/g dry matter).

The model correction and a posteriori correction methods were likely less efficient because a moisture variation does not only induce a clear-cut bias. The exhaustive calibration method with EPO correction provided the most robust correction. In practice, the major drawback of exhaustive calibration is that the base of the initial spectrum cannot be processed. This can be a serious problem when this base is already known and when the operator plans to apply it under conditions in which the disturbance parameter is not controlled. A new data acquisition process must be initiated in such cases. This could turn out to be laborious because variability in the parameter encountered during the application as well as that of the property to calibrate should be covered by the new spectral base. Moreover, this parameter is only accounted for through the regression method.

The EPO correction method was better for two reasons. First, this method enabled control of the type of information eliminated from the spectral base. Indeed, it is easy to discuss the principal components of the PCA performed for the EPO. These components are linear combinations of absorptions at different spectral wavelengths and can be readily linked with the absorption bands of the different chemical bonds. EPO is also of practical interest. It is possible to establish a learning spectral base that spans the variation range of the disturbance parameter prior to performing PCA. The principal components are determined and then removed from the initial spectral base. This spectral set is then used for the calibration.

## 5. CONCLUSION

Near infrared absorption increases with the moisture content of the material. This moisture impact involves two effects:

- a **global additive effect** corresponding to an increase in absorption throughout the spectrum.

- a **local multiplicative effect** linked with near infrared absorption bands specific to O-H chemical bonds of water.

These phenomena could be explained by modifications in the wood meal structure (shrinkage and swelling phenomena) and refraction indices induced by moisture variations. Reflection, refraction and diffraction mechanisms thus have different features and near infrared radiation scattering occurs.

Of all the correction methods assessed, exhaustive calibration and EPO were found to be the most efficient for offsetting disturbances caused by moisture variations. Exhaustive calibration is based on a batch of samples spanning the variation range of the disturbance parameter, and EPO relies on orthogonalization, separating relevant information and disturbance



in the spectral data. The results obtained with these methods were as efficient as those obtained for the reference calibration. However, EPO had two advantages: the initial calibration base was not lost and could be corrected with respect to disturbances due to moisture variations – it was also possible to analyse information removed from this spectral base.

## REFERENCES

- Ben-Gera I. and Norris K.H., 1968. Determination of moisture content in soybeans by direct spectrophotometry. *Isr. J. Agric. Res.* 18: 125–132.
- Bertrand D. and Dufour E., 2000. La spectroscopie infrarouge et ses applications analytiques. *Tec & Doc, Coll. STAA*, 223 p.
- Birkett M.D. and Gambino M.J.T., 1988. Potential applications for near infrared spectroscopy in the pulping industry. *Paper Southern Africa*, Nov-Dec: 34–38.
- Birkett M.D. and Gambino M.J.T., 1989. Estimation of pulp kappa number with near infrared spectroscopy. *TAPPI J.* 72-9: 193–197.
- Büning-Pfaue H., 2003. Analysis of water in food by near-infrared spectroscopy. *Food Chem.* 82: 107–115.
- Burns D. and Ciurczak E., 2001. *Handbook of Near-Infrared Analysis* 2nd Edition, Marcel-Dekker, Inc. New York, 2001, 280 p.
- Chauchard F., Roger J.-M., and Bellon-Maurel V., 2004. Correction of the temperature effect on near infrared calibration – application to soluble solid content prediction. *J. Near Infrared Spectrosc.* 12-3: 199–205.
- Easty D.B., Berben, S.A., Dethomas, F.A., and Brimmer P.J., 1990. Near-infrared spectroscopy for the analysis of wood pulp: quantifying hardwood-softwood mixtures and estimating lignin content. *TAPPI J.* 73-10: 257–261.
- Hansen W.G., Wiedemann S.C.C., Snieder M., and Wortel V.A.L., 2000. Tolerance of near infrared calibrations to temperature variations; a practical evaluation. *J. Near Infrared Spectrosc.* 8-2: 125–132.
- Hart J.R., Norris K.H., and Golumbic C., 1962. Determination of the moisture content of seeds by near-infrared spectrophotometry of their methanol extracts. *Cereal Chem.* 39: 94–99.
- Kawano S., Abe H., and Iwamoto M., 1995. Development of a calibration equation with temperature compensation for determining the Brix value in intact peaches. *J. Near Infrared Spectrosc.* 3: 211–218.
- Naes T., Isaksson T., Fearn T. and Davies T., 2001. *A User Friendly Guide to Multivariate Calibration and Classification*. NIR Publications, Chichester, 344 p.
- Roger J.-M., Chauchard F. and Bellon-Maurel V., 2003. EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. *Chemometr. Intell. Lab. Syst.* 66-2: 191–204.
- Thygesen L.G. and Lundqvist S.O., 2000a. NIR measurement of moisture content in wood under unstable temperature conditions. Part 1. Thermal effects in near infrared spectra of wood. *J. Near Infrared Spectrosc.* 8: 183–189.
- Thygesen L.G. and Lundqvist S.O., 2000b. NIR measurement of moisture content in wood under unstable temperature conditions. Part 2. Handling temperature fluctuations. *J. Near Infrared Spectrosc.* 8: 191–199.
- Tsuchikawa S., 2007. A review of recent near infrared research for wood and paper. *Appl. Spectrosc. Rev.* 42: 43–71.
- Williams P.C. and Sobering D.C., 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near Infrared Spectrosc.*
- Williams P.K. and Norris K., 2001. *Near-infrared Technology in the Agricultural and Food Industries* (2nd ed.), American Association of Cereal Chemists Inc., St. Paul, MN, USA, 129 p.