

APPLICATION NOTE

KEYWORDS

- Chemical composition
- Extractable compounds
- Resinous forest
- Wood
- Pinus Pinaster
- Resin

TECHNIQUES

- Chemometrics
- Near infrared spectroscopy
- Reflectance
- Non destructive method

APPLICATIONS

- Forest-wood sector
- Silviculture
- Logging

Near infrared spectroscopy applied to pine wood

In this note, we explore tools for detecting the rate of extractable compounds in wood samples in a non-destructive manner.



By Morandise Rubini, Thomas Cabaret and Bertrand Charrier. CNRS/Université de Pau des Pays de L'Adour, institut des Sciences Analytiques et de Physico-chimie pour L'environnement et Les Matériaux, Xylomat, UMR5254 and IDIL Fibres Optiques (France)

The forest-wood industry is present in many applications: logging and woodworking (sawing, carpentry, joinery, veneer, panels and parquet floors), consumer goods (construction, furniture, paper and cardboard), extraction of molecules, wood chemistry and energy.

Scientific researches are being carried out to support this forest-wood sector in France (Nouvelle Aquitaine). Non-destructive testing systems are developed to measure raw wood. The objective is to be able to determine extractable compounds in pine wood planks in order to be able to orient their uses and carry out quality control.

The use of near infrared spectroscopy coupled with statistical models has proven its effectiveness in measuring these parameters. Results are presented in this application note.

EXPÉRIENCES ON PINUS PINASTER

For this study, planks of maritime pine (sapwood and heartwood) were selected based on our knowledge: « the more a sample is colored, the richer it will be in resin ».

The parameter that interests us is the extractable content (in % of extractables relative to dry wood). It is determined by acetone extraction. In this case, the extractables consist mainly of monoterpenes, sesquiterpenes, dipterene and resin acids.

Equipment

The device provided by IDIL Fibres Optiques makes it possible to collect reflectance spectra:

- 1 MEMS-based near infrared spectrometer (1300–2500nm)
- 1 fiber optic probe (Y furcation)
- 1 halogen light source (360 –2400nm)
- Spectralon as reference material



All samples were measured 3 times for 10 seconds (integration time), then averaged, to obtain a characteristic spectrum per sample. Spectral data were exported with NanoQuest SW TM software as ASCII format files for further analysis.

RESULTS

Spectral pre-processing

Its goal is to eliminate or minimize variables that are unrelated to the property of interest under study (in this case, the extractable rate). Figure 1a shows raw spectra acquired on 78 wood samples.

On these spectra, we notice a shift of the baseline, with a curvilinear overall shape. This is due to variations in the optical path length during light scattering due to interactions between near infrared radiation and the sample. *Multiplicative Scatter Correction* (MSC) pre-processing is in this case chosen. MSC is a signal processing algorithm particularly useful in NIR spectroscopy due to the nonlinear scattering of light present in spectra acquired in diffuse reflection. The MSC performs a linear transformation of each spectrum to better match the individual spectra to the average of all spectra in the data set. This method is often used to correct the spectra measured in diffuse reflection, as well as the differences in the overall intensities of the signals. The result of this pretreatment is shown in Figure 1b.

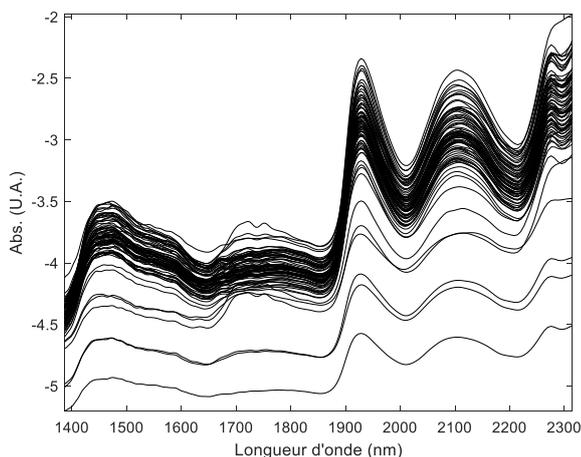


Figure 1a: Raw spectra.

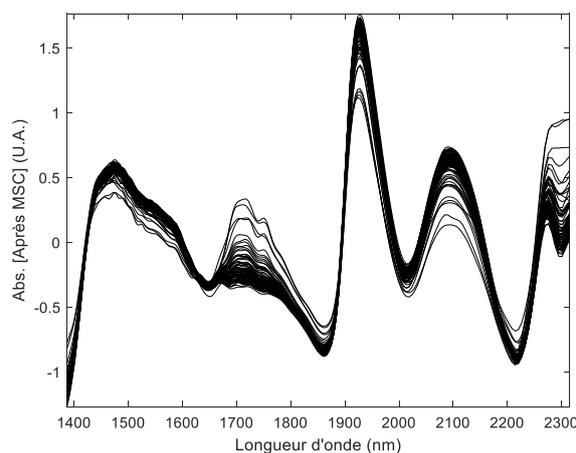


Figure 1b: Spectra after Multiplicative Scatter Correction (MSC) pre-processing

Partial Least Squares (PLS) regression

Partial Least Squares (PLS) model results on the pretreated spectra are shown in Table 1. PLS regression is one of the chemometric methods commonly used for the quantitative analysis of spectroscopic data. The PLS is a linear model that describes a dependent variable (parameter of interest) from independent variables (wavelength of the NIR spectrum).

To develop a reliable model, a large set of spectra (independent variables) and reference values (dependent variables) is necessary. The performance of the model depends on both the preprocessing and the range of selected wavelengths.

The critical decision regarding the development of prediction models is the selection of the optimal number of factors (components, latent variables). The choice of the number of latent variables must be done in a parsimonious manner. Indeed, risks of under-modeling (few latent variables) or over-modeling (too many latent variables) are possible.

Number of latent variables	3
R^2_{Cal}	0,81
R^2_{Val}	0,79
R^2_{Cv}	0,80
RMSEC	2,88
RMSEP	3,27
RMSECV	2,97

Table 1: PLS model performances

In practice, the optimal number of latent variables corresponds to the model which will give, for the smallest number of latent variables, the smallest prediction error, with the highest coefficient of determination. In our case, 3 latent variables are selected (See Table 1).

Validating the performance of chemometric models such as PLS regression is essential in the development of robust models for routine use.

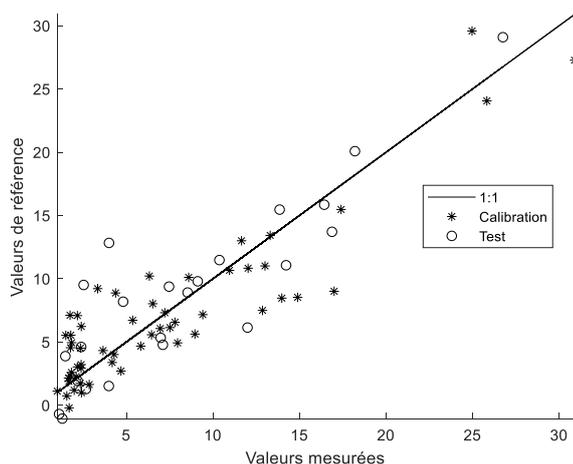
Two validation strategies are carried out:

1) Cross validation

In a calibration set, a defined set of samples is excluded from the same data set (cross validation set). Once the calibration model is developed with the remaining samples, the cross-validation samples are predicted. Successive iterations are performed, each time removing a different set of samples, until each sample has been omitted once. Cross validation is only effective as a proof of concept for the calibration performance. The statistics used to describe the calibration model are the coefficients of determination R^2_{Cal} , R^2_{CV} , and the square roots of the root mean square errors RMSEC and RMSECV for the calibration and cross validation, respectively.

RMSEC and RMSECV values are quite close, they are respectively 2,88 %, and 2,97 %. R^2_{Cal} , and R^2_{CV} values are very high, they exceed 0,8.

2) Testing the model on an independent dataset



The initial data set is divided into two sets consisting of a calibration set (calibration set) and validation (test set), with approximately two-thirds of the samples assigned to the calibration set, and one-third to the test game. The algorithms used to divide the initial data set allow for a test set independent of the calibration set, as well as promote representativeness of the data.

RMSEP values are 3,27 %. These values are quite close to those of the RMSEC, it means that the model is quite robust. As before, the values of R^2_{Val} are greater than 0.8.

CONCLUSION

The NanoQuest spectrometer allows us to quantify the level of extractables in wood from a non-destructive near infrared measurement. In order to achieve this objective, the spectra from the spectrometer were preprocessed, and a PLS regression method was used.

