

# Determination of soil content in chlordecone (organochlorine pesticide) using near infrared reflectance spectroscopy (NIRS)

Didier Brunet <sup>a,\*</sup>, Thierry Woignier <sup>b,c</sup>, Magalie Lesueur-Jannoyer <sup>d</sup>, Raphaël Achard <sup>d</sup>,  
Luc Rangon <sup>b</sup>, Bernard G. Barthès <sup>a</sup>

<sup>a</sup> IRD (Institut de recherche pour le développement), UMR Eco&Sols (Ecologie fonctionnelle et biogéochimie des sols), Montpellier SupAgro, bâtiment 12, 2 place Viala, 34060 Montpellier Cedex 1, France.

<sup>b</sup> IRD, UMR Eco&Sols, PRAM (Pôle de recherche agronomique de la Martinique), BP 213, Petit Morne, 97232 Le Lamentin, Martinique (French West Indies), France.

<sup>c</sup> CNRS (Centre national de la recherche scientifique) - Université Montpellier 2, place Eugène Bataillon, 34095 Montpellier cedex 5, France.

<sup>d</sup> CIRAD (Centre de coopération internationale en recherche agronomique pour le développement), PRAM, BP 213, Petit Morne, 97232 Le Lamentin, Martinique (French West Indies), France.

## Abstract

Chlordecone is a toxic organochlorine insecticide that was used in banana plantations until 1993 in the French West Indies. This study aimed at assessing the potential of near infrared reflectance spectroscopy (NIRS) for determining chlordecone content in Andosols, Nitisols and Ferralsols from Martinique. Using partial least square regression, chlordecone content conventionally determined through gas chromatography-mass spectrometry could be correctly predicted by NIRS ( $Q^2 = 0.75$ ,  $R^2 = 0.82$  for the total set), especially for samples with chlordecone content  $< 12 \text{ mg kg}^{-1}$  or when the sample set was rather homogeneous ( $Q^2 = 0.91$ ,  $R^2 = 0.82$  for the Andosols). Conventional measures and NIRS predictions were poorly correlated for chlordecone content  $> 12 \text{ mg kg}^{-1}$ , nevertheless ca. 80% samples were correctly predicted when the set was divided into three or four classes of chlordecone content. Thus NIRS could be considered a time- and cost-effective method for characterizing soil contamination by chlordecone.

## Capsule

Soil content in chlordecone, an organochlorine insecticide, can be determined time- and cost-effectively using near infrared reflectance spectroscopy (NIRS).

## Keywords

Near infrared reflectance spectroscopy (NIRS); Organochlorine pesticide; Chlordecone; Soil contamination

## 1. Introduction

Soil contamination by agricultural pesticides is of increasing concern. In Martinique and Guadeloupe islands (French West Indies) chlordecone was used from 1971 to 1993, mainly for protecting banana plantations against root borers. Chlordecone, also known as kepone, is a chlorinated polycyclic ketone pesticide with the formula  $C_{10}Cl_{10}O$  (Figure 1, according to Fournier, 2008). It is toxic, persistent, and induces cumulative and delayed toxicity (Sterrett and Boss, 1977; Epstein, 1978; Dawson et al., 1979; Hulster et al., 1994). In the French West Indies, previously contaminated agricultural soils have become new diffuse sources of chlordecone contamination for natural water and cultivated roots and tubers (Snegaroff, 1977). Chlordecone has strong lipid affinity and accumulates in the food chain (Sterrett and Boss, 1977). Studies have even been carried out to evaluate its concentration in the mother's milk (Multigner et al., 2007).

As complement and support to researches on pesticide adsorption on soils (Ahmad et al., 2001; Fushiwaki and Urano, 2001) and transfer from soil to water and plants (Topp et al., 1986; Hulster et al., 1994), there is a need to develop time- and cost-effective methods for analysing pesticide content in soil, water, plants, etc. Indeed, analysing pesticides in soils is often expensive and/or time consuming. As regards chlordecone, the usual method involves extraction with two solvents and quantification by gas chromatography-mass spectroscopy (GC-MS). An important advantage of the method is its low detection limit ( $0.01 \text{ mg kg}^{-1}$ ); however the whole process, including extraction, takes several hours, which does not easily allow analysing large amounts of samples such as those required for characterizing soil contamination and its variations on vast areas.

Near infrared reflectance spectroscopy (NIRS) has been reported to provide inexpensive, fast and accurate determination of many soil properties (Salgó et al., 1998; Reeves and McCarty, 2001; Ludwig et al., 2002). Several authors have even reported accurate prediction of pesticide sorption to soil using NIRS or mid infrared reflectance spectroscopy (MIRS): Bengtsson et al. (2007) used NIRS to predict lindane and linuron sorption in soils from Sweden; Forouzanoghar et al. (2008) and Kookana et al. (2008) used MIRS to determine sorption coefficient for diuron and atrazine in soils from Australia, respectively. In these studies, pesticide sorption to soil was determined using the batch sorption/desorption equilibrium technique.

The present study aimed at assessing the potential of NIRS for determining soil content in chlordecone in Andosols, Nitisols and Ferralsols that had been contaminated under banana plantation in Martinique.

## **2. Material and methods**

### *2.1. Soil samples*

A total of 236 samples were collected at 0-30 and 30-60 cm depths in Andosols, Nitisols and Ferralsols, all over the eastern part of Martinique (French West Indies), using an auger (Table 1). The samples were air-dried then crushed and sieved through a 2-mm mesh. Aliquots were finely ground to pass a 0.2-mm mesh and oven-dried at 40°C for 48 h before spectrum acquisition.

### *2.2. Conventional analysis of chlordecone*

Samples were analysed in the analysis laboratory of the Drôme county (LDA26 at Valence, France), which works under the norm NF17025. After sample homogenisation, chlordecone was extracted using dichloromethane and acetone (volume ratio 50/50). It was then assayed by gas chromatography-mass spectroscopy (GC-MS) using a Triple Quadripole Varian MS1200 analyser (Palo Alto, CA, USA), which calibration used the standard addition method and two internal standards, hexabromobenzene and triphenylphosphate. The resulting data are given with a 30% confidence interval.

### *2.3. Spectrum acquisition and pre-processing*

Soil samples (ca. 5 g) were scanned in the near infrared region between 1100 and 2500 nm at 2 nm intervals using a Foss NIRSystems 5000 spectrophotometer (Silver Spring, MD, USA) in order to determine their reflectance. More than 200 samples could be scanned daily, without any consumable. Each sample spectrum, automatically averaged from 32 spectra, was recorded as absorbance, which is the logarithm of the inverse of the reflectance ( $\log[1/R]$ ). Data analysis was conducted using the Winisi III-v1.61e software (Foss NIRSystems/Tecator International, LLC, Silver Spring, MD, USA).

In order to reduce spectral data, spectra were condensed by keeping the first out of four adjacent spectral points, yielding 173 data points per spectra (Barthès et al., 2006). First derivative of spectra, calculated over a 4-point gap with 4-point smoothing, was used either alone (None 144) or in conjunction with a standard normal variate transform (SNV 144) or a standard multiplicative correction (MSC 144). Derivatives reduce baseline variation and enhance spectral features (Reeves et al., 2002), SNV transform reduces the particle size effect (Barnes et al., 1989), and MSC removes additive and/or multiplicative signal effects (Martens et al., 1983).

#### 2.4. Processing methods

Statistical analyses were performed on the total set and on the Andosol set, but not on the Nitisol and the Ferralsol sets because they did not include enough samples. A principal component analysis (PCA) was carried out on the spectral dataset to calculate the Mahalanobis distance  $H$  (Mark and Tunnell, 1985). Samples with  $H > 3$  were considered spectral outliers and eliminated from further investigations (Shenk and Westerhaus, 1991a). Modified partial least square (mPLS) regression was used to correlate spectral data to conventional values (Shenk and Westerhaus, 1991b). The sample set was divided into a calibration subset selected by the software to include the most representative samples (150 for the total set and 100 for the Andosols), and a remaining validation subset (Shenk and Westerhaus, 1991a). A calibration equation was first built using all calibration samples, and provided standard error of calibration (SEC) and determination coefficient between predicted and conventional values (RSQ). Cross validation was then performed on the calibration subset, which was split into four groups, three being used for developing the model and one for prediction. The procedure was performed four times to use all calibration samples for both model development and prediction, then the residuals of the four predictions were pooled to calculate the standard error of cross validation (SECV). The outliers for calibration (i.e. samples with  $t > 2.5$ ) were removed and another cross validation was performed. This procedure was carried out twice. The number of factors giving the lowest final SECV determined the optimal number of terms to be used for the calibration. The accuracy of the cross validation was assessed using SECV and determination coefficient between predicted and conventional values ( $Q^2$ ). Model accuracy was then evaluated on the validation subset using the standard error of prediction (SEP), determination coefficient between predicted and conventional data ( $R^2$ ), and RPD, which is the ratio of standard deviation of validation subset to SEP (Islam et al., 2003).

### 3. Results and discussion

For the total set, spectrum pre-processing using None 144, SNV 144 and MSC 144 yielded respectively 13, 6 and 4 spectral outliers (i.e. 6%, 3% and 2%) and 14, 20 and 12 calibration outliers (i.e. 9%, 13% and 8% of the calibration subset); for the Andosols, it yielded respectively 9, 7 and 9 spectral outliers (i.e. 5%, 4% and 5%), and 10, 12 and 8 calibration outliers (i.e. 10%, 12% and 8% of the calibration subset).

Considering the three pre-processing procedures, RSQ (simple calibration) ranged from 0.80 to 0.87,  $Q^2$  (cross validation) from 0.70 to 0.75, and  $R^2$  (validation subset) from 0.77 to 0.82 for the total set (Table 2), which indicates reasonable fit. For the Andosols, RSQ ranged from 0.88 to 0.95 and  $Q^2$  from 0.82 to 0.91, which indicates better fit for calibration, while validation was not clearly affected, with  $R^2$  ranging from 0.80 to 0.82 (Table 3). Standard error or prediction (SEP) ranged from 2.4 to 3.3 mg kg<sup>-1</sup> for the total set, representing 35% to 47% of the mean, and was 2.4 mg kg<sup>-1</sup> for the Andosols, representing 30-32% of the mean. For the total set, RPD was 2.1-2.2 except with pre-treatment SNV 144, and it was 2.2 for the Andosols, which denotes accurate predictions, as predictions with  $RPD \geq 2$  have been considered reliable for soil properties (Chang et al., 2001; Brown et al., 2005). More accurate calibration for the Andosol set resulted probably from its greater homogeneity, as the total set included several soil types. Validation results were less affected by set homogeneity. In a previous work regarding NIRS determination of carbon and nitrogen contents, we showed that prediction accuracy increased with set homogeneity, regarding soil texture especially (Brunet et al., 2007).

Comparison between conventional determinations and NIRS predictions carried out with None 144 pre-processing on validation subsets are presented in Figure 2 for the total set and in Figure 3 for the Andosols. The figures show that NIRS predictions correlated well with conventional measures up to 12 mg kg<sup>-1</sup>; then predictions remained almost constant at 10-12 mg kg<sup>-1</sup> while conventional measures increased up to 20 mg kg<sup>-1</sup>, indicating that high values were under-predicted (this was also the case with other pre-processing procedures; data not shown). Nevertheless, when the validation subset of the total set was divided into four classes of chlordecone content (< 5, 5-10, 10-15 and > 15 mg kg<sup>-1</sup>), the comparison between measured and predicted classes yielded good results (Table 4): the proportion of properly classified samples was 77% for the whole subset, and 88%, 80%, 82% and 0% for samples with measured chlordecone content < 5, 5-10, 10-15 and > 15 mg kg<sup>-1</sup>, respectively. When three classes only were distinguished (i.e. < 5, 5-10 and > 10 mg kg<sup>-1</sup>), the proportion of properly classified samples reached up to 81% for the whole subset, and up to 71% for the > 10-mg kg<sup>-1</sup> class.

As a matter of fact, the calibration subsets of the total and Andosol sets included only samples with chlordecone content  $\leq 12.4$  mg kg<sup>-1</sup>, those with higher content being considered calibration outliers and removed. As a consequence, samples with content  $> 12.4$  mg kg<sup>-1</sup> were out of the calibration range and could not be properly analysed. Considering the total set, an attempt was made to keep samples with high chlordecone content in the calibration subset,

but the results were not improved: the performance of simple calibration ( $RSQ = 0.62-0.78$ ,  $SEC = 1.4-1.8 \text{ mg kg}^{-1}$ ) and cross validation ( $Q^2 = 0.55-0.64$ ,  $SECV = 1.8-2.0 \text{ mg kg}^{-1}$ ) decreased and that of validation did not change ( $R^2 = 0.75-0.82$ ,  $SEP = 2.3-3.1 \text{ mg kg}^{-1}$ ).

Pesticides such as chlordecone are organic compounds that tend to adsorb onto soil colloids, and several authors have underlined the role of soil organic matter and clay content in the retention of pesticides (Cheng, 1990; Spark and Swift, 2002; Li et al., 2003). In a report on the storage of organochlorine compounds in the soils of the French West Indies, Cabidoche et al. (2006) showed the close relation between soil contents in organic carbon ( $C_{org}$ ) and chlordecone. Organic carbon content is one of the soil properties that correlates most to absorbance in the NIR region (Salgó et al., 1998; Reeves and McCarty, 2001; Ludwig et al., 2002), as confirmed in the present study: with pre-processing None 144,  $RSQ$ ,  $Q^2$  and  $R^2$  ranged from 0.96 to 0.99,  $SEP$  from 8 to 6%, and  $RPD$  from 4.8 to 5.2 for the Andosols and the total set (data not shown). Nevertheless, NIRS prediction of chlordecone content did not result from NIRS prediction of  $C_{org}$  and correlation between  $C_{org}$  and chlordecone content, which could have been suspected. Indeed, determination coefficient between  $C_{org}$  and chlordecone content was 0.033 for Andosols and 0.002 for the total set, thus chlordecone content did not correlate to  $C_{org}$  (Figure 4). Moreover, wavelengths that contributed most to NIRS predictions of  $C_{org}$  and chlordecone content were not the same (Figure 5). The 20 wavelengths (out of 173) that contributed most to the calibration models of  $C_{org}$  and chlordecone in the Andosol and the total sets are presented in Table 5: four wavelengths were the same for  $C_{org}$  and chlordecone in the Andosol set (1468, 1508, 2036 and 2380 nm), five wavelengths were the same for both variables in the total set (1612, 1620, 1628, 1724 and 1780 nm); thus most wavelengths that contributed most to NIRS predictions of chlordecone content and  $C_{org}$  were not the same. For both the Andosol and the total sets, the regions 1470-1480, 1710-1720, 1780-1800 and 1830-1840 nm had major contributions to NIRS prediction of chlordecone content. The 1860 nm region has been attributed to the absorption due to the sixth overtone of C-Cl bond stretching (Shenk et al., 2001). This probably explains the very important contribution of the region 1830-1840 nm to NIRS prediction of chlordecone content, as the molecule includes ten C-Cl bonds (Figure 1). Additionally, the 1160, 1450 and 2030 nm regions have been assigned to the fourth, third and second overtones of C=O stretching, respectively (Shenk et al., 2001). This might explain the contributions of the regions 1196 and 2020-2040 nm in the Andosol set, and 1470-1480 nm in both sets, due to the presence of a C=O bond in chlordecone. However, the regions 2040 nm (for Andosols) and 1450-1470 nm (for both sets) also had important contributions to NIRS prediction of  $C_{org}$ ,

which is not surprising as soil organic matter includes C=O bonds. By contrast, the region 1830-1840 nm assigned to C-Cl did not contribute highly to NIRS prediction of  $C_{org}$ . Moreover, in both sets, nine out of the 20 wavelengths (i.e. ca. 50%) that contributed most to NIRS prediction of chlordecone content were included in the region 1700-1850 nm, which contributed much less to  $C_{org}$  prediction (Table 5).

The performance of NIRS prediction of soil chlordecone content was similar to those reported for NIRS or MIRS prediction of pesticide sorption to soil determined in batch sorption/desorption equilibrium experiments:  $R^2 = 0.84-0.85$  for NIRS prediction of lindane and linuron sorption to a range of soils from Sweden (Bengtsson et al., 2007);  $Q^2 = 0.69-0.72$  for MIRS prediction of atrazine and diuron sorption to soils from Australia (Forouzanoghar et al., 2008; Kookana et al., 2008). Using MIRS, Forouzanoghar et al. (2008) also reported under-prediction of diuron sorption to soils at high values of sorption coefficient. In the present study, under-prediction at high values could be attributed to non-linearity between pesticide concentration and absorbance. Fitting absorbance to the decimal logarithm of chlordecone content, according to a procedure carried out by Zornoza et al. (2008), yielded high validation  $R^2$  between predicted and measured log-concentration in chlordecone ( $R^2 = 0.89$ ) for the total set. Nevertheless, after back-conversion from log-concentration to concentration, validation  $R^2$  between predictions and measures was lower than initial  $R^2$  (0.76 vs. 0.82 for None 144 pre-processing). From a general viewpoint, Shepherd and Walsh (2002) considered that conventional laboratory analyses might fail to predict properly high contents in several soil analytes. To some extent, under-estimation by NIRS could thus correspond to over-estimation by conventional methods (confidence interval was 30% for conventional determination of chlordecone content in soil).

## Conclusion

The study demonstrated the interest of NIRS for the rapid and low-cost characterisation of soil content in chlordecone, an organochlorine insecticide, in contaminated Andosols, Nitisols and Ferralsols from Martinique. The performance of NIRS tended to be higher for the rather homogeneous Andosol set ( $Q^2 = 0.82-0.91$ ;  $R^2 = 0.80-0.82$ ) than for the total sample set ( $Q^2 = 0.70-0.75$ ;  $R^2 = 0.77-0.82$ ). Conventional measures and NIRS predictions correlated poorly at high chlordecone contents ( $> 12 \text{ mg kg}^{-1}$ ), which might either be due to NIRS or conventional analyses (confidence interval for conventional analyses was 30%). Nevertheless, when the validation set was divided into four classes according to chlordecone content, NIRS prediction of chlordecone class was correct (ca. 80% samples were properly classified).

## Acknowledgments

The authors thank the CIRAD physicochemical analysis laboratory, based in Montpellier (France), for soil carbon analysis. They also thank two anonymous referees for their comments and recommendations, which have helped improve the paper significantly.

## References

- Ahmad, R., Kookana, R.S., Alston, A.M., Skjemstad, J.O., 2001. The nature of soil organic matter affects sorption of pesticides. 1. Relationships with carbon chemistry as determined by  $^{13}\text{C}$  CPMAS NMR spectroscopy. *Environmental Science & Technology* 35, 878-884.
- Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and detrending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy* 43, 772-777.
- Barthès, B.G., Brunet, D., Ferrer, H., Chotte, J.L., Feller, C., 2006. Determination of total carbon and nitrogen content in a range of tropical soils using near infrared spectroscopy: influence of replication and sample grinding and drying. *Journal of Near Infrared Spectroscopy* 14, 341-348.
- Bengtsson, S., Berglöf, T., Kylin, H., 2007. Near infrared reflectance spectroscopy as a tool to predict pesticide sorption in soil. *Bulletin of Environmental Contamination and Toxicology* 78, 295-298.
- Brown, D.J., Brickley, R.S., Miller, P.R., 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma* 129, 251-267.
- Brunet, D., Barthès, B.G., Chotte, J.L., Feller, C., 2007. Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma* 139, 106-117.
- Cabidoche, Y.M., Clermont-Dauphin, C., Lafont, A., Sansoulet, J., Cattani, P., Achard, R., Caron, A., Chabrier, C., 2006. Stockage dans les Sols à Charges Variables et Dissipation dans les Eaux de Zoocides Organochlorés Autrefois Appliqués en Bananeraies aux Antilles : Relation avec les Systèmes de Culture. Rapport Final d'Exécution INRA-CIRAD dans le Cadre du Programme 2003-2005 "Evaluation et Réduction des Risques Liés à l'Utilisation des Pesticides". Ministère de l'Ecologie et du Développement Durable, Paris.
- Chang, C.W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R.J., 2001. Near-infrared reflectance spectroscopy - Principal components regression analyses of soil properties. *Soil Science Society of America Journal* 65, 480-490.
- Cheng, H.H., 1990. Organic residues in soils: mechanisms of retention and extractability. *International Journal of Environmental Analytical Chemistry* 39, 165-171.
- Dawson, G.W., Weimer, W.C., Shupe, S.J., 1979. Kepone - A case study of a persistent material. *The American Institute of Chemical Engineers (AIChE). Symposium Series* 75, 366-374.
- Epstein, S.S., 1978. Kepone-hazard evaluation. *The Science of the Total Environment* 9, 1-62.
- FAO-ISRIC-ISSS (Food and Agriculture Organization of the United Nations, International Soil Reference and Information Centre, International Society for Soil Science), 1998. *World Reference Base for Soil Resources*. FAO, Rome.
- Forouzanoghar, M., Kookana, R.S., Forrester, S.T., Smernik, R.J., Chittleborough, D.J., 2008. Midinfrared spectroscopy and chemometrics to predict diuron sorption coefficients in soils. *Environmental Science & Technology* 42, 3283-3288.
- Fournier, J., 2008. Le chlordécone, insecticide du charançon du bananier. *L'Actualité Chimique* 316, 46.



- Fushiwaki, Y., Urano, K., 2001. Adsorption of pesticides and their biodegraded products on clay minerals and solids. *Journal of Health Science* 47, 429-432.
- Hulster, A., Muller, J.F., Marschner, H., 1994. Soil plant transfer of polychlorinated dibenzo-p-dioxins and dibenzofurans to vegetables of the cucumber family (*Curcubitaceae*). *Environmental Science & Technology* 28, 1110-1115.
- Islam, K., Singh, B., McBratney, A., 2003. Simultaneous estimation of several soil properties by ultra-violet, visible and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research* 41, 1101-1114.
- Kookana, R.S., Janik, L.J., Forouzangohar, M., Forrester, S.T., 2008. Prediction of atrazine sorption coefficients in soils using mid-infrared spectroscopy and partial least-squares analysis. *Journal of Agricultural and Food Chemistry* 56, 3208-3213.
- Li, H., Sheng, G., Teppen, B.J., Johnston, C.T., Boyd, S.A., 2003. Sorption and desorption of pesticides by clay minerals and humic acid-clay complexes. *Soil Science Society of America Journal* 67, 122-131.
- Ludwig, B., Khanna, P.K., Bauhus, J., Hopmans, P., 2002. Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. *Forest Ecology and Management* 171, 121-132.
- Mark, H.L., Tunnell, D., 1985. Qualitative near-infrared reflectance analysis. *Analytical Chemistry* 58, 379-384.
- Martens, H., Jensen, S.A., Geladi, P., 1983. Multivariate linearity transformation for near infrared reflectance spectrometry, in: *Proceedings of the Nordic Symposium on Applied Statistics*, Stokkand Forlag Publishers, Stavanger, Norway, pp. 205-234.
- Multigner, L., Cordier, S., Kadhel, P. Huc-Terki, F., Blanchet, P., Bataille, H., Janky, E., 2007. Pollution par le chlordécone aux Antilles. Quel impact sur la santé de la population ? *Environnement, Risques et Santé* 6, 405-407.
- Reeves, J.B., McCarty, G.W., 2001. Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and a fibre-optic probe. *Journal of Near Infrared Spectroscopy* 9, 25-34.
- Reeves, J. III, McCarty, G., Mimmo, T., 2002. The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. *Environmental Pollution* 116, S277-S284.
- Salgó, A., Nagy, J., Tarnóy, J., Marth, P., Pálmai, O., Szabó-Kele, G., 1998. Characterisation of soils by the near infrared technique. *Journal of Near Infrared Spectroscopy* 6, 199-203.
- Shenk, J.S., Westerhaus, M.O., 1991a. Population definition, sample selection and calibration procedures for near infrared reflectance spectroscopy. *Crop Science* 31, 469-474.
- Shenk, J.S., Westerhaus, M.O., 1991b. Population structuring of near infrared spectra and modified partial least square regression. *Crop Science* 31, 1548-1555.
- Shenk, J.S., Workman Jr., J.J., Westerhaus, M.O., 2001. Application of NIR spectroscopy to agricultural products, in: Burns, D.A., Ciurczak, E.W. (Eds.), *Handbook of Near-Infrared Analysis*, second ed. CRC Press, New York, pp. 419-474.
- Shepherd, K.D., Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soils properties. *Soil Science Society of America Journal* 66, 988-998.
- Snegaroff, J., 1977. Les résidus d'insecticides organochlorés dans les sols et les rivières de la région bananière de Guadeloupe. *Phytiatrie-Phytopharmacie* 26, 251-268.
- Spark, K.M., Swift, R.S., 2002. Effect of soil composition and dissolved organic matter on pesticide sorption. *Science of the Total Environment* 298, 147-161.
- Sterrett, F.S, Boss, C.A., 1977. Careless Kepone. *Environment* 19, 30-37.

- Topp, E., Scheunert, I., Attar, A., Korte, F., 1986. Factors affecting the uptake of  $^{14}\text{C}$ -labelled organic chemicals by plants from soil. *Ecotoxicology and Environmental Safety* 11, 219-228.
- Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K.M., Arcenegui, V., Mataix-Beneyto, J., 2008. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. *Soil Biology & Biochemistry* 40, 1923-1930.

**Table 1**

Minimum, maximum, mean and standard deviation (SD) of soil content in chlordecone ( $\text{mg kg}^{-1}$ , conventional determination) in the studied sample set and according to soil type (FAO-ISRIC-ISSS classification, 1998).

	<i>n</i>	Min	Max	Mean	SD
Total set	236	0.0	19.8	3.9	4.3
Andosols	169	0.0	19.8	4.7	4.8
Nitisols	12	0.0	2.3	0.8	0.8
Ferralsols	55	0.1	6.2	2.2	1.4

*n* is the number of samples.

**Table 2**

Calibration and validation results for soil content in chlordecone ( $\text{mg kg}^{-1}$ ) for the total sample set.

Pre-processing method	Calibration subset					Validation subset			
	<i>n</i> <sub>1</sub>	SEC	RSQ	SECV	Q <sup>2</sup>	<i>n</i> <sub>2</sub>	SEP	R <sup>2</sup>	RPD
None 144	136	0.9	0.87	1.3	0.73	73	2.4	0.82	2.2
SNV 144	130	0.7	0.83	0.9	0.75	80	3.3	0.77	1.6
MSC 144	138	1.0	0.80	1.3	0.70	82	2.5	0.81	2.1

*n*<sub>1</sub> is the number of samples in the calibration subset after elimination of calibration outliers.

*n*<sub>2</sub> is the number of samples in the validation subset.

SEC, SECV and SEP are standard errors of calibration, cross validation and prediction, respectively.

RSQ, Q<sup>2</sup> and R<sup>2</sup> are determination coefficients of calibration, cross validation and prediction, respectively.

RPD is the ratio of standard deviation (of the validation subset) to SEP.

**Table 3**

Calibration and validation results for soil content in chlordecone ( $\text{mg kg}^{-1}$ ) for the Andosol set.

Pre-processing method	Calibration subset					Validation subset			
	$n_1$	SEC	RSQ	SECV	$Q^2$	$n_2$	SEP	$R^2$	RPD
None 144	90	0.6	0.95	0.9	0.91	60	2.4	0.82	2.2
SNV 144	88	1.0	0.88	1.2	0.82	62	2.4	0.81	2.2
MSC 144	92	1.0	0.88	1.2	0.82	60	2.4	0.80	2.2

$n_1$  is the number of samples in the calibration subset after elimination of calibration outliers.

$n_2$  is the number of samples in the validation subset.

SEC, SECV and SEP are standard errors of calibration, cross validation and prediction, respectively.

RSQ,  $Q^2$  and  $R^2$  are determination coefficients of calibration, cross validation and prediction, respectively.

RPD is the ratio of standard deviation (of the validation subset) to SEP.

**Table 4**

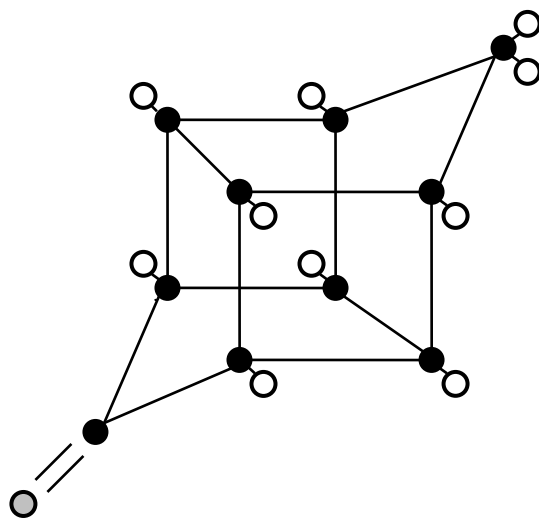
Comparison between the size of measured and predicted classes of chlordecone content for the validation subset of the total set with pre-processing None 144.

		Predicted chlordecone content ( $\text{mg kg}^{-1}$ )				Total
		0-5	5-10	10-15	15-20	
Measured chlordecone content ( $\text{mg kg}^{-1}$ )	0-5	30	4	0	0	34
	5-10	0	12	3	0	15
	10-15	0	3	14	0	17
	15-20	0	4	3	0	7
	Total	30	23	20	0	73

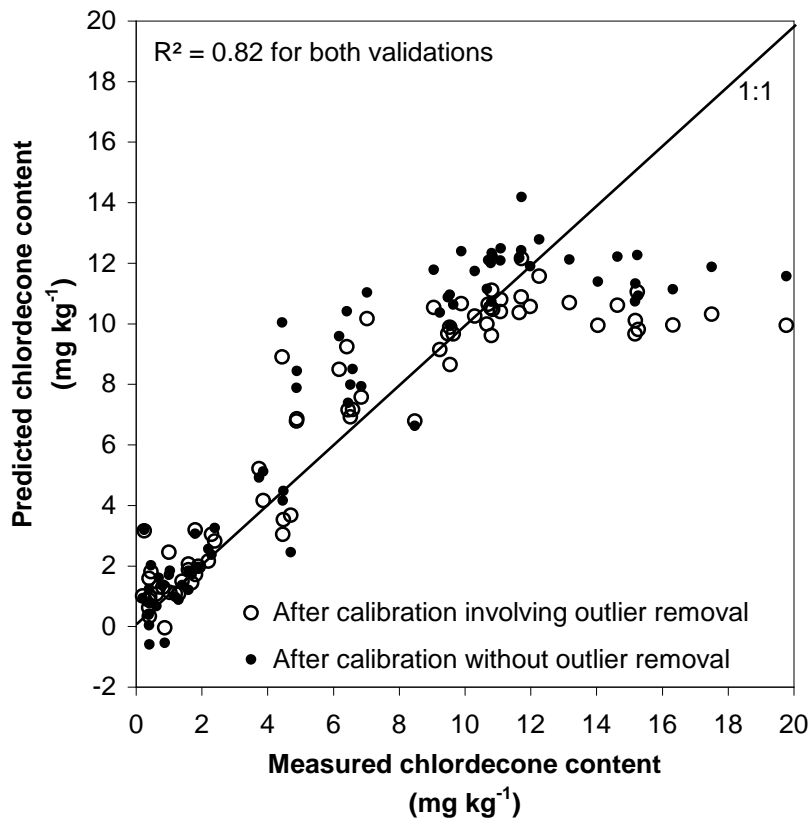
**Table 5.** Wavelengths with highest contributions (b-coefficients) to NIRS predictions of C<sub>org</sub> and chlordecone contents using pre-treatment None 144 for the Andosol and the total sets (wavelengths in bold have high contribution to the prediction of both variables in a given set; those underlined have high contribution to the prediction of a given variable in both sets).

Andosol set				Total set			
Organic carbon prediction		Chlordecone prediction		Organic carbon prediction		Chlordecone prediction	
Wave-length (nm)	Regression coeff.	Wave-length (nm)	Regression coeff.	Wave-length (nm)	Regression coeff.	Wave-length (nm)	Regression coeff.
<u>1620</u>	9336	1740	-4461	<u>1580</u>	11148	<u>1844</u>	-9097
<u>1636</u>	8346	<u>1796</u>	3279	1556	-10621	<u>1796</u>	7589
<b>1468</b>	7917	<u>1828</u>	-3274	<b>1620</b>	10488	1668	7322
<u>1580</u>	7044	<b>2380</b>	3170	1604	9555	<b>1724</b>	-6466
1676	-6966	<b>2036</b>	3134	<b>1780</b>	-8800	<b>1612</b>	6208
2260	-6371	<u>1476</u>	-2797	<b>1628</b>	8569	<u>1828</u>	-6029
2340	6153	<u>1844</u>	-2621	<b>1612</b>	7649	<u>1708</u>	5985
<b>2380</b>	6129	1748	-2533	1340	-7345	<b>1628</b>	4916
2300	6055	<b>1508</b>	-2511	<u>1636</u>	7075	<b>1780</b>	-4856
1524	-5856	1716	2471	1300	6831	1836	-4579
<b>2036</b>	-5331	2020	2420	1276	-6811	1700	4493
1492	5329	2172	2418	1540	-6530	<u>1468</u>	-4454
<u>1612</u>	5281	1764	2312	1548	-6446	1788	4446
1452	5194	<u>1780</u>	-2191	<u>2036</u>	-6419	<b>1620</b>	4381
1564	5182	1196	2156	1596	6260	1436	-4237
<b>1508</b>	-5157	2076	-2122	2004	6185	2108	-3961
2404	5137	<u>1708</u>	2112	1572	6163	1516	-3851
2244	-5080	<b>1468</b>	-2093	2020	-6059	<u>1476</u>	-3715
1460	4981	2236	2068	<b>1724</b>	-6058	1676	3695
1588	4929	2388	2036	2028	-6023	1692	3535

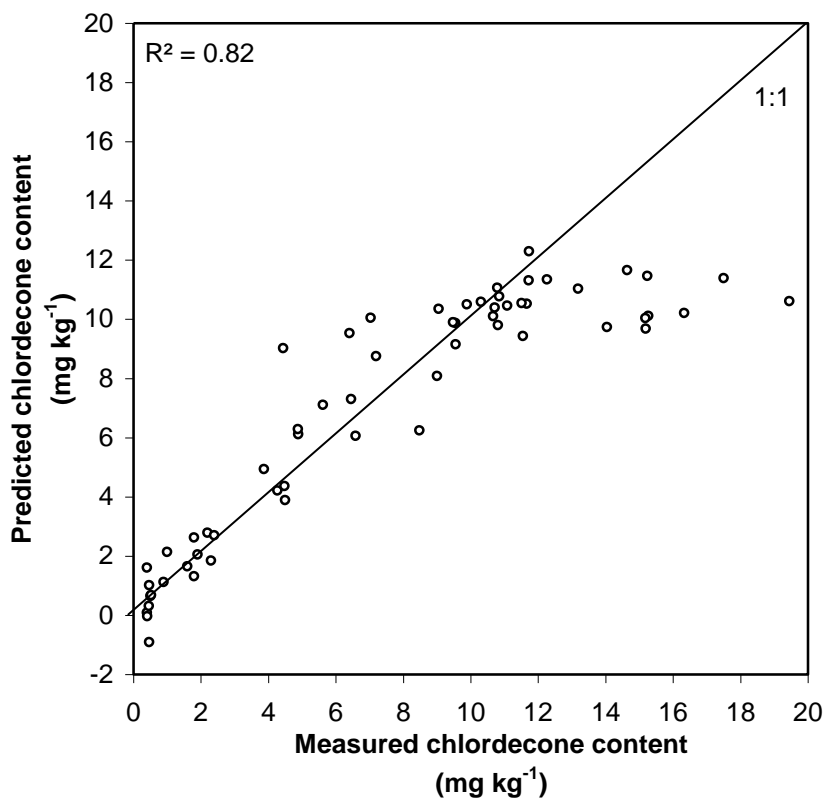
**Figure 1.** Molecular structure of chlordecone (black, white and grey circles represent carbon, chlorine and oxygen atoms, respectively).



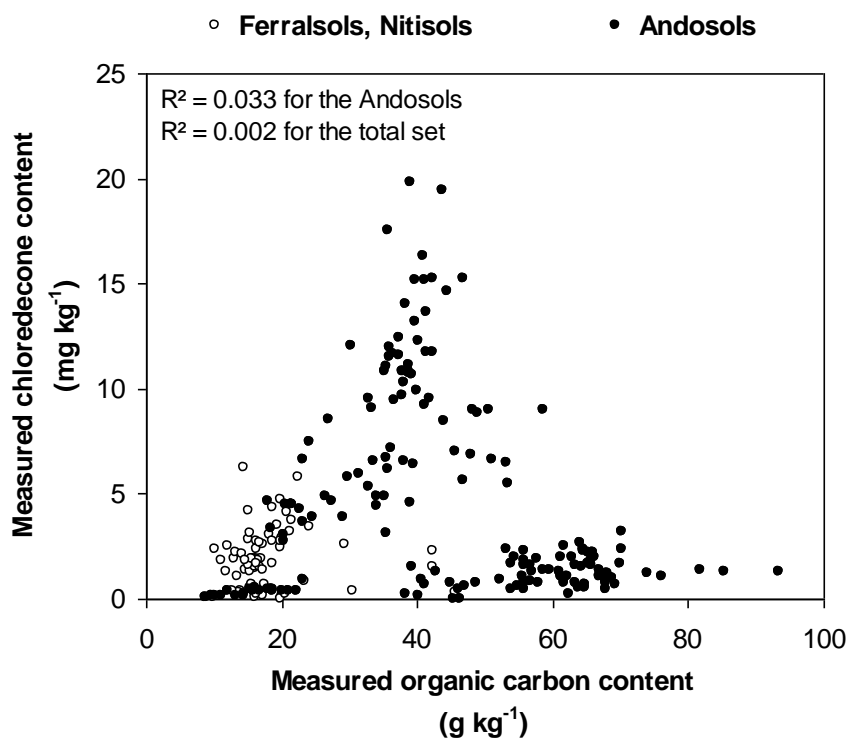
**Figure 2.** Comparison between predicted (NIRS) and measured (conventional) values of soil content in chlordecone ( $\text{mg kg}^{-1}$ ) for the validation subset of the total set with spectrum pre-processing None 144, after calibration either involving outlier removal or not.



**Figure 3.** Comparison between predicted (NIRS) and measured (conventional) values of soil content in chlordecone ( $\text{mg kg}^{-1}$ ) for the validation subset of the Andosol set with spectrum pre-processing None 144.



**Figure 4.** Relationship between soil organic carbon and chlordecone contents.



**Figure 5.** Comparison between coefficients of regression of chlordecone vs. organic carbon contents on absorbance at every wavelength (b-coefficients) for the Andosol set, with pre-treatment None 144.

