FTIR spectroscopy can predict organic matter quality in regenerating cutover peatlands.

To cite this version:
**FTIR spectroscopy can predict organic matter quality in regenerating cutover peatlands**

(Running title: Whole soil FTIR on peat)

Rebekka R.E. Artz¹, Stephen J. Chapman¹, A.H. Jean Robertson¹, Jacqueline M. Potts², Fatima Laggoun-Défarge³, Sébastien Gogo³, Laure Comont³, Jean-Robert Disnar³ and Andre-Jean Francez⁴

¹ Soils Group, The Macaulay Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK.
² Biomathematics and Statistics Scotland (BioSS), The Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK.
³ Institut des Sciences de la Terre d’Orléans (ISTO), UMR 6113 CNRS/Université d’Orléans, Bt Géosciences, Rue de St Amand, BP 6759, 45067 Orléans cedex 2, France.
⁴ Interactions biologiques et transferts de matières, UMR CNRS 6553 Ecobio, Université de Rennes 1, 35042 Rennes cedex, France.

Date of receipt:

**Keywords:** Fourier-Transform Infrared Spectroscopy, peat, organic matter quality, carbohydrates, C:N ratio, organic micro-remains

**Corresponding author:** Rebekka R.E. Artz, The Macaulay Institute, fax: (+44) (0)1224 498207; e-mail: r.artz@macaulay.ac.uk
Abstract

Vegetational changes during the restoration of cutover peatlands leave a legacy in terms of the organic matter quality of the newly formed peat. Current efforts to restore peatlands at a large scale therefore require low cost, and high throughout, techniques to monitor the evolution of organic matter. In this study, we assessed the Fourier Transform Infrared (FTIR) spectra of the organic matter in peat samples at various stages of peatland regeneration from five European countries. Using predictive partial least squares analyses, we were able to reconstruct both peat C:N ratio and carbohydrate signatures, but not the micromorphological composition of vegetation remains, from the FTIR datasets. Despite utilising different size fractions, both carbohydrate (< 200 µm fraction) and FTIR (bulk soil) analyses report on the composition of plant cell wall constituents in the peat and therefore essentially reveal the composition of the parent vegetational material. This suggests that FTIR analysis of peat may be used successfully for evaluation of the present and future organic matter composition of peat in monitoring of restoration efforts.

1. Introduction

Northern Peatlands are composed almost entirely of decomposing plant material and store approximately a third of all soil organic matter (Gorham, 1991) even though their total cover only extends to 3-5 % of the global land area. Peat extraction for fuel and horticultural use has steadily diminished this carbon stock, with the largest quantities of peat having been extracted in the mid to late 20th century (Chapman et al., 2003). Various restoration programs have since been designed to encourage revegetation of cut-over peatlands (Gorham and Rochefort, 2003). Although some of these programs have demonstrated that annual gaseous emissions
show a return to net carbon sequestration (Tuittila et al., 1999) or at least reduce net emissions (Waddington and Warner, 2001), it is not known how peatland restoration affects the pool of soil organic matter and hence the long-term regeneration of the carbon sequestration potential. Increased losses of dissolved organic carbon (DOC) have been observed from many peatland ecosystems in the past decades (Freeman et al., 2001), and some of this can be ascribed to increased turnover of the soil organic matter (Glatzel et al., 2003). Currently, monitoring efforts of the evolution of soil organic matter quality during restoration of peatlands have only a limited array of tools. Generally, bulk measures such as total and soluble organic carbon and nitrogen, and their ratios, have been most often used to assess restoration success (Andersen et al., 2006; Comont et al., 2006). Similarly, a technique often employed in peat organic matter compositional studies is analysis of the patterns of carbohydrate monomers derived from plant cellulose and hemicelluloses as these are indicative of the source plant composition and the preservation status of these remains (Cheshire, 1979; Moers et al., 1989, 1990; Bourdon et al., 2000). Comont et al., (2006) used peat C:N ratios combined with micromorphological and carbohydrate composition of peat in a pioneering study to elucidate the evolution of organic matter with regeneration. These techniques, however, are expensive and time consuming processes. FTIR spectroscopy is a commonly used technique capable of distinguishing the principal chemical classes in soil organic matter, such as carbohydrates, lignins, cellulose, fats and/or lipids and proteinaceous compounds, through the vibrational characteristics of their structural chemical bonds. The use of attenuated total reflectance accessories, in particular those utilising very hard crystals such as diamond, has further advanced the use of FTIR in soils and other solid residues. Dilution with KBr is no longer necessary, reproducibility is increased and the nondestructive nature of this analysis allows the sample to be re-used for other analyses. FTIR spectroscopy has been used
successfully on whole soils to describe the status of decomposition in different horizons (Haberhauer et al., 1998, 1999; Chapman et al., 2001), for example through following the reduction of the carbohydrate markers with depth. Using multivariate statistics, FTIR data can be used as quantitative indicators of the composition of the soil organic matter to distinguish soil horizons (Haberhauer et al., 1999, 2000). Models utilising partial least squares (PLS) analysis have been applied to FTIR data to predict various chemical and physical qualities of organic materials, including studies of the lignin and carbohydrate contents of wood and woody peat (Durig et al., 1988; Tucker et al., 2001; Bjarnestad and Dahlman, 2002) and the phenolic and carbohydrate contents of food (e.g. Coimbra et al., 2005). This study investigated the potential use of FTIR spectroscopy data as indicators of peat organic matter quality in regenerating peatlands. We determined various chemical and micromorphological characteristics of peat samples from profiles at sites at different stages of regeneration from five cutover European peatlands and tested the power of partial least squares analysis using FTIR data to predict these organic matter characteristics. In large scale restoration projects, it would be advantageous to be able to use low cost and high throughput techniques in order to assess the success of restoration efforts. Our results are therefore discussed with respect to the utility of FTIR spectroscopy coupled to predictive PLS in the assessment of organic matter quality with peatland regeneration.

2. Materials and Methods

2.1. Sampling procedure

Sites within gradients of unaided regeneration were selected in previously cut-over peatlands in five countries in Europe (Table 1). Cores \( n = 3 \) were obtained from each site with a double-skinned peat corer (to avoid compaction) and were sectioned into 4 horizons of different stages of decomposition. The horizons were designated...
horizons 3 (surface layer 0-5 cm), 4 (5-10 cm), 6 (22.5-27.5 cm) and 8 (42.5 to 47.5 cm). Core samples were cut into 1 cm$^3$ subsamples and the subsamples mixed to ensure homogeneity. Portions were shipped on ice packs to partner laboratories for the relevant analyses contributing to this study. For the purpose of this comparative study, only a single replicate from each Country x Site x Horizon combination was analysed for all analytes. Samples where not all analyses could be completed due to low sample size were excluded from statistical analyses, reducing the dataset for statistical analyses to $n = 68$ (Table 1).

2.2. FTIR spectroscopy

Spectral characterisation of peat samples was performed by diamond attenuated total reflectance FTIR spectroscopy using a Nicolet Magna-IR 550 FTIR spectrometer (Thermo Electron, Warwick, U.K.) fitted with a potassium bromide beam splitter and a deuteroglycine sulphate detector. A Diamond Attenuated Total Reflectance (DATR) accessory, with a single reflectance system, was used to produce transmission-like spectra. The samples were dehydrated by freeze drying and powdered by ball milling with zirconium balls. Samples were placed directly on a DATR/KRS-5 crystal and a flat tip powder press was used to achieve even distribution and contact. Spectra were acquired by averaging 200 scans at 4 cm$^{-1}$ resolution over the range 4000 – 350 cm$^{-1}$. A correction was made to spectra for the ATR to allow for differences in depth of beam penetration at different wavelengths (Omnic software, version 7.2, Thermo Electron). All spectra were also corrected for attenuation by water vapour and CO$_2$. Minor differences in the amplitude and baseline between runs were corrected by normalisation of the data by subtraction of the sample minimum followed by division by the average of all data points per sample. First and
second derivatives were calculated to determine and test correlations of organic matter variables which formed ‘shoulders’ rather than distinct peaks in the FTIR profiles.

2.3. Micromorphological analysis

Micromorphological identification and quantification of peat microremains were carried out using a DMR XP Leica photonic microscope under transmitted light. Wet bulk peat samples were mounted as smearslides and analysed with ×20 and ×50 objectives. The surfaces of the main categories were counted (in relative numeric frequencies) through a grid reticule, used as surface unit, and placed on the microscope ocular. Three thousand to 5000 items per sample were counted with an estimated counting error of about 10%.

2.4. C and N determination

Carbon and nitrogen contents were determined by combustion at 1100°C with a CNS-2000 LECO apparatus, on dried and crushed peat samples. Due to the total lack of carbonates, total carbon (TC) was assumed to be total organic carbon (TOC).

2.5. Characterisation of carbohydrate signatures

Carbohydrate analyses were generally performed on fine-grained peat fractions (<200 µm, isolated by wet-sieving at 200 µm under positive pressure using water circulation), although bulk peat samples were also analysed in some cases, for comparison. Cellulosic and hemicellulosic sugars were identified and quantified by gas chromatography after appropriate hydrolysis. Total sugars were determined by hydrolysis after treatment with concentrated acid (see below) whereas labile (hemicellulosic) sugars were determined independently without this treatment.
Cellulosic sugars were determined by difference between the total and hemicellulosic sugars. A detailed procedure is given in Comont et al. (2006). Briefly: for total sugar determination, 1 ml of 24 N H$_2$SO$_4$ was added to 100 mg sample dry weight. After 12 h at room temperature, samples were diluted to 1.2 M H$_2$SO$_4$ and heated at 100°C for 4 h (hemicellulosic sugar analysis begins directly at this hydrolysis stage). After cooling, deoxy-6-glucose was added as an internal standard and samples were neutralised with CaCO$_3$. Precipitate was discarded removed following centrifugation and the supernatant evaporated to dryness. After resuspension in methanol, the solution was purified by centrifugation and the supernatant transferred and evaporated under vacuum. The resulting carbohydrates were dissolved in trimethylsilylated pyridine (Sylon BFT, Supelco) and immediately analysed by GC-FID using a 25 m x 0.25 mm CPSil5CB (0.25 µm film thickness) column. Oven settings were as follows: an initial oven temperature (60°C) was ramped at 30°C min$^{-1}$ to 120°C where it was maintained for 1 min, then ramped to 240°C at 3°C min$^{-1}$ and finally at 20°C min$^{-1}$ up to 310°C where it was maintained for 10 min. The injector split was off at the start time and turned on after 2 min. The injector and detector were maintained at 240°C and 300°C, respectively. A mixture of eight monosaccharides (ribose, arabinose, xylose, rhamnose, fucose, glucose, mannose and galactose) was used as an external standard for compound identification through peak retention times and for individual response coefficient determination.

2.6. Statistical analyses

All statistical analyses were performed using Genstat for Windows (8$^{th}$ edition, VSN International). FTIR spectral data in the diamond interference region (2200-1900 cm$^{-1}$) were excluded from analyses. Relationships between FTIR spectra (‘x’ variate, as zero, first and second order derivatives) and the corresponding organic
matter (micromorphological and carbohydrate signature) datasets as well as carbon and nitrogen contents and their ratios ('y' variates) were assessed using partial least squares (PLS) analyses. We investigated both each parameter separately in univariate PLS and also within multivariate datasets in multivariate PLS, with leave-one-out validation. The Genstat procedure returned the number of latent roots (dimensions), the predicted residual error sum of squares (PRESS), percentage of variance explained and PLS loadings. The number of roots for each PLS analysis was set at the number that returned minimum PRESS. The root mean square error of cross-validation (RMSECV) was calculated from PRESS using the square root of PRESS/n. Significance levels were estimated using Osten’s F-test. Assessment of the predictive qualities of PLS was performed by principal component analysis (PCA) of the observed and predicted values for both micromorphological analyses and plant-derived carbohydrate monomer signatures.

3. Results

3.1. Patterns of FTIR carbon chemistry signatures

Sample characterisation using FTIR spectroscopy concerned the correct assignment of the observed spectral characteristics to the most likely origin of the absorption bands. A summary of the most characteristic bands observed in peat and their assignment is presented in Table 2. Generally, FTIR analysis on the peat horizon samples showed a decline of the main polysaccharide markers (absorption bands around 3400 and 1040 cm\(^{-1}\)) and relative increase of the main bands assigned to lignin-like (1513, 1450, 1371, 1265 and 835 cm\(^{-1}\)) and aliphatic structures (2920 and 2850 cm\(^{-1}\)) with depth, as expected with increasing humification. Figure 1 shows an example from a Scottish peat core at an advanced stage of regeneration. Spectral bands indicative of ‘carboxylates’, which include contributions from vibrations of
aromatic and aliphatic carboxylates (R-COO¯) and/or aromatic C=C structures also increased in relative terms with depth (1650-1600 and 1426 cm⁻¹).

3.2. PLS calibrations with chemical and organic matter parameters

The FTIR data were assessed against the data obtained from organic matter analyses (variation of the original data shown in Table 3) using both univariate and multivariate PLS analyses. In the cases of the elemental and carbohydrate analyses, the percentage variance explained was > 60% for the majority of the analytes (Table 4) and all were highly significant in PLS of the zero order FTIR data. PLS using the first or second derivative did increase the percentage variance explained, but only marginally (data not shown). The RMSECV values (Table 4) were generally lower than the standard deviation of the original datasets (Table 3). Total C content was associated with positive loadings in the main polysaccharide envelopes at 3300 and 1030 cm⁻¹ and negative loadings of the wax markers at 2920 and 2850 cm⁻¹ (Fig. 2A). Total N content was associated with positive loadings in the bands representing the amide I and II regions (Table 2) as previously reported by Chapman et al (2001) and there an additional negative correlation with the wax markers at 2920 and 2850 cm⁻¹ (Fig 2B). The loading plots for hemicellulosic sugars (Fig 2C) showed a similar positive relationship with the main polysaccharide bands but there was also a strong negative relationship with the carboxylate marker region. Loadings generated for fucose (Fig 2D) primarily demonstrated a strong negative correlation with the wax markers at 2920 and 2850 cm⁻¹.

In some cases, the differences in loadings were more subtle. For example, the PLS loadings for mucilage were visually very similar to those of total C content (not shown). We therefore also examined the relative differences in PLS loadings between organic matter parameters using subtraction. Examples of relative differences in PLS
loadings are shown in Fig 3. The relative difference in PLS loadings between those generated for mucilage and those for total C (Fig 3A) showed that both the main polysaccharide envelope around 1100 cm⁻¹ and the wax markers (2920 and 2850 cm⁻¹) were less discriminatory for prediction of mucilage content than for total C. Similarly, for the prediction of structureless Cyperaceae versus preserved Cyperaceae, the wax marker bands and various other bands indicative of more humified tissue (1710-1707, 1650-1600, and 1515-1513 cm⁻¹) were more discriminatory (Fig 3B). For most neutral sugars, however, the main discriminatory region was within the main 1200-800 cm⁻¹ polysaccharide envelope (Fig 3C-F). Within this polysaccharide envelope, there were subtle differences in the bands which contributed more or less to the discrimination between different neutral sugars.

3.3. Predictive PLS of FTIR spectra as a tool in organic matter studies of regenerating peatlands

To assess the potential of predictive PLS of FTIR spectra for organic matter parameters, we reconstructed the composition of the organic matter datasets (i.e. micromorphological fingerprints, carbohydrate signatures) using the PLS outputs. Examples of the changes of the organic matter parameters and correlation with the predicted values with depth are shown for two contrasting sites of the peatland at Le Russey, France (Fig. 4). The fitted and observed values were in close agreement for the carbohydrate parameters.

Over the entire dataset of 68 samples from all sites in Europe, values for the observed versus predicted C:N ratio’s were highly correlated and linear regression explained 70.6 % of the variance (Fig 5A). To test whether PLS analyses were adequate in predicting the organic matter composition of each of the peat samples, we combined the predicted values from univariate PLS analyses for each subset of
multivariate organic matter analyses (i.e. micromorphological and carbohydrates) and analysed these reconstructed organic matter datasets using PCA. We subsequently compared them to PCA of the observed OM characteristics using linear regression of the 1st principal components (Fig. 5B and C). Linear regression explained 51.0 and 82.1% of the variance in comparisons of the first principal components of the micromorphological and carbohydrate composition obtained by univariate PLS models, respectively (Fig. 5B and C). We also tested the predicted values of multivariate PLS. The regressions of multivariate PLS explained a marginally lower percentage of the variance in the datasets (data not shown). The first principal components of the carbohydrate datasets explained the majority of the variance in both the observed and fitted datasets (88.7 and 94.9%, respectively). The observed strength of the regression therefore indicates that PLS is able to predict the carbohydrate signature of a wide range of peat samples.

4. Discussion

4.1. FTIR spectral characteristics and the potential effect of spectral interferences on PLS models

We observed a few samples which showed spectral interference from silicate minerals. A notable example is shown in Fig 1A in the Scottish sample from an advanced stage of regeneration at horizon 6 (22.5-27.5 cm depth), which shows the diagnostic peaks at 3700 and 467 cm\(^{-1}\) of kaolinite. Where mineral interferences manifest themselves in samples obtained from deeper horizons, they may have originated from wind-blown material from nearby exposed mineral surfaces during the formation of the peatlands. We also observed spectral signals from silicate minerals in a few surface samples from sites on, or close to, nearly exhausted peatlands. Mineral interference also manifests itself in the 1030 cm\(^{-1}\) polysaccharide
band (Farmer, 1974) and could therefore potentially skew the accuracy of prediction of the polysaccharide markers. Other notable results were the low relative absorption values in the 1707 cm\(^{-1}\) region in the samples from the Baupte peatland (data not shown). Peat samples from Baupte had consistently higher pH values and we therefore attributed this lack of absorption in the 1707 cm\(^{-1}\) region to the majority of acids being present in the carboxylate form. A reduction in intensity of the 1720 cm\(^{-1}\) absorption band with increasing pH values was been previously shown in FTIR analyses on peat samples where the sample pH was moderated (Gondar et al., 2005). In low pH environments however, for example as observed in the profile of the Scottish peat samples (e.g. Fig.1), this variation is related to peat decomposition rather than pH changes and increases in intensity of this band illustrate progressive free acid release with increasing humification. Analysis without samples that were characterised by spectral interference by silicate minerals, or those with elevated pH values, did not increase the amount of variability explained by PLS (data not shown).

4.2. Principal calibrations

Infrared spectroscopic data from peat samples, both in the mid and near IR ranges, have been used previously to predict various organic matter parameters. Good correlations by PLS between various parameters such as total C and N, pH, ash content, total organic matter etc, and IR analyses of organic soils have been shown on numerous occasions (Palmborg and Nordgren, 1993; Chapman et al., 2001; Tremblay and Gagne, 2002, Couteaux et al., 2003). Some studies have attempted to predict a small range of functional chemical signatures of peat, such as relative concentrations of amino acids and amino sugars (Holmgren and Norden, 1988) or humic and fulvic acids (Tremblay and Gagne, 2002). The use of ATR accessories produces more consistent data and the use of an additional water correction has been shown to
increase the accuracy of total C predictions (Tucker et al., 2001). A possible reason for the generally poor relationships of FTIR data with micromorphological analyses may be the size of sample used in direct microscopic analysis compared to the use of a homogenised sample as for FTIR spectroscopy. Another possible factor is the three-dimensionality of the chemical analyses (as these are based on mass), compared to micromorphological analyses which essentially extrapolate the mass of microremains based on the area they occupy within smearslides. The variance explained by PLS with some of the more prevalent tissues observed (e.g. proportion of preserved *Cyperaceae* and mucilage) are generally more encouraging (Table 4). Other reports have, however, shown the ability to separate major vegetational differences using near infrared spectroscopy, where the authors were able to build good PLS models with high levels of explained variance for the content of leaf material from Ericales and *Sphagnum* spp. in a single peat core sectioned to 1 cm samples (McTiernan et al., 1998). In the near infra-red, the vibrational characteristics of the N-H and O-H stretch regions are more separated, while the mid-IR is dominated by the O-H stretch regions. Use of data from near infrared analyses may therefore improve the prediction of organic matter parameters.

We also observed cases where the variance explained within the carbohydrate parameters was rather low (Table 4). The carbohydrate chemistry was assessed on the fine fraction (< 200 micron) because this size fraction offers greater sensitivity as it is composed of the biodegraded plant material admixed with products of secondary microbial production. The bulk fraction signature is effectively ‘swamped’ by intact or only partially degraded plant tissues (Comont et al., 2006). This may, however, also offer an explanation for the relatively low variance explained for the total cellulosic sugar content as well as the relative amounts of arabinose and xylose by PLS on the bulk sample derived FTIR data. The latter monomers are the principal
biomarkers of sedges (e.g. *Cyperaceae*; Moers et al., 1989; Comont et al., 2006), which in most samples appear to be the dominant preserved tissues (Table 3). Similarly, single correlations for each of the principal biomarkers of intact bryophytes, mannose, rhamnose and galactose (Popper and Fry, 2003) were rather poor (Table 3). FTIR spectroscopy has been successfully used previously to predict the neutral sugar monomers of plant cell wall polysaccharides in foodstuffs, including such monomers as those described as vegetation biomarkers in peatlands. Examples are prediction of mannose content (from mannans and mannoproteins) in wines (Coimbra et al., 2005) and xylose content from olive pulp polysaccharides (Coimbra et al., 1999). That we were able to successfully reassemble the neutral sugar profiles (Fig. 5) of a large variety of peatland samples from five European locations with large differences in plant cover and degree of decomposition despite observing low correlation of each principal monomer with FTIR spectra may at first seem puzzling. Kačuráková et al. (2000) investigated individual plant cell wall compounds (pectic polysaccharides, hemicelluloses and monosaccharides) by FTIR spectroscopy and were able to determine spectral marker regions for a large number of these compounds within the 1200-800 cm\(^{-1}\) region. They attributed the main differences to both C-OH relative steric positions within the monomer side chains as well as the vibrational characteristics of the pyranose backbones. Carbohydrate analysis reports on the individual components of each (hemi)cellulose type present in the more biologically decomposed size fractions, whereas the FTIR spectra report on the spectral properties of both the side-chain monomers and the pyranose backbones. In agreement with Kačuráková et al., we showed in this study that the spectral markers within the 1200-800 cm\(^{-1}\) region were discriminatory for the relative differences in PLS prediction between neutral sugars (Fig. 3). The correlation of the multivariate datasets is
therefore most likely explained by both methodologies essentially reporting on the
original composition of the parent vegetational material.

4.3. Predictive potential of peat FTIR-based PLS models and applications in
restoration monitoring

The main strength of this study is the ability to satisfactorily reconstruct (Fig. 5) the relative differences between peat samples from different stages of regeneration and from widely differing locations with their associated differences in vegetation structures, and hence micromorphological and chemical composition. In peatland restoration projects, it is often difficult to ascertain the regeneration boundary, i.e. the interphase between the cut horizon of remaining catotelm and the accumulated organic matter during regeneration. The C:N ratio and micromorphological composition of peat have been previously used as an indicator of this regeneration boundary by Comont et al., (2006). The composition of carbohydrate monomers is also indicative of the origin of the OM (i.e. vegetation-type specific, Moers et al., 1989: Wicks et al., 1991) and fucose has been proposed as an indicator of microbially produced polysaccharide monomers (Murayama et al., 1988; Comont et al., 2006). A simple ratio of the polysaccharide to carboxylate FTIR band intensity has previously been shown to explain 68.7 % of the community catabolic response to a suite of simple carbon sources (Artz et al., 2006) and other studies have shown good predictive properties of PLS analyses on FTIR and NIR data with microbial biomass (Chapman et al., 1998; Couteaux et al., 2003). There are other reports where IR data have been used in prediction of various organic matter parameters such as cellulose content and cellulose decomposition rates (Hartmann and Appel, 2006) or the relative quantities of phenolic compounds and leaf litter decomposition rates (Stolter et al., 2006). The botanical composition of peat, i.e. the composition of litter entering the
organic matter pool, has been implicated in the degree of decomposition on numerous occasions (Verhoeven and Toth, 1995; Belyea, 1996; Frolking et al., 2001) and therefore the composition of the organic matter in restored peatlands may be critical to their long term carbon sequestration potential. Indeed, Andersen et al. (2006) showed that the substrate quality (in terms of availability of N, P and soluble organic carbon) of the organic matter in vegetational successions on cutover peatlands was related to the evolution of the total microbial biomass and their respiration activity. Therefore, monitoring of restoration projects should include evaluations whether organic matter parameters return to values more closely resembling intact peatland systems. FTIR spectroscopy coupled with predictive PLS analysis may be a useful, low-cost, addition to the toolbox in the assessment and monitoring of restoration success in peatland ecosystems.
Acknowledgements

This work was part of the RECIPE initiative, funded through a grant by the EU Framework 5. SJC is funded by the Scottish Executive Environment and Rural Affairs Department. The authors gratefully acknowledge analytical assistance provided by M. Hatton and N. Lottier. We would like to thank Dr. David Elston (BioSS) and Dr. Charlie Shand (Macaulay Institute) for constructive comments on this manuscript.
References


Figure legends:

**Fig. 1.** FTIR spectra of peat profile samples from the Scottish Site D (an advanced stage of regeneration), ranging from the surface moss layer at 0-5 cm (solid line) through decomposing plant litter at 5-10 cm (dashed line) and highly humified peat at 22.5 - 27.5 cm (dash-dotted line) and at 42.5 - 47.5 cm (dotted line). Relative abundances for each spectral signal were obtained by normalisation of data (see text). Characteristic FTIR bands of the major biochemical descriptors have been marked on the whole spectrum (A) and in the region containing the lignin, carboxylate and peptide markers (B). The left and right insets represent magnified sections of the spectrum of the 22.5 - 27.5 cm horizon. Spectral markers indicative of mineral interference in this sample have been marked with arrows (see text).

**Fig. 2.** Loadings generated by partial least squares analysis of zero-order FTIR-ATR absorbances against A) Total C and B) total N content. Loadings against micromorphological parameters such as percentage preserved (solid line) and structureless (dotted line) *Cyperaceae* tissues and mucilage are shown in Figs. C and D, respectively. Figures E and F present loadings generated against hemicellulosic sugars and fucose, respectively.

**Fig. 3.** Relative differences in PLS loadings between different organic matter parameters as assessed by subtraction of loadings. Variations over the entire FTIR spectral range are shown for the differences in loadings between mucilage and total C content (A) and the difference between structureless versus preserved *Cyperaceae* (B). Relative differences in PLS loadings within the polysaccharide envelope are
characteristic for neutral sugars. Differences in loadings of fucose versus xylose (C), arabinose versus mannose (D), fucose versus ribose (E) and mannose versus rhamnose (F) are shown as examples.

**Fig. 4.** Fitted (open symbols) versus observed (closed symbols) values for various organic matter parameters with depth (cm) for the site in the early stage of regeneration (FR-A) and the intact site (FR-D) in the peatland at Le Russey, France.

**Fig. 5.** Predictive properties of univariate PLS models based on FTIR data to describe the following organic matter properties of peat. **A:** Predicted (x) vs. observed (y) values of C/N ratios **B:** Regression plot of the first dimensions of PCA performed on the PLS predicted (x) versus observed (y) values of the micromorphological remains. Each micromorphological parameter was used in separate univariate PLS analyses and the predicted data were used for PCA based reconstruction. Variance explained for the first dimensions for each PCA are shown in brackets on each axis. **C:** Regression plot of the first dimensions of PCA performed on the PLS predicted (x) versus observed (y) values of the carbohydrate monomer analyses. Each carbohydrate parameter was used in separate univariate PLS analyses and the predicted data were used for PCA based reconstruction. Variance explained for the first dimensions for each PCA are shown in brackets on each axis. Samples from the different countries are indicated by the following symbols: Finnish samples (downward, filled triangles), France Baupte (empty circles), France Russey (upward, empty triangles), Switzerland Chaux d’Abel (filled circles) and Scotland (filled squares). The solid lines indicate the mean regression, the dashed lines indicate the 95% confidence interval and the dotted lines indicate the 95% prediction intervals.
### Table 1. Origin and general characteristics of peat samples

<table>
<thead>
<tr>
<th>Country</th>
<th>Site</th>
<th>Regeneration stage</th>
<th>Dominant vegetation</th>
<th>Years since abandonment (estimate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td>A</td>
<td>Early</td>
<td><em>Eriophorum vaginatum</em></td>
<td>10</td>
</tr>
<tr>
<td>(Aitoneva)</td>
<td>B</td>
<td>Early</td>
<td><em>Eriophorum vaginatum</em></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>Early</td>
<td>Carex rostrata</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>Early</td>
<td>Sphagnum fallax</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>Bare</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>A</td>
<td>Bare</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>(Baupte)</td>
<td>B</td>
<td>Early</td>
<td><em>Eriophorum vaginatum</em></td>
<td>5-10</td>
</tr>
<tr>
<td>France</td>
<td>C</td>
<td>Advanced</td>
<td><em>Sphagnum fallax, Eriophorum angustifolium</em></td>
<td>5</td>
</tr>
<tr>
<td>(Le Russey)</td>
<td>D</td>
<td>Reference</td>
<td>Sphagnum fallax, Eriophorum angustifolium, Calluna vulgaris</td>
<td>Intact</td>
</tr>
<tr>
<td>Switzerland</td>
<td>A</td>
<td>Early</td>
<td>S. fallax (discontinuous), <em>E. vaginatum</em>, Polytrichum commune, <em>P. strictum</em></td>
<td>&gt;25</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>Reference</td>
<td>S. fallax (continuous), <em>E. vaginatum</em>, Polytrichum commune, <em>P. strictum</em>, Vaccinium spp</td>
<td>Intact</td>
</tr>
<tr>
<td>Scotland</td>
<td>A</td>
<td>Early</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>(Middlemuir)</td>
<td>B</td>
<td>Early</td>
<td><em>E. vaginatum</em>, Agrostis canina, Calluna vulgaris, <em>E. angustifolium</em>, <em>S. auriculatum</em>, <em>S. cuspidatum</em></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>Early</td>
<td><em>E. angustifolium</em>, <em>S. auriculatum</em>, <em>E. vaginatum</em>, <em>S. cuspidatum</em></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>Advanced</td>
<td><em>S. palustre, C. vulgaris, E. vaginatum, Erica tetralix, Deschampsia flexuosa, Molinia caerulea</em></td>
<td>50</td>
</tr>
</tbody>
</table>

*a* Determined by assessment of vegetation diversity and depth of newly formed peat

*b* Based on % cover estimates. Only vegetation with >10% cover reported

* Horizon 3 samples not analysed.
<table>
<thead>
<tr>
<th>Wavenumber, cm(^{-1})</th>
<th>Assignment</th>
<th>Characterisation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3340</td>
<td>γ (O-H) stretching</td>
<td>Cellulose, in samples with defined 3340 peak</td>
<td>Cocozza et al., 2003</td>
</tr>
<tr>
<td>2920</td>
<td>antisymmetric CH(_2)</td>
<td>Fats, wax, lipids</td>
<td>Niemeyer et al., 1992; Cocozza et al., 2003</td>
</tr>
<tr>
<td>2850</td>
<td>symmetric CH(_2)</td>
<td>Fats, wax, lipids</td>
<td>Niemeyer et al., 1992; Cocozza et al., 2003</td>
</tr>
<tr>
<td>1720</td>
<td>C=O stretch of COOH or COOR</td>
<td>Carboxylic acids, aromatic esters</td>
<td>Niemeyer et al., 1992; Haberhauer et al., 1998; Cocozza et al., 2003; Gondar et al., 2005</td>
</tr>
<tr>
<td>1710-1707</td>
<td>C=O stretch of COOH</td>
<td>Free organic acids</td>
<td>Gondar et al., 2005</td>
</tr>
<tr>
<td>1653</td>
<td>C=O of amide I</td>
<td>Proteinaceous origin</td>
<td>Ibarra et al., 1996; Zaccheo et al., 2002</td>
</tr>
<tr>
<td>1650-1600</td>
<td>Aromatic C=C stretching and/or asymmetric C-O stretch in COO-</td>
<td>Lignin and other aromatics, or aromatic or aliphatic carboxylates</td>
<td>Niemeyer et al., 1992; Cocozza et al., 2003</td>
</tr>
<tr>
<td>1550</td>
<td>N-H in plane (amide-II)</td>
<td>Proteinaceous origin</td>
<td>Ibarra et al., 1996; Zaccheo et al., 2002</td>
</tr>
<tr>
<td>1515-1513</td>
<td>Aromatic C=C stretching</td>
<td>Lignin/Phenolic backbone</td>
<td>Cocozza et al., 2003</td>
</tr>
<tr>
<td>1426</td>
<td>Symmetric C-O stretch from COO- or stretch and OH deformation (COOH)</td>
<td>Carboxylate/Carboxylic structures (humic acids)</td>
<td>Parker, 1971</td>
</tr>
<tr>
<td>1450, 1371</td>
<td>C-H deformations</td>
<td>Phenolic (lignin) and aliphatic structures</td>
<td>Parker, 1971</td>
</tr>
<tr>
<td>1265 (approximately)</td>
<td>C-O stretching of phenolic OH and/or arylmethylene</td>
<td>Indicative of lignin backbone</td>
<td>Niemeyer et al., 1992; Ibarra et al., 1996</td>
</tr>
<tr>
<td>1080-1030</td>
<td>Combination of C-O stretching and O-H deformation</td>
<td>Polysaccharides</td>
<td>Grube et al., 2006</td>
</tr>
<tr>
<td>900</td>
<td>Out of phase ring stretching (ring ‘breathing’)</td>
<td>Cellulose, corresponding band to sharpened 3340 peak</td>
<td>Zaccheo et al., 2002</td>
</tr>
<tr>
<td>720</td>
<td>CH(_2) wag</td>
<td>Long chain (&gt;C4) alkanes</td>
<td>Ibarra et al., 1996</td>
</tr>
<tr>
<td>835</td>
<td>Aromatic CH out of plane</td>
<td>Lignin</td>
<td>Zaccheo et al., 2002</td>
</tr>
</tbody>
</table>
Table 3. Organic matter parameters measured with average, standard deviation and range observed.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Average</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elemental</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total C (%)</td>
<td>52.5</td>
<td>4.0</td>
<td>42.9-60.6</td>
</tr>
<tr>
<td>Total N (%)</td>
<td>1.9</td>
<td>0.6</td>
<td>0.9-3.2</td>
</tr>
<tr>
<td>C/N ratio</td>
<td>30.6</td>
<td>9.9</td>
<td>15.1-59.7</td>
</tr>
<tr>
<td><strong>Micromorphology</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amorphous organic matter (%)</td>
<td>20.7</td>
<td>15.6</td>
<td>0.9-71.8</td>
</tr>
<tr>
<td>Preserved <em>Cyperaceae</em> (%)</td>
<td>9.6</td>
<td>19.0</td>
<td>0-79.8</td>
</tr>
<tr>
<td>Preserved <em>Sphagnum</em> (%)</td>
<td>4.6</td>
<td>6.9</td>
<td>0-27.3</td>
</tr>
<tr>
<td>Preserved <em>Polytrichum</em> (%)</td>
<td>1.1</td>
<td>3.9</td>
<td>0-22.1</td>
</tr>
<tr>
<td>Undetermined preserved tissue (%)</td>
<td>2.6</td>
<td>4.6</td>
<td>0-19.1</td>
</tr>
<tr>
<td>Structureless <em>Cyperaceae</em> (%)</td>
<td>4.5</td>
<td>9.4</td>
<td>0-41.8</td>
</tr>
<tr>
<td>Structureless <em>Sphagnum</em> (%)</td>
<td>0.5</td>
<td>1.3</td>
<td>0-5.2</td>
</tr>
<tr>
<td>Structureless <em>Polytrichum</em> (%)</td>
<td>0.1</td>
<td>1.0</td>
<td>0-8.1</td>
</tr>
<tr>
<td>Undetermined structureless (%)</td>
<td>10.9</td>
<td>9.9</td>
<td>0.2-43.6</td>
</tr>
<tr>
<td>Mucilage (%)</td>
<td>40.8</td>
<td>24.1</td>
<td>0-94.4</td>
</tr>
<tr>
<td>Microorganisms (%)</td>
<td>3.1</td>
<td>2.4</td>
<td>0.1-12.9</td>
</tr>
<tr>
<td>Cuticles/Spores/Pollen (%)</td>
<td>1.1</td>
<td>3.6</td>
<td>0-28.4</td>
</tr>
<tr>
<td>Gelified or oxidised debris (%)</td>
<td>0.1</td>
<td>0.4</td>
<td>0-2.8</td>
</tr>
<tr>
<td><strong>Carbohydrates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total sugars (mg g⁻¹)</td>
<td>134.9</td>
<td>60.2</td>
<td>40.6-346.5</td>
</tr>
<tr>
<td>Hemicellulosic sugars (mg g⁻¹)</td>
<td>86.9</td>
<td>40.1</td>
<td>27.2-179.9</td>
</tr>
<tr>
<td>Cellulosic sugars (mg g⁻¹)</td>
<td>47.9</td>
<td>28.2</td>
<td>10.7-173.2</td>
</tr>
<tr>
<td>Xylose (%)</td>
<td>19.4</td>
<td>4.1</td>
<td>10.5-31.0</td>
</tr>
<tr>
<td>Arabinose (%)</td>
<td>6.6</td>
<td>4.9</td>
<td>1.7-19.6</td>
</tr>
<tr>
<td>Hemicellulosic glucose (%)</td>
<td>38.2</td>
<td>11.3</td>
<td>16.4-58.5</td>
</tr>
<tr>
<td>Mannose (%)</td>
<td>10.4</td>
<td>1.8</td>
<td>6.3-16.4</td>
</tr>
<tr>
<td>Rhamnose (%)</td>
<td>7.5</td>
<td>1.8</td>
<td>2.7-12.6</td>
</tr>
<tr>
<td>Galactose (%)</td>
<td>6.4</td>
<td>6.5</td>
<td>0.2-26.9</td>
</tr>
<tr>
<td>Ribose (%)</td>
<td>2.0</td>
<td>1.8</td>
<td>0-6.3</td>
</tr>
<tr>
<td>Fucose (%)</td>
<td>9.5</td>
<td>6.2</td>
<td>1.8-23.2</td>
</tr>
</tbody>
</table>
Table 4. Partial least squares analysis results of zero-order FTIR spectra vs. peat micromorphological and chemical properties.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>No. of latent roots in univariate PLS (variance accounted for, $r^2$ in %)</th>
<th>RMSECV $^a$ in univariate PLS</th>
<th>No. of latent roots in multivariate PLS (variance accounted for, $r^2$ in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elemental</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total C (%)</td>
<td>6 (81.6) *** $^b$</td>
<td>2.1</td>
<td>6 (75.6)</td>
</tr>
<tr>
<td>Total N (%)</td>
<td>6 (68.8) *** $^b$</td>
<td>0.4</td>
<td>6 (67.5)</td>
</tr>
<tr>
<td>C/N ratio</td>
<td>6 (70.6) *** $^b$</td>
<td>6.4</td>
<td>6 (70.4)</td>
</tr>
<tr>
<td><strong>Micromorphology</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amorphous organic matter (%)</td>
<td>1 (9.0) *** $^b$</td>
<td>15.2</td>
<td>2 (8.9)</td>
</tr>
<tr>
<td>Preserved Cyperaceae (%)</td>
<td>2 (51.0) *** $^b$</td>
<td>14.3</td>
<td>2 (50.4)</td>
</tr>
<tr>
<td>Preserved Sphagnum (%)</td>
<td>1 (8.0) ***</td>
<td>6.8</td>
<td>2 (7.7)</td>
</tr>
<tr>
<td>Preserved Polytrichum (%)</td>
<td>1 (19.3) *** $^b$</td>
<td>3.7</td>
<td>2 (22.9)</td>
</tr>
<tr>
<td>Undetermined preserved tissue (%)</td>
<td>1 (2.4) ns $^b$</td>
<td>4.7</td>
<td>2 (2.6)</td>
</tr>
<tr>
<td>Structureless Cyperaceae (%)</td>
<td>1 (16.5) *** $^b$</td>
<td>8.7</td>
<td>2 (16.4)</td>
</tr>
<tr>
<td>Structureless Sphagnum (%)</td>
<td>1 (9.2) **</td>
<td>1.3</td>
<td>2 (0.2)</td>
</tr>
<tr>
<td>Structureless Polytrichum (%)</td>
<td>1 (4.1) ns $^b$</td>
<td>1.0</td>
<td>2 (5.4)</td>
</tr>
<tr>
<td>Undetermined structureless (%)</td>
<td>1 (9.7) **</td>
<td>9.6</td>
<td>2 (6.4)</td>
</tr>
<tr>
<td>Mucilage (%)</td>
<td>2 (40.7) *** $^b$</td>
<td>19.7</td>
<td>2 (37.8)</td>
</tr>
<tr>
<td>Microorganisms (%)</td>
<td>1 (11.4) **</td>
<td>2.3</td>
<td>2 (11.7)</td>
</tr>
<tr>
<td>Cuticles/Spores/Pollen (%)</td>
<td>1 (10.9) ns $^b$</td>
<td>3.7</td>
<td>2 (10.7)</td>
</tr>
<tr>
<td>Gelified or oxidised debris (%)</td>
<td>1 (12.1) *** $^b$</td>
<td>0.4</td>
<td>2 (2.9)</td>
</tr>
<tr>
<td><strong>Carbohydrates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total sugars (mg g$^{-1}$)</td>
<td>4 (81.2) *** $^b$</td>
<td>29.5</td>
<td>4 (80.7)</td>
</tr>
<tr>
<td>Hemicellulosic sugars (mg g$^{-1}$)</td>
<td>4 (83.5) *** $^b$</td>
<td>17.7</td>
<td>4 (82.9)</td>
</tr>
<tr>
<td>Cellulosic sugars (mg g$^{-1}$)</td>
<td>2 (43.8) *** $^b$</td>
<td>22.2</td>
<td>4 (46.9)</td>
</tr>
<tr>
<td>Xylose (%)</td>
<td>3 (41.2) *** $^b$</td>
<td>3.4</td>
<td>4 (36.9)</td>
</tr>
<tr>
<td>Arabinose (%)</td>
<td>2 (45.7) *** $^b$</td>
<td>3.8</td>
<td>4 (41.4)</td>
</tr>
<tr>
<td>Hemicellulosic glucose (%)</td>
<td>4 (60.1) *** $^b$</td>
<td>7.8</td>
<td>4 (47.8)</td>
</tr>
<tr>
<td>Mannose (%)</td>
<td>3 (17.7) *** $^b$</td>
<td>1.8</td>
<td>4 (8.2)</td>
</tr>
<tr>
<td>Rhamnose (%)</td>
<td>4 (46.3) *** $^b$</td>
<td>1.6</td>
<td>4 (21.0)</td>
</tr>
<tr>
<td>Galactose (%)</td>
<td>2 (59.3) *** $^b$</td>
<td>4.3</td>
<td>4 (58.0)</td>
</tr>
<tr>
<td>Ribose (%)</td>
<td>4 (64.0) *** $^b$</td>
<td>1.2</td>
<td>4 (57.5)</td>
</tr>
<tr>
<td>Fucose (%)</td>
<td>4 (84.9) *** $^b$</td>
<td>2.8</td>
<td>4 (81.8)</td>
</tr>
</tbody>
</table>

$^a$ Root mean square error of cross-validation in univariate PLS. $^b$ Significance at $p < 0.05$ - *, $p < 0.01$ - **, $p < 0.001$ - ***, ns - not significant.