Catena 167 (2018) 18–27

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena

Using laboratory-based hyperspectral imaging method to determine carbon functional group distributions in decomposing forest litterfall

Iman Tahmasbian[a,*], Shahla Hosseini Bai[b], Yuzhe Wang[c], Sue Boyd[a], Jun Zhou[d], Roya Esmaeilani[e], Zhihong Xu[a,*]

[a] Environmental Futures Research Institute, School of Environment and Science, Griffith University, Nathan, Brisbane, Queensland 4111, Australia
[b] Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore DC, QLD 4558, Australia
[c] Key Laboratory of Pollution Processes and Environmental Criteria, College of Environmental Science and Engineering, Nankai University, Tianjin 300071, People's Republic of China
[d] School of Information and Communication Technology, Nathan Campus, Griffith University, 170 Kessels Road, QLD 4111, Australia
[e] Faculty of Computing, University Technology Malaysia, 81300 Skudai, Johor, Malaysia

ARTICLE INFO

Keywords:
Chemometrics
Decomposition
Toohey Forest
ANN
PLSR
NMR

ABSTRACT

Studying C functional group distributions in decomposing litterfall samples is one of the common methods of studying litterfall decomposition processes. However, the methods of studying the C functional group distributions, such as ^{13}C NMR spectroscopy, are expensive and time consuming and new rapid and inexpensive technologies should be sought. Therefore, this study examined whether laboratory-based hyperspectral image analysis can be used to predict C functional group distributions in decomposing litterfall samples. Hyperspectral images were captured from ground decomposing litterfall samples in the visible to near infrared (VNIR) spectral range of 400–1000 nm. Partial least-square regression (PLSR) and artificial neural network (ANN) models were used to correlate the VNIR reflectance data measured from the litterfall samples with their C functional group distributions determined using ^{13}C NMR spectroscopy. The results showed that alkyl-C, *O*,N-alkyl-C, di-*O*-alkyl-C_1, di-*O*-alkyl-C_2, aryl-C_1, aryl-C_2 and carboxyl derivatives could be acceptably predicted using the PLSR model, with R^2 values of 0.72, 0.73, 0.71, 0.74, 0.76, 0.75 and 0.63 and ratio of prediction to deviation (RPD) values of 1.86, 1.82, 1.78, 1.71, 1.90, 1.76 and 1.43, respectively. Predicted *O*,N-alkyl-C, di-*O*-alkyl-C_1, di-*O*-alkyl-C_2, aryl-C_1 and aryl-C_2 using the ANN model provided R^2 values of 0.62, 0.68, 0.69, 0.82 and 0.67 and the RPDs of 1.54, 1.76, 1.52, 2.10 and 1.72, respectively. With the exception of aryl-C_1, the PLSR model was more reliable than the ANN model for predicting C functional group distributions given limited amount of training data. Neither the PLSR nor the ANN model could predict the carbohydrate-C and *O*-aryl-C acceptably. Overall, laboratory-based hyperspectral imaging in combination with the PLSR modelling can be recommended for the analysis of C functional group distribution in the decomposing forest litterfall samples.

1. Introduction

Litterfall decomposition is an ecological process affecting carbon (C) and nutrient cycling in forest ecosystems (Swift et al., 1979; Trumbore and Czimczik, 2008; Wang et al., 2015). Litterfall decomposition improves soil quality through the addition of organic matter enhancing soil physicochemical properties such as soil water holding capacity and filtration, soil biodiversity and soil microorganism activity (Bossa et al., 2005; Ngoran et al., 2006; Wang et al., 2015). Litterfall decomposition rate is considered to be influenced by different factors such as litterfall quality, which can be investigated through analysing litterfall C functional groups (Bejarano et al., 2014; Bonanomi et al., 2013; Finn et al., 2015; Knorr et al., 2005; Xiao et al., 2013). Significant correlations have been reported among litterfall C functional groups including alkyl-C, N-alkyl-C, *O*-alkyl-C, di-*O*-alkyl-C and litterfall decomposition rate (Bonanomi et al., 2013). Studying the C functional groups in decomposing litterfall is therefore relevant to studies addressing C and nutrient cycling in forest ecosystems.

Solid-state ^{13}C nuclear magnetic resonance (^{13}C NMR) spectroscopy has been used successfully to examine the C cycling and to study the effects of litterfall quality on decomposition processes (Almendros et al., 2000; Baldock et al., 2004; Bell-Dereske et al., 2016; Incerti et al., 2017; Mathers et al., 2007). However, ^{13}C NMR spectroscopy is an expensive and time-consuming technique for studying large numbers of

* Corresponding authors.
E-mail addresses: Iman.tahmasbian@griffithuni.edu.au (I. Tahmasbian), Zhihong.xu@griffith.edu.au (Z. Xu).

https://doi.org/10.1016/j.catena.2018.04.023
Received 13 August 2017; Received in revised form 17 April 2018; Accepted 20 April 2018

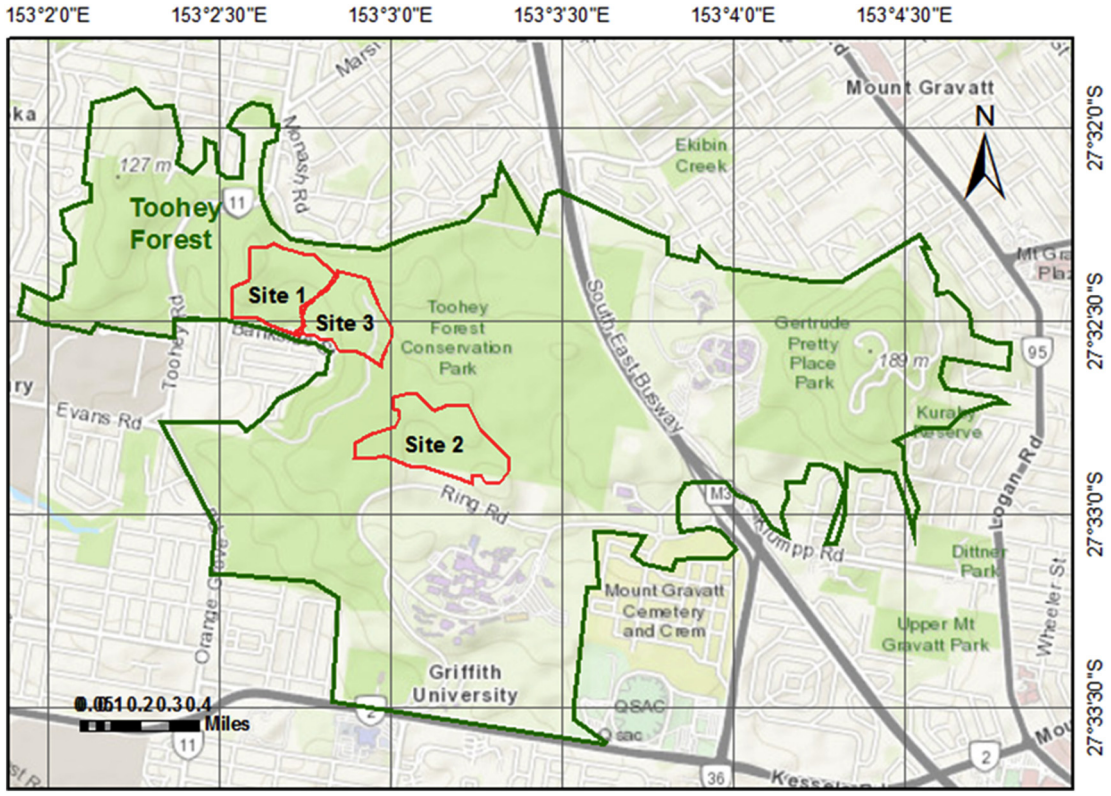

Fig. 1. The study area and location of the three study sites at Toohey Forest, Queensland, Australia.

samples. Therefore, it is important to develop complementary, rapid, inexpensive and reliable technologies for the study of decomposing litterfall quality.

Hyperspectral imaging, which utilises the advantages of both conventional imaging and spectroscopy in a single instrument, has emerged as one of the most promising methods for rapid, facile and non-destructive evaluation of diverse materials (Elmasry et al., 2012; Manley, 2014). Hyperspectral imaging systems use a wide range of electromagnetic spectra, from ultraviolet to infrared, to acquire information on both the spatial and the spectral distribution of the materials (Elmasry et al., 2012; Koehler et al., 2002; Manley, 2014). Hyperspectral reflectance values are correlated with reference values using multivariate analyses, such as linear and non-linear regression models (Kamruzzaman et al., 2015; Manley, 2014; Tahmasbian et al., 2017). Dimensionality reduction or feature selection (band selection) techniques, such as principal component analyses (PCA), stepwise feature selection and particle swarm optimisation (PSO), are usually employed prior to multivariate analyses. This is due to the high dimensionality of the hyperspectral data, with many non-informative bands (Cen et al., 2016; Fenzández et al., 2016; Naganathan et al., 2016; Qi et al., 2017; Yang et al., 2007).

After selecting the most important bands by using an appropriate band selection technique, the raw data can then be subjected to multivariate analyses directly. However, transformation of spectral data is often required to reduce the impact of artefacts, such as light scattering, random noise, and baseline shift effects (Kamruzzaman et al., 2015; Manley, 2014). The most common data transformation methods include first and second derivatives, normalization and scaling, multiplicative scatter correction (MSC), standard normal variate (SNV) and orthogonal signal correction (OSC) (Fearn, 2000; Kamruzzaman et al., 2016; Rinnan et al., 2009; Siripatrawan et al., 2011; Tahmasbian et al., 2017). The transformed data are then used to develop the multivariate model. This step employs appropriate chemometric algorithms, such as partial least squares regression (PLSR) and artificial neural network (ANN), as linear and non-linear calibration methods, respectively (Balabin et al., 2011; Guo et al., 2016; Kuang et al., 2015).

The PLSR model has been used, as linear regression algorithms, in chemometrics for a variety of subjects (Antonucci et al., 2010; Higashikawa et al., 2014; Kamruzzaman et al., 2016; Zhu et al., 2016). Unlike multiple linear regression, PLSR can analyse data which are strongly correlated with multiple variables (Höskuldsson, 1988; Wold et al., 1984; Wold et al., 2001). The PLSR model performs well with noisy and small datasets and chooses the response variable by generating independent latent variables (LV) from the strongest prediction of cross-validation (Höskuldsson, 1988; Pirouz, 2006; Wold et al., 1984; Wold et al., 2001). The ANN model, inspired by the animal brain, is a powerful machine learning algorithm and intelligent prediction tool used in chemometrics (Balabin et al., 2011; Guo et al., 2016; Larose and Larose, 2014). The ANN model can work well when using noisy and uninformative datasets, as it assigns weights to each artificial neurone connection (Larose and Larose, 2014).

Hyperspectral imaging technology has been used to study a diverse range of systems, including medical and pharmaceutical systems, forest ecosystems and food and agricultural products (Kamruzzaman et al., 2016; Lin et al., 2015; Gama et al., 2018; Song et al., 2012; Sytar et al., 2017; Tahmasbian et al., 2017). To the best of our knowledge, however, this is the first paper studying the possibility of using a laboratory-based hyperspectral image analysis and multivariate regression models to predict C functional group distribution in decomposing litterfall samples.

2. Methods and materials

2.1. Study site description

The decomposing litterfall samples used in the current experiment were taken from a litterfall decomposition experiment conducted in Toohey Forest, a suburban native forest located in south-eastern