

# SCIENTIFIC REPORTS



OPEN

## Hyperspectral near infrared imaging quantifies the heterogeneity of carbon materials

Mikko Mäkelä<sup>1,2</sup> & Paul Geladi<sup>2</sup>

For many applications heterogeneity is a direct indicator of material quality. Reliable determination of chemical heterogeneity is however not a trivial task. Spectral imaging can be used for determining the spatial distribution of an analyte in a sample, thus transforming each pixel of an image into a sampling cell. With a large amount of image pixels, the results can be evaluated using large population statistics. This enables robust determination of heterogeneity in biological samples. We show that hyperspectral imaging in the near infrared (NIR) region can be used to reliably determine the heterogeneity of renewable carbon materials, which are promising replacements for current fossil alternatives in energy and environmental applications. This method allows quantifying the variation in renewable carbon and other biological materials that absorb in the NIR region. Reliable determination of heterogeneity is also a valuable tool for a wide range of other chemical imaging applications.

For many applications variation or heterogeneity is a direct indicator of material quality. Reliable estimation of chemical heterogeneity is however not a trivial task. The material needs to be correctly sampled to describe its true time or space dependent variation, which requires obtaining and analysing a sufficient number of samples while minimizing the uncertainty of the analytical procedure. Although an entire theory of representative sampling has been developed for this purpose<sup>1,2</sup>, fast and reliable quantification of heterogeneity from individual samples or real-time monitoring of entire material streams remains challenging.

Spectral imaging can combine e.g. vibrational spectroscopy with the spatial attributes of an image. Mathematical operations can thus be performed in both the spectral and spatial domains, which enables relating chemical information of different analytes with individual image pixels. As an example, vibrational modes of different molecular bonds can be measured within an entire image area through absorbance in the near infrared (NIR) or infrared (IR) region. As opposed to fundamental vibrational modes detected with IR, NIR wavelengths provide information on the overtones and combination bands of hydrogen and other bonds. Molecular absorptivity in the NIR region is thus orders of magnitude weaker than in the IR, which requires longer path lengths for obtaining undistorted spectra, but provides an advantage in sample preparation and in analysing larger heterogeneous samples<sup>3,4</sup>. Spectral interpretation in the NIR is however more complicated and is often performed through calibration based on a known reference method. The term hyperspectral imaging is generally used when spectral images are continuously recorded on tens or even several hundred different wavelengths<sup>5</sup>.

After a calibration model has been determined based on the known concentrations of calibration samples, the properties of individual image pixels can be predicted. In addition to the average concentration of an analyte in a sample, spectral imaging thus enables determining its spatial distribution<sup>6</sup>. Each pixel then becomes a sampling cell. With images that contain a large number of sample pixels, the results can be evaluated using large population statistics<sup>7</sup>. This enables developing a reliable variation metric for describing material heterogeneity.

Although spectral imaging has its roots in remote sensing<sup>3,6,8</sup>, it has been finding its way also to the chemical and materials science communities. As an example, in-line hyperspectral NIR imaging has recently been used for studying the spatial properties of adhesive layers in textile laminates<sup>9</sup>. NIR imaging below the diffraction limit has recently been reported for studying the behavior of DNA walkers<sup>10</sup>. NIR and Raman imaging have also been used for identification of counterfeit drugs<sup>11,12</sup>. Wilczyński *et al.*<sup>12</sup> used the analysis of gray value intensities for quantitative evaluation of heterogeneity in counterfeit tablets. No calibration model was determined and differences

<sup>1</sup>Aalto University, Department of Bioproducts and Biosystems, PO Box 16300, 00076, Aalto, Finland. <sup>2</sup>Swedish University of Agricultural Sciences, Department of Forest Biomaterials and Technology, Skogsmarksgränd, 90183, Umeå, Sweden. Correspondence and requests for materials should be addressed to M.M. (email: [mikko.makela@aalto.fi](mailto:mikko.makela@aalto.fi)) or P.G. (email: [paul.geladi@slu.se](mailto:paul.geladi@slu.se))

Experiment	Carbonization temperature (°C)	Moisture content (kg H <sub>2</sub> O kg <sup>-1</sup> db)	Dry solids (%)
1	180	10	8.9
2	260	10	8.9
3	180	4.4	19
4	260	4.4	19
5	180	7.3	12
6	260	7.3	12
7	220	10	8.9
8	220	4.4	19
9	220	7.3	12
10	220	7.3	12
11	220	7.3	12

**Table 1.** The hydrothermal experiments. db = dry basis.

in heterogeneity were based on spatial and spectral information alone. Recently de Moura França *et al.*<sup>13</sup> also illustrated the use of homogeneity curves for analyzing heterogeneity of binary images.

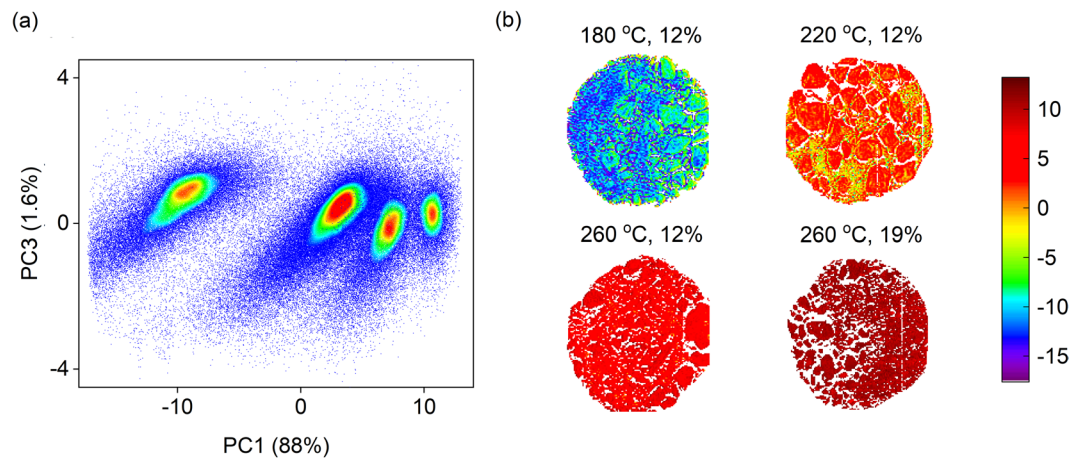
Here, we demonstrate for the first time the use of hyperspectral NIR imaging for quantifying the heterogeneity of renewable carbon materials after multivariate image regression. Biomass-derived carbon is a promising material as it can be prepared from renewable feedstocks and used for replacing current fossil alternatives in a range of energy and environmental applications<sup>14</sup>. The chemical and physical properties of carbon materials can be fine-tuned for a specific application through the use of additives or subsequent activation steps<sup>15,16</sup>. Independent of the choice of method or final application, variation in the properties of the carbon material affect the quality of the final product. In the future, hyperspectral NIR imaging and heterogeneity estimates can be used for monitoring the quality of renewable carbon and other biological materials that absorb in NIR region.

## Methods

**Carbon materials.** Carbon materials were hydrothermally prepared from chemical sludge produced at a wastewater treatment plant of a pulp and paper mill. An experimental design was used to systematically produce carbon materials with different properties. In the design carbonization temperature (180–260 °C) and the moisture content of sludge (10.2–4.4 kg H<sub>2</sub>O kg<sup>-1</sup> db) were varied on three different levels. A total of 11 experiments were performed including three repeated center-points (Table 1). The hydrothermal experiments were performed in a 280 mL autoclave (Büchi Limbo, Büchi AG) using 200 g of wet sludge. The reactor was purged with 2 MPa nitrogen and heated to treatment temperature under autogenous pressure followed by an isothermal holding time of 1 h. Once the reactor was cooled to room temperature, the non-condensable gases were discarded and the slurry was vacuum-filtered using a 25 µm pore size Whatman filter paper. The filtered carbon material was dried at 105 °C overnight. The ash and carbon contents of the dried samples were determined according to respective European standards SS-EN 14775 and SS-EN 15407. The final carbon content was given as a percentage of dry, ash-free (daf) char. After the analyses, the samples were sieved through a 250 µm sieve (Retsch GmbH), which provided a total of 33 samples in three different size fractions including the original unsieved material.

**Spectral imaging.** Hyperspectral imaging was performed with a Specim (Specim, Spectral Imaging, Ltd.) camera equipped with an OLES30 lens provided by Specim. Three images were taken, one for each size fraction. Two rows of quartz halogen lamps generated polychromatic light and the reflected wavelengths were separated by a grating-prism monochromator followed by a HgCdTe detector array. The spectral range was limited to 950–2550 nm with a spectral resolution of approximately 6 nm. The samples were placed in black spherical polypropylene screw caps with a diameter of 23 mm and a depth of 8 mm. The camera was operated in line-scanning mode where a line of 388 pixels was continuously recorded on 288 wavelengths. The samples moved under the camera on a moving belt generating an image. The number of lines and samples on the moving belt thus determined image size. The acquisition time was 10 ms per line, resulting in approximately 14 s per image. The field of view was set to 50 mm, which resulted in a pixel size of 0.13 × 0.13 mm. The absorbance in each pixel was calculated based on measured Spectralon white reference and dark current intensities.

**Image and data analysis.** Backgrounds, sample holders and potential dead pixels were removed from the individual images using principal component analysis (PCA)<sup>17</sup>. The hyperspectral image of each unsieved sample was split into four individual 90 degree segments, which resulted in a total of 44 calibration samples on 276 wavelengths. The respective median spectra were calculated and further split into 33 calibration and 11 validation objects. A calibration model was then determined based on partial least squares (PLS) regression<sup>18</sup> on determined reference values. Spectral preprocessing was performed through the standard normal variate (SNV) transformation and mean centering to minimize the effects of light scattering. The reference carbon values were mean centered. The root mean squared errors of prediction (RMSEP) based on the validation set were calculated as:



**Figure 1.** (a) Principal components scores of the cleaned sample image and (b) a score image of four samples based on the first principal component which explained 88% of data variation. Samples labels in (b) show carbonization conditions given in Table 1.

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

where  $y_i$  and  $\hat{y}_i$  denote the measured and predicted values, respectively, and  $n$  the number of predictions. In addition, prediction bias was calculated as:

$$Bias = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (2)$$

RMSEP was used to describe model range error ratio (RER) and residual prediction deviation (RPD) parameters through:

$$RER = \frac{(y_{max} - y_{min})}{RMSEP} \quad (3)$$

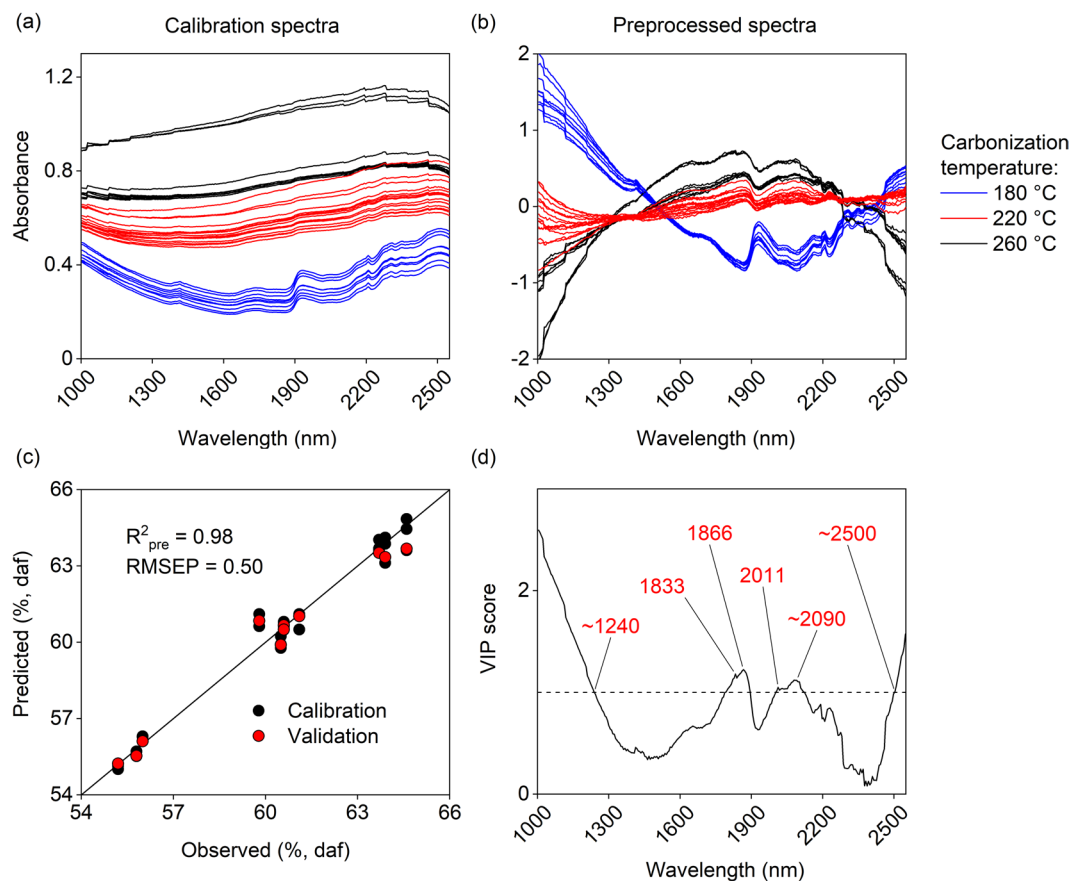
$$RPD = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}} RMSEP^{-1} \quad (4)$$

Data analysis and plotting were performed with the Evince (Prediktera AB), Matlab (The Mathworks, Inc.) PLS Toolbox (Eigenvector Research, Inc.) and OriginPro (Originlab Corp.) software packages. Data availability; the datasets analysed during this study are available from the corresponding author on reasonable request.

## Results and Discussion

Backgrounds and unnecessary pixels were removed from the images with the help of PCA. The first three principal components explained 96–97% of variation in the cleaned images. The first component described the mean spectra, while the second and third components included more detailed spectral features. Based on the score images, the first component described differences between the different samples as the second and third mainly included variation within the individual samples due to e.g. uneven illumination of the sample surfaces. PCA scores of the first and third component showed that the carbon samples contained four different groups, which were mainly separated by carbonization temperature, Fig. 1a. The samples carbonized at 260 °C were further separated into two groups. The effect of carbonization temperature was also clearly visible in the score images (Fig. 1b). This indicated that their chemical differences were visible in the NIR spectra. Once the images had been cleaned, the samples were split into four different parts and the respective median spectra were used for calibration against the reference carbon contents of the samples. The first wavelengths within 950–1000 nm were discarded as they mainly consisted of noise.

The calibration spectra before preprocessing showed increased loss of spectral features with increasing carbonization temperature, Fig. 2a. Especially the peaks at approximately 1410, 1740, 1930 and 2510 nm that correspond with the stretching of O-H, S-H, C-O and C-H from C-H and CH<sub>2</sub> groups<sup>19,20</sup> completely disappeared from the samples produced at higher temperatures. A calibration model was however successfully determined based on preprocessed spectra (Fig. 2b). The final PLS model was composed of 4 latent variables and explained 98% of variation in the calibration and validation data (Fig. 2c). The determined RMSEP were 0.50% carbon (daf) combined with low prediction bias (−0.13). The model RER and ratio RPD parameters related the prediction errors to the original distribution of reference values and were 19 and 6.6, respectively. This indicated that the model performed well and was suitable for quality control applications<sup>3</sup>.



**Figure 2.** (a) Original and (b) preprocessed calibration spectra. Preprocessing was based on SNV transformation and mean centering. (c) Predicted vs. observed carbon contents based on the final PLS model, where the 45° line illustrates a perfect fit. (d) VIP scores of the PLS model.

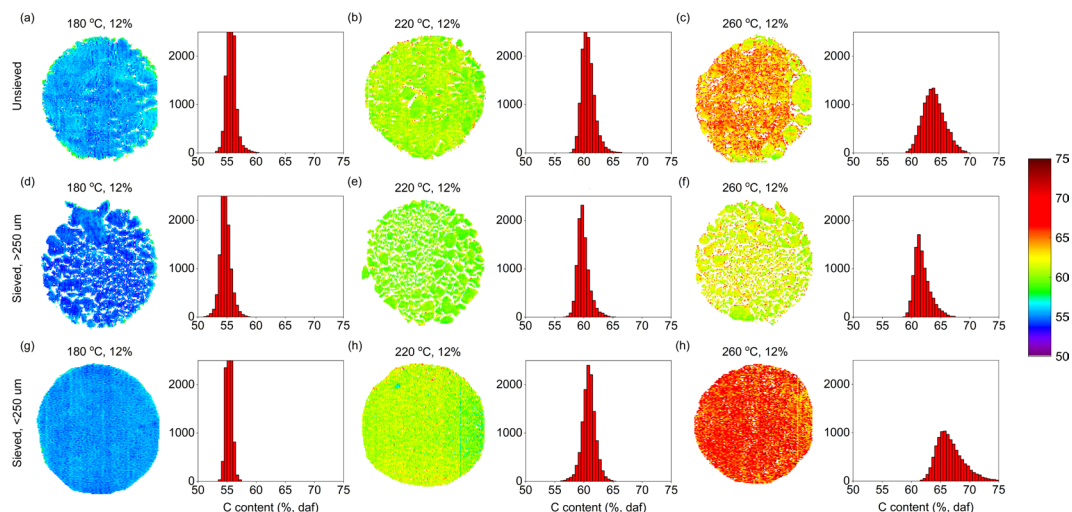
PLS VIP scores based on the model suggested that wavelengths below 1240, within 1830–1870 and 2010–2090, and above 2500 nm were especially important for the predictive ability of the model (Fig. 2d). Wavelengths below 1240 nm generally contain the second overtones of C-H strains of CH and CH<sub>2</sub> groups and the CH<sub>3</sub> groups of aromatic moieties<sup>19</sup>. The preprocessed calibration spectra showed significant variation between the different sample groups in this range (Fig. 2b), which was also indicated by the loading vector of the first latent variable explaining 94% of the variation in the reference values. Shorter NIR wavelengths are also more strongly affected by light scattering and sample color. The effects of scattering are however complicated due to the presence of absorbing constituents and variation in physical properties such as particle size and shape, surface properties and sample packing<sup>21,22</sup>. Increasing carbonization temperature generally leads to a higher carbon content and a darker sample color, which can manifest as increased absorbance on shorter wavelengths<sup>23</sup>. The O-H and C-O strains and the second overtone of C-O strain can normally be detected at 1830 nm<sup>20</sup>. The second overtone of the C=O strain and the O-H deformation and strain generally absorb within 2030–2080 nm<sup>19</sup>. Absorbance of C-O strains and C-H strains from C-H and CH<sub>2</sub> groups have been reported at 2506–2519 nm<sup>19</sup>. Based on additional chemical analyses, the O/C ratio of the feed material decreased from 0.61 (daf) to 0.29–0.46 (daf) during the experiments, which explains changes in the absorbance of C-O and C=O bonds. The respective H/C ratio decreased slightly from 1.7 (daf) to 1.4–1.6 (daf) suggesting changes in C-H and possibly O-H bonds likely through the decomposition of carbohydrates. It should be noted that the spectral resolution of hyperspectral cameras is generally not as high as those of spectrometers, which can generate a level of uncertainty to the interpretation of spectral assignments. The PLS model however enabled reliable calibration as indicated by the determined performance parameters.

Once the suitable calibration model was obtained, the carbon contents of individual image pixels were predicted. Two more images were obtained after sieving the original samples through a 250 μm sieve. The sieving enabled testing the hypothesis that variation in the different size fractions could be determined based on the predicted pixel data. The pixel populations of individual samples were further processed to remove potential outliers. Observations that situated >|3| standard deviations from the population mean based on a normal distribution were excluded. This still enabled using 99.7% of the original pixel data.

In general, heterogeneity can be used for describing the chemical or physical properties of a material. Chemical heterogeneity generally relates to composition, while physical heterogeneity is often related to particle size or shape distribution. For many materials heterogeneity is a mixture of the two. We described heterogeneity of the samples through the concept of constitutional heterogeneity ( $h_c$ ), originally developed within the theory of

Source	Sum of squares	Degrees of freedom	Mean square	F-ratio	p-value
Total corrected	2.2	32			
Model	1.97	7	0.28	31	<0.01
Residual	0.23	25	9.1 10 <sup>-3</sup>		
Lack of fit	0.16	19	8.6 10 <sup>-3</sup>	0.79	0.68
Pure error	0.07	6	0.01		

**Table 2.** Analysis of variance for the regression model on the effects of carbonization conditions on determined heterogeneities. The heterogeneity values were log10 transformed.



**Figure 3.** Examples of the predicted carbon contents (% daf) of different samples. The histograms illustrate pixel count for different carbon contents. Sample labels show carbonization conditions given in Table 1. Determined heterogeneities,  $h_i \cdot 10^4$ : (a) 3.8, (b) 5.8, (c) 13, (d) 3.4, (e) 2.9, (f) 4.2, (g) 1.2, (h) 5.3 and (i) 14.

sampling to account for weight differences of different sub-samples<sup>24</sup>. However, we assumed constant mass and penetration depth across all pixels and expressed  $h_i$  as:

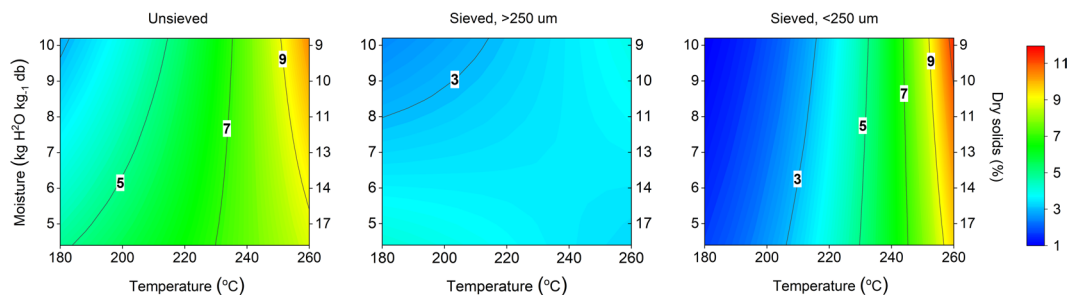
$$h_i = N \sum_{i=1}^N \frac{(c_i - \bar{c})^2}{\bar{c}^2} \quad (5)$$

where  $c_i$  denoted the carbon concentration in a single pixel (% daf),  $\bar{c}$  the average concentration of all pixels in a sample, i.e., predicted sample concentration (% daf) and  $N$  the number of pixels within a sample. This definition provided a dimensionless heterogeneity metric corrected for the number of pixels between different samples that can be used for comparing different sample properties. In addition, the metric took into account increases in spatial detail as heterogeneity generally increases with an increasing number of image pixels.

As illustrated in Fig. 3, the carbon contents and respective heterogeneity increased gradually with increasing carbonization temperature. Higher temperatures hence generated more variation in the carbon content of the samples. The distributions also seemed to be affected by the different size fractions. It is currently well known that temperature governs hydrothermal reactions and higher temperatures lead to increased decomposition of the lignocellulosic components normally present in biomass<sup>25,26</sup>. The raw calibration spectra indicated the loss of O-H, C-O and C-H bonds suggesting the decomposition of carbohydrates. The histograms in Fig. 3 give an indication of the progress of hydrothermal reactions. Higher temperatures suggested increased reaction rates especially for smaller particles. We have previously shown that the variation in pixel predictions decreased with the severity of carbonization after the char samples from different waste feeds were milled to obtain undistorted spectra<sup>27</sup>. However, a robust heterogeneity metric should be able to determine changes in heterogeneity without specific sample pretreatment.

As the original samples were prepared based on an experimental design, the determined heterogeneities could also be described based on a regression model. This allows evaluating the effects of carbonization conditions on the heterogeneity metric and to illustrate changes through response surfaces. An interaction model based on log10 transformed response values explained 90% of data variation with no significant lack of fit. Different sample size fractions were described by two dummy variables. The model showed that both temperature and sample size fraction had a statistically significant effect ( $p < 0.001$ ) on heterogeneity and also showed a significant interaction ( $p < 0.001$ ). Moisture content was the least significant variable ( $p = 0.10$ ) but showed a significant interaction with temperature ( $p = 0.01$ ). The standard error of the model was 0.10 within a range of 1.16 log10 transformed





**Figure 4.** Response surfaces of predicted heterogeneity for different sample fractions based on carbonization conditions.

heterogeneities (Table 2). As illustrated in Fig. 4, the model predictions confirmed that heterogeneity increased with carbonization temperature. The sieving procedure also significantly decreased heterogeneity in the larger particle size fraction. Particles smaller than 250  $\mu\text{m}$  included a wider range of carbon contents that were nearly comparable to the original unsieved samples. As sample heterogeneity can be determined and modelled, it can also be controlled. This opens up doors for enhancing the quality of renewable carbon materials.

To the better of our knowledge, this is the first time that hyperspectral NIR imaging has been used to quantify the constitutional heterogeneity of a predicted analyte distribution. The concept of heterogeneity based on the theory of sampling has previously been discussed by Piqueras *et al.*<sup>28</sup> in an imaging context. The authors used the standard deviations of predicted constituent distributions determined through multivariate curve resolution as constitutional heterogeneity estimates. In our work, their method did not provide robust heterogeneity estimates as the metric is highly sensitive to outlier removal. However, the authors made important observations on the differences between constitutional and distributional heterogeneity when applied to images. As constitutional heterogeneity describes the variation of independent pixel values around the mean, the spatial structure of the image is lost<sup>28</sup>. Thus heterogeneity is best interpreted by also considering the predicted concentration image (Fig. 3), which together provide information on both the constitutional and distributional heterogeneities. For renewable carbon materials, determination of heterogeneity enables developing methods to minimize variation and increase the quality of the carbon material for subsequent applications. This will help to develop renewable carbon materials for replacing respective fossil alternatives in a wide range of energy and environmental applications. This method can also be used for characterizing other materials that absorb in the NIR region.

## Conclusions

Heterogeneity affects material quality. Reliable quantification of heterogeneity has traditionally been a challenging task, as it requires rigorous sampling and analysis procedures. Spectral imaging can determine the spatial distribution of an analyte in sample, thus transforming each pixel into a sampling cell. With a large number of image pixels, the results can be evaluated using population statistics, which allows determining a robust heterogeneity metric for biological materials. Here we have shown that hyperspectral NIR imaging can be used for reliable determination of heterogeneity of biomass-derived carbon after multivariate image regression. Heterogeneity can also be modelled and controlled based on carbonization conditions. In the future, the variation and quality of carbon materials can potentially be monitored online within production systems. The concept of heterogeneity should however be extended beyond the imaged surface a material. Future work should also determine the effect of pixel size on reliable heterogeneity estimates and separate physical and chemical heterogeneities.

## References

- Gy, P. M. The analytical and economic importance of correctness in sampling. *Anal. Chim. Acta* **190**, 13–23, [https://doi.org/10.1016/S0003-2670\(00\)82864-0](https://doi.org/10.1016/S0003-2670(00)82864-0) (1986).
- Gy, P. M. Introduction to the theory of sampling I. Heterogeneity of a population of uncorrelated units. *Trends Anal. Chem.* **14**, 67–76, [https://doi.org/10.1016/0165-9936\(95\)91474-7](https://doi.org/10.1016/0165-9936(95)91474-7) (1995).
- Manley, M. Near-infrared and hyperspectral imaging: non-destructive analysis of biological materials. *Chem. Soc. Rev.* **43**, 8200–8214, <https://doi.org/10.1039/c4cs00062e> (2014).
- Pelletier, M. J. & Pelletier, C. C. in *Raman, infrared and near-infrared chemical imaging* (eds S Šašić & Y Ozaki) (John Wiley & Sons, Inc., 2010).
- Boldrini, B., Kessler, W., Rebner, K. & Kessler, R. W. Hyperspectral imaging: review of best practice, performance and pitfalls for in-line and on-line applications. *J. Near Infrared Spectrosc.* **20**, 483–508, <https://doi.org/10.1255/1003> (2012).
- Amigo, J. M., Babamoradi, H. & Elcoroaristizabal, S. Hyperspectral image analysis. A tutorial. *Anal. Chim. Acta* **896**, 34–51, <https://doi.org/10.1016/j.aca.2015.09.030> (2015).
- Burger, J. & Geladi, P. Hyperspectral NIR image regression part II: dataset preprocessing diagnostics. *J. Chemom.* **20**, 106–119, <https://doi.org/10.1002/cem.986> (2006).
- Badgley, G., Field, C. B. & Berry, J. A. Canopy near-infrared reflectance and terrestrial photosynthesis. *Sci. Adv.* **3**, e1602244, <https://doi.org/10.1126/sciadv.1602244> (2017).
- Mirschel, G., Daikos, O., Scherzer, T. & Steckert, C. Near-infrared chemical imaging used for in-line analysis of inside adhesive layers in textile laminates. *Anal. Chim. Acta* **932**, 69–79, <https://doi.org/10.1016/j.aca.2016.05.015> (2016).
- Pan, J. *et al.* Visible/near-infrared subdiffraction imaging reveals the stochastic nature of DNA walkers. *Sci. Adv.* **3**, e1601600, <https://doi.org/10.1126/sciadv.1601600> (2017).
- Pebriere, H., Ghyselink, C., Lempereur, L. & Brenier, C. Investigation of the composition of anabolic tables using near infrared spectroscopy and Raman chemical imaging. *Drug Test. Anal.* **8**, 370–377, <https://doi.org/10.1002/dta.1843> (2015).

12. Wilczyński, S., Koprowski, R., Marmion, M., Duda, P. & Błońska-Fajfrowska, B. The use of hyperspectral imaging in the VNIR (400–1000 nm) and SWIR range (1000–2500 nm) for detecting counterfeit drugs with identical API composition. *Talanta* **160**, 1–8, <https://doi.org/10.1016/j.talanta.2016.06.057> (2016).
13. de Moura França, L., Amigo, J. M., Cairós, C., Bautista, M. & Pimentel, M. F. Evaluation and assessment of homogeneity in images. Part 1: Unique homogeneity percentage for binary images. *Chemom. Intell. Lab. Syst.* **171**, 26–39, <https://doi.org/10.1016/j.chemolab.2017.10.002> (2017).
14. Titirici, M.-M. *et al.* Sustainable carbon materials. *Chem. Soc. Rev.* **44**, 250–290, <https://doi.org/10.1039/c4cs00232f> (2015).
15. Deng, J., Li, M. & Wang, Y. Biomass-derived carbon: synthesis and applications in energy storage and conversion. *Green Chem.* **18**, 4824–4854, <https://doi.org/10.1039/c6gc01172a> (2016).
16. Jain, A., Balasubramanian, R. & Srinivasan, M. P. Hydrothermal conversion of biomass waste to activated carbon with high porosity: a review. *Chem. Eng. J.* **283**, 789–805, <https://doi.org/10.1016/j.cej.2015.08.014> (2016).
17. Wold, S., Esbensen, K. H. & Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **2**, 37–52, [https://doi.org/10.1016/0169-7439\(87\)80084-9](https://doi.org/10.1016/0169-7439(87)80084-9) (1987).
18. Geladi, P. & Kowalski, B. R. Partial least squares regression: a tutorial. *Anal. Chim. Acta* **185**, 1–17, [https://doi.org/10.1016/0003-2670\(86\)80028-9](https://doi.org/10.1016/0003-2670(86)80028-9) (1986).
19. Osborne, B. G., Fearn, T. & Hindle, P. H. *Practical NIR spectroscopy with applications in food and beverage analysis*. 2nd edn, 227 (Pearson Education Limited, 1993).
20. Schwanninger, M., Rodrigues, J. C. & Fackler, K. A review of band assignments in near infrared spectra of wood and wood components. *J. Near Infrared Spectrosc.* **19**, 287–308, <https://doi.org/10.1255/jnirs.955> (2011).
21. Engelsen, S. B. Near infrared spectroscopy - a unique window of opportunities. *NIR news* **27**, 14–17, <https://doi.org/10.1255/nirn.1620> (2016).
22. Tsu, D. V., Muehle, M., Becker, M., Schuelke, T. & Slagter, J. Quantification of diffuse scattering in glass and polymers by parametric power law analysis of UV to NIR light. *Surf. Coat. Technol.* **336**, 39–53, <https://doi.org/10.1016/j.surfcoat.2017.08.054> (2018).
23. Mäkelä, M., Volpe, M., Volpe, R., Fiori, L. & Dahl, O. Spatially resolved spectral determination of polysaccharides in hydrothermally carbonized biomass. *Green Chemistry* **20**, 1114–1120, <https://doi.org/10.1039/C7GC03676K> (2018).
24. Petersen, L. & Esbensen, K. H. Representative process sampling for reliable data analysis - a tutorial. *J. Chemom.* **19**, 625–647, <https://doi.org/10.1002/cem.968> (2005).
25. Falco, C., Baccile, N. & Titirici, M.-M. Morphological and structural differences between glucose, cellulose and lignocellulosic biomass derived hydrothermal carbons. *Green Chem.* **13**, 3273–3281, <https://doi.org/10.1039/c1gc15742f> (2011).
26. Reza, M. T., Becker, W., Sachsenheimer, K. & Mumme, J. Hydrothermal carbonization (HTC): Near infrared spectroscopy and partial least-squares regression for determination of selective components in HTC solid and liquid products derived from maize silage. *Bioresour. Technol.* **161**, 91–101, <https://doi.org/10.1016/j.biortech.2014.03.008> (2014).
27. Mäkelä, M. & Geladi, P. Hyperspectral imaging to determine the properties and homogeneity of renewable carbon materials. *ChemSusChem* **10**, <https://doi.org/10.1002/cssc.201700777> (2017).
28. Piqueras, S., Burger, J., Tauler, R. & de Juan, A. Relevant aspects of quantification and sample heterogeneity in hyperspectral image resolution. *Chemom. Intell. Lab. Syst.* **117**, 169–182, <https://doi.org/10.1016/j.chemolab.2011.12.004> (2012).

## Acknowledgements

The authors wish to thank Julia Forsberg from RISE Processum AB for her contribution in the preparation of hydrothermal carbon. The funding provided by the Swedish Energy Agency through project 42372-1 and Bio4Energy, a strategic research environment appointed by the Swedish government, are gratefully acknowledged.

## Author Contributions

Both M.M. and P.G. contributed to the design of the study. M.M. performed the sample preparation, image acquisition, and data analysis and wrote the initial manuscript text. Both M.M. and P.G. finalized and reviewed the manuscript.

## Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018