



Estimation of structural disorder degree from carbon materials' Raman spectra by Multivariate Curve Resolution and Partial Least Squares Regression

S. Cerruti,^a E. Robotti,^a L. Cagna,^a G. Gatti,^b A. Croce,^a V. Polliotto,^c E. Marengo^a

^a Dipartimento di scienze e innovazione tecnologica, Università degli Studi del Piemonte Orientale, Viale Teresa Michel 11, 15121 Alessandria (Italy)

^b Dipartimento per lo sviluppo sostenibile e la transizione ecologica, Università degli Studi del Piemonte Orientale, Complesso San Giuseppe, Piazza Sant'Eusebio 5, 13100 Vercelli (Italy)

^c ITT Italia S.R.L., Motion Technologies Innovation Center, Via Molini 19, 12032 Barge (Italy)
simone.cerruti@uniupo.it

Raman spectroscopy is the most frequently used technique to evaluate the disorder degree inside the chemical structure of carbon materials: whereas a perfect graphite sample shows only a narrow band at ca. 1580 cm⁻¹ (*G* band), other components (*D* bands) become Raman-active in the presence of structural defects¹, the most prominent being at ca. 1350 cm⁻¹ (*D*₁), thus broadening the global profile of the spectrum between 1000 and 1800 cm⁻¹. Usually, the spectra are treated via a curve-fitting process, using a pre-established set of mathematical functions (at least 5, with different shapes²⁻³), and then some specific disorder indices are derived from their properties (height, width, position, etc.). However, there is no general agreement on the number or shape(s) of those components; moreover, the fitting is usually done by hand or by non-automatic methods, thus resulting in a tedious and time-consuming process and leading to a high uncertainty of the results⁴.

To overcome some of these problems, two different chemometric procedures based respectively on MCR and PLSR techniques, were applied in order to evaluate the disorder degree directly from the Raman spectra of many different carbon materials: petroleum coke, partially graphitized coke, synthetic and natural graphite (ca. 350 measures in total). The indices were calculated via a semi-automatic software⁵ and consist of the ratio between the intensities of *D* and *G* (*R*_{*D*}), the sum of the areas of all the components divided by the intensity of *G* (*Standard Total Area, STA*), and the ratio between the average intensity value in 1575 – 1595 cm⁻¹ and the average intensity value in 1610 – 1630 cm⁻¹ (*G-shape-factor, GSF*)⁵. Monte Carlo cross-validation was used to estimate the prediction ability and in the case of PLSR to establish the optimal number of Latent Variables to retain, whereas for MCR a set of 3 pure components represent the best compromise between description completeness and model complexity: in any case, the models show good performances both in description and prediction, always leading to a coefficient of determination *R*² > 0.99 in fitting and > 0.98 in cross-validation.

References:

- [1] A.C. Ferrari. *Solid State Communications* 2007, 143, 47–57.
- [2] G. Rantitsch et al. *International Journal of Coal Geology* 2014, 130, 1-7.
- [3] A. Croce et al. *Research on Chemical Intermediates* 2021, 47, 419–431.
- [4] Z.E. Brubaker et al. *Journal of Materials Science* 2021, 56.27, 15087–15121.
- [5] N.K. Lünsdorf et al. *Geostandards and Geoanalytical Research* 2017, 41.4, 593–612.