



Supplement of

Estimating soil carbon sequestration potential with mid-IR spectroscopy and explainable machine learning

Yang Hu and Raphael A. Viscarra Rossel

Correspondence to: Yang Hu (yang.hu4@postgrad.curtin.edu.au) and Raphael A. Viscarra Rossel (r.viscarra-rossel@curtin.edu.au)

The copyright of individual parts of the supplement might differ from the article licence.

MAOC model interpretation with SHAP

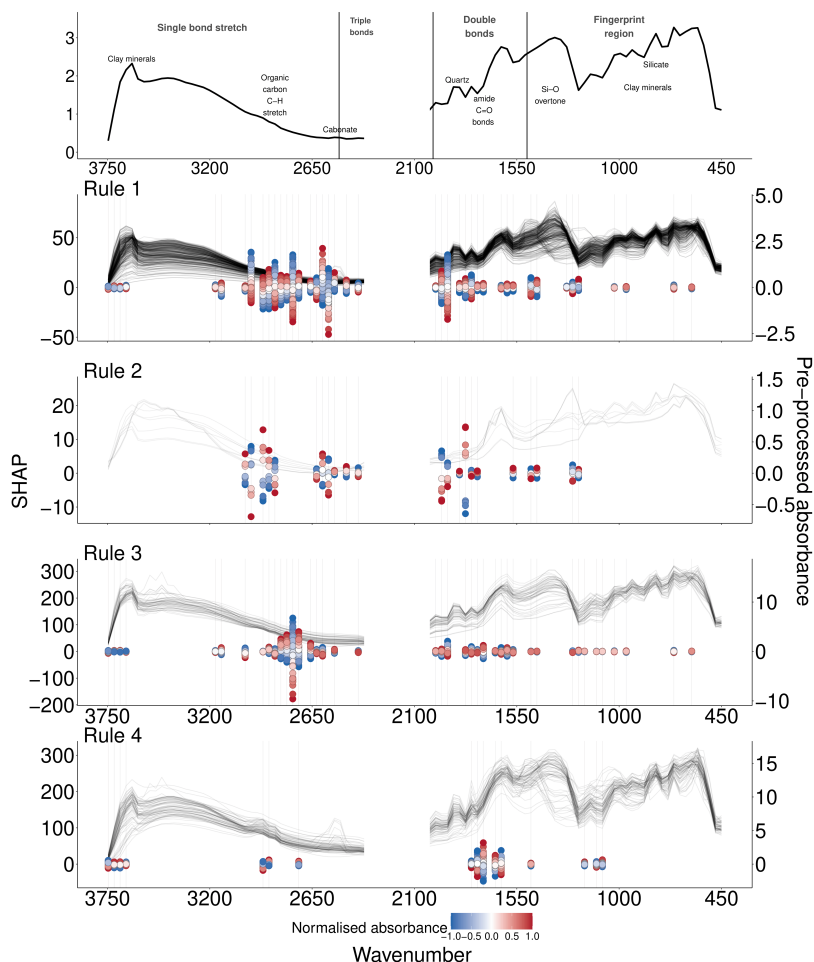


Figure S1. The mean spectra, key spectral assignment, and the SHAP contribution of the spectral regions used in each linear model of each rule of MAOC _{CUBIST} model. A positive SHAP value indicates a positive contribution to a model with increased absorbance, while a negative SHAP value indicates a negative contribution with increased absorbance. The magnitude of SHAP indicates the strength of the contribution. The SHAP values are plotted over the pre-processed spectra of each rule set. The SHAP values are coloured by the normalised absorbance value at each wavenumber, ranging from -1 (lowest absorbance at each wavenumber) to 1 (highest absorbance at each wavenumber). SHAP values of each rule are plotted in different y-axes to accommodate differences in magnitude across rules.

5 The organic C-H region ($2946\text{--}2850\text{ cm}^{-1}$) shows the strongest and most consistent SHAP contributions across all four rules, with alternating positive and negative values at peaks and troughs (Figure S1). Especially in Rule 3, where high absorbance in the two peaks associated with organic C-H region have positive contribution with negative contribution from the trough in between (Figure S1). The fingerprint region has comparatively weaker contributions across all rules (Figure S1). Rule 4 is distinct in showing stronger SHAP contributions in the double bonds region and fingerprint region relative to the other rules (Figure S1).