



**Figure S1: Sensitivity of the  $\text{PARTY}_{\text{SOCv2.0EU}}$  machine-learning model to the test set. Model sensitivity to the test set was assessed as its sensitivity to independent test samples (1) from a reference soil group not existing in the training set (*i.e.* excluding Chernozem soil samples from the test set); (2) that are unlikely to be encountered in agricultural soils (*i.e.* excluding from the test set soils sampled at late dates of bare fallow treatments, more than 25 years after the experiment onset, which cannot represent soils with regular carbon input). This figure is built with a test set composed exclusively of soils from independent Cambisols and Luvisols of Northwestern Europe. It should be compared to Fig. 2b (for  $\text{PARTY}_{\text{SOCv2.0EU}}$ ) that is built using the exact same training set and validation procedure (leave-one-site-out), but that uses a different test set including 32 additional independent test samples (15 Chernozem soils samples and 17 soil samples collected at late dates in bare fallow treatments).**