TropSOC Database

2.2.3. Forest – Mineral Soil Layers – Physicochemical soil properties from NIR-MIR spectroscopy

When using these data, please cite the database and the key publication in ESSD:

Doetterl, S.; Bukombe, B.; Cooper, M.; Kidinda, L.; Muhindo, D.; Reichenbach, M.; Stegmann, A.; Summerauer, L.; Wilken, F.; Fiener, P. (2021): TropSOC Database. V. 1.0. GFZ Data Services.
https://doi.org/10.5880/fidgeo.2021.009

Doetterl S., Asifiwe R.K., Baert G., Bamba F., Bauters M., Boeckx P., Bukombe B., Cadisch G., Cizungu L.N., Cooper M., Hoyt A., Kabaseke C., Kalbitz K., Kidinda L., Maier A., Mainka M., Mayrock J., Muhindo D., Mujinya B.B., Mukotanyi, S.M., Nabahungu L., Reichenbach M., Rewald B., Six J., Stegmann A., Summerauer L., Unseld R., Vanlauwe B., Van Oost K., Verheyen K. Vogel C., Wilken F., Fiener P. Organic matter cycling along geochemical, geomorphic and disturbance gradients in forests and cropland of the African Tropics - Project TropSOC Database Version 1.0. *Earth System Science Data*. https://doi.org/10.5194/essd-2021-73, 2021.

Introduction

The dataset comprises a unique plot and sample identifier and 60 additional soil variables derived from physical and chemical laboratory analyses for TropSOC's forest plots. The dataset represents the predictions of depth explicit physicochemical soil properties using near and mid-infrared spectroscopy at the plot level for TropSOC's forest plots. Predictions are given for 10 soil depth increments of 10 cm each for the top 1 meter of soil, calibrated on data from 4 depth intervals (0 - 10 cm, 30 - 40 cm, and 60 - 70 cm, 90 - 100 cm). Missing values are indicated by -9999. Note that some variables report negative values. In most cases, these values are the result of samples with very low concentrations of the measured parameter or samples outside the calibrated range of our measurements. We recommend to consider those data points as "no data" or "below detection limit". Information on wet chemistry calibration data and its methodology can be found in *221_soil_carbon.csv/pdf* and *222_soil_phy_chem.csv/pdf*. Details regarding plots and plot design can be found in *2_forest.pdf*.

Data structure

No.	Variables	Explanation	Unit	Variable names in non-spectrometric analysis[1]
1	plotID	unique identifier of each plot and point where data were collected	-	plotID
2	sampleID	unique identifier of any soil or vegetation sample taken in the field	-	sampleID
3	BD_spec	bulk density of mineral soil layer	g cm^{-3}	BD_m_soil
4	clay_spec	clay [<2 µm] fraction of fine soil [<2 mm]	wt %	clay
5	silt_spec	silt [2-53 µm] fraction of fine soil [<2 mm]	wt %	silt
6	sand_spec	sand [2000-53 µm] fraction of fine soil [<2 mm]	wt %	sand
7	pH_KCL_spec	pH measured in 1 M KCl solution	-	pH_KCl

8	TC_spec	carbon content of the bulk soil	wt %	mean_C_bulk
9	TN_spec	nitrogen content of the bulk soil	wt %	mean_N_bulk
10	CN_spec	C/N ratio	-	-
11	P_avail_spec	soil plant available phosphorus (Bray-P)	mg kg^{-1}	P_avail
12	bases_in_CEC_spec	base saturation of potential cation exchange capacity	%	bases_in_CEC
13	CEC_spec	potential cation exchange capacity	meq 100 g^{-1} [= cmol(+) kg^{-1}]	CEC
14	bases_in_ECEC_spec	base saturation of effective cation exchange capacity	%	bases_in_ECEC
15	ECEC_spec	effective cation exchange capacity	meq 100 g^{-1} [= cmol(+) kg^{-1}]	ECEC
16	exch_bases_Mg_spec	exchangeable Mg^{2+} of ECEC	meq 100 g^{-1} [= cmol(+) kg^{-1}]	exch_bases_Mg
17	exch_bases_Ca_spec	exchangeable Ca^{2+} of ECEC	meq 100 g^{-1} [= cmol(+) kg^{-1}]	exch_bases_Ca
18	exch_bases_K_spec	exchangeable K^{+} of ECEC	meq 100 g^{-1} [= cmol(+) kg^{-1}]	exch_bases_K
19	exch_acidity_Al_spec	exchangeable Al^{3+} of ECEC	meq 100 g^{-1} [= cmol(+) kg^{-1}]	exch_acidity_Al
20	Al_py_extract_spec	mass percent of sodium-pyrophosphate extractable Al in mass percent of the bulk soil	%	mean_Al_py_extract
21	Al_ox_extract_spec	mass percent of ammonium oxalate-oxalic acid extractable Al in mass percent of the bulk soil	%	mean_Al_ox_extract
22	Al_dcb_extract_spec	mean mass percent of dithionite-citrate-bicarbonate extractable Al in mass percent of the bulk soil	%	mean_Al_dcb_extract
23	Fe_py_extract_spec	mass percent of sodium-pyrophosphate extractable Fe in mass percent of the bulk soil	%	mean_Fe_py_extract
24	Fe_ox_extract_spec	mass percent of ammonium oxalate-oxalic acid extractable Fe in mass percent of the bulk soil	%	mean_Fe_ox_extract
25	Fe_dcb_extract_spec	mass percent of dithionite-citrate-bicarbonate extractable Fe in mass percent of the bulk soil	%	mean_Fe_dcb_extract
26	Mn_py_extract_spec	mass percent of sodium-pyrophosphate extractable Mn in mass percent of the bulk soil	%	mean_Mn_py_extract
27	Mn_ox_extract_spec	mass percent of ammonium oxalate-oxalic acid extractable Mn in mass percent of the bulk soil	%	mean_Mn_ox_extract
28	Mn_dcb_extract_spec	mass percent of dithionite-citrate-bicarbonate extractable Mn in mass percent of the bulk soil	%	mean_Mn_dcb_extract
29	Ca_ICPOES_spec	Ca concentration in bulk soil	%	mean_Ca_bulk_ICPOES
30	K_ICPOES_spec	K concentration in bulk soil	%	mean_K_bulk_ICPOES

31	Mg_ICPOES_spec	Mg concentration in bulk soil	%	mean_Mg_bulk_ICPOES
32	Na_ICPOES_spec	Na concentration in bulk soil	%	mean_Na_bulk_ICPOES
33	P_ICPOES_spec	P concentration in bulk soil	%	mean_P_bulk_ICPOES
34	Ti_ICPOES_spec	Ti concentration in bulk soil	%	mean_Ti_bulk_ICPOES
35	AL_ICPOES_spec	Al concentration in bulk soil	%	mean_Al_bulk_ICPOES
36	Fe_ICPOES_spec	Fe concentration in bulk soil	%	mean_Fe_bulk_ICPOES
37	Mn_ICPOES_spec	Mn concentration in bulk soil	%	mean_Mn_bulk_ICPOES
38	Si_XRF_spec	Si concentration in bulk soil	%	Si_XRF
39	S_XRF_spec	S concentration in bulk soil	µg g^{-1}	S_XRF
40	Zr_XRF_spec	Zr concentration in bulk soil	µg g^{-1}	Zr_XRF
41	Ti_XRF_spec	Ti concentration in bulk soil	%	Ti_XRF
42	TI_ZR_ratio_spec	ratio between Ti and Zr in bulk soil	-	-
43	clay_SI_ratio_spec	ratio between clay and Si	-	-
44	Fe_tot_Fe_dcb_ratio_spec	ration between Fe total and Fe dcb	-	-
45	delta_14C_CO2_spec	delta ^{14}C of CO_2	‰	delta_14C_CO2
46	delta_14C_soil_spec	delta ^{14}C of the bulk soil	‰	delta_14C_soil
47	FMC_CO2_spec	fraction modern carbon of CO_2	%	FMC_CO2
48	FMC_soil_spec	fraction modern carbon of the bulk soil	%	FMC_soil
49	Cmass_CPOM_spec	fraction of organic carbon mass associated with coarse particulate organic matter	-	no equivalent (can be calculated from mean_C_bulk multiplied with mean_weight_CPOM)
50	Cmass_m_spec	fraction of organic carbon mass associated with microaggregate	-	no equivalent (can be calculated from mean_C_bulk multiplied with mean_weight_m)
51	Cmass_s+c_spec	fraction of organic carbon mass associated with free silt and clay	-	no equivalent (can be calculated from mean_C_s+c multiplied with mean_weight_s+c)
52	Cconc_bulk_spec	organic carbon concentration of the bulk soil	%TC	mean_C_bulk
53	Cconc_CPOM_spec	organic carbon concentration of coarse particulate organic matter fraction	%TC	mean_C_CPOM
54	Cconc_m_spec	organic carbon concentration of microaggregates fraction	%TC	mean_C_m
55	Cconc_s+c_spec	organic carbon concentration of free silt and clay fraction	%TC	mean_C_s+c
56	Nmass_CPOM_spec	fraction of total nitrogen mass associated with coarse particulate organic matter fraction	-	no equivalent (can be calculated from mean_N_bulk multiplied with mean_weight_CPOM)
57	Nmass_m_spec	fraction of total nitrogen mass associated with microaggregate	-	no equivalent (can be calculated from mean_N_bulk multiplied with mean_weight_m)
58	Nmass_s+c_spec	fraction of total nitrogen mass associated with free silt and clay	-	no equivalent (can be calculated from mean_N_s+c

				multiplied with mean_weight_s+c)
59	Nconc_bulk_spec	total nitrogen concentration of the bulk soil	%TN	mean_N_bulk
60	Nconc_CPOM_spec	total nitrogen concentration of coarse particulate organic matter fraction	%TN	mean_N_CPOM
61	Nconc_m_spec	total nitrogen concentration of microaggregates fraction	%TN	mean_N_m
62	Nconc_s+c_spec	total nitrogen concentration of free silt and clay fraction	%TN	mean_N_s+c

[1)] see 221_soil_carbon.csv/pdf; 222_soil_phy_chem.pdf/csv; 253_c14.csv/pdf

Methods

Preparation

All samples (n=1458) were finely ground using a ball mill and measured with a VERTEX70 Fourier Transform-Infrared (FT-IR) spectrometer with a High Throughput Screening Extension (HTS-XT) (Bruker Optics GmbH, Germany). Two replicates per sample were filled into the cups of a 24-well plate and the surface was flattened without compression using a spatula. IR reflectance was acquired for a wavenumber range of 7500 cm^{-1} and 600 cm^{-1} (wavelengths of 1333.3 nm to 16666.7 nm) with an effective resolution of 2 cm_{-1}. Gold was used as a background material on every well to normalize the reflectance spectra of all subsequent soil samples of the same well. For each sample, 32 co-added internal measurements were averaged and corrected for atmospheric CO_2 and H_2O using the OPUS spectrometer software (Bruker Optics GmbH, Ettingen, Germany). Reflectance was transformed into absorbance using log(1/reflectance) prior to further processing and subsequent modelling.

Analyses

All spectra were processed using the R packages 'simplerspec' (Baumann, 2020), 'prospectr' (Stevens and Ramirez-Lopez, 2020) and 'caret' (Kuhn, 2020) in the R statistical computing environment (R Core Team, 2020). The two replicates of spectral measurements per soil sample were averaged. The spectra were then resampled to a resolution of 16 cm^{-1} and trimmed to the 4000–600 cm^{-1} spectral range. Smaller resolutions (4cm^{-1}, 2 cm^{-1}) were compared as well, but did not result in more accurate calibration models. Spectral pre-processing was optimized using the suggested workflow of Summerauer et al. (2021), minimizing the reconstruction error when back-transforming the PLS projection of the pre-processed spectra of the prediction set into a spectral matrix, using the projection model built with the pre-processed spectra of the reference set. A Savitzki-Golay filter combined with a second derivate, second order polynomial approximation with a window size of 15 cm^{-1} with a subsequent multiplicative scatter correction resulted in the smallest reconstruction error and was therefore selected as pre-processing for further analyses.

For each response variable, the available reference data was merged with the corresponding spectral data. The data were then split up into calibration sets (2/3 of each reference set) and into a validation set (1/3 of each reference set) using k-means sampling, keeping the number of principal components which explain 99 % of the total variance. K-means sampling selects one sample per cluster defined in a cluster analyses on a PCA of pre-processed spectra (Naes et al. 2002). Set sizes varied, since different reference samples were chemically analysed for different soil properties.

For every soil property, Partial Least Squares (PLS) regressions were trained with the calibration sets. To determine the optimal number of components for the regression models, a five times repeated 10-

fold cross-validation was applied. The best number of components was chosen as the one within one standard error of the optimal model with the lowest average Root Mean Square Error (RMSE) for all evaluated components. The optimal number of components was then used to train the final model with the complete calibration set. The maximal number of components was set to 15, the final selected number of components varied between 1 and 14, depending on the response variable. Model validation was independently tested with the held-out validation set. Goodness of fit ranged with an R^2 for the calibration (Table 1) between 0.08 and 0.99 and for the validation (Table 2) between 0.08 and 0.96. The RPIQ (ratio of performance to interquartile distance) ranged for the calibration between 0.28 and 14.00 and for the model validation between 0.34 and 7.68. RPD (Ratio of standard error of performance to standard deviation) ranged from 1.0 to 11.1 for the calibration and from 1.0 to 4.6 for the validation. **Note that no parameters were included, where $R^2 < 0.3$ for validation and calibration data, which generally resembles also low RPD values <1.2 .**

Table 1. Statiscial performance evaluation for the calibration dataset for all measured parameters (n= 52 - 963) described above for TropSOC's soil data and following the workflow suggested by Summerauer et al. (2021). Abbreviations: n= number of observations used; CV = Coefficient of variation; R^2 = explained variation; RMSE = root mean squared error; RPD = Ratio of standard error of performance to standard deviation; RPIQ= ratio of performance to interquartile distance (RPIQ).

Variable no.	Variable	Unit	n	CV	RMSE	R^2	RPD	RPIQ
3	BD_spec	g cm^{-3}	963	25.64	0.27	0.42	1.32	1.62
4	clay_spec	wt %	156	37.41	4.43	0.93	3.78	5.45
5	silt_spec	wt %	156	52.75	4.02	0.80	2.24	1.99
6	sand_spec	wt %	156	43.32	5.28	0.90	3.13	4.58
7	pH_KCL_spec	-	156	17.87	0.27	0.87	2.76	3.51
8	TC_spec	wt %	202	78.75	0.53	0.92	3.49	4.04
9	TN_spec	wt %	202	80.24	0.05	0.89	3.08	3.33
10	CN_spec	-	202	190.20	11.73	0.84	2.50	0.29
11	P_avail_spec	mg kg^{-1}	156	170.09	29.96	0.81	2.28	1.09
12	bases_in_CEC_spec	%	156	102.25	10.61	0.86	2.71	3.77
13	CEC_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	156	51.65	4.62	0.83	2.46	3.74
14	bases_in_ECEC_spec	%	156	68.11	13.45	0.86	2.67	5.69
15	ECEC_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]]	156	63.03	1.92	0.89	3.01	2.99
16	exch_bases_Mg_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	156	113.56	0.80	0.71	1.87	2.33
17	exch_bases_Ca_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	156	116.05	1.44	0.93	3.67	4.81
18	exch_bases_K_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	156	113.86	0.20	0.53	1.46	1.47
19	exch_acidity_Al_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	156	112.40	0.88	0.82	2.39	3.62
20	Al_py_extract_spec	%	52	71.13	0.04	0.90	3.20	3.96
21	Al_ox_extract_spec	%	52	68.35	0.07	0.88	2.94	4.34
22	Al_dcb_extract_spec	%	52	79.98	0.30	0.97	5.53	5.62
23	Fe_py_extract_spec	%	52	158.74	0.17	0.67	1.76	1.20
24	Fe_ox_extract_spec	%	52	81.87	0.35	0.77	2.13	2.95
25	Fe_dcb_extract_spec	%	52	63.57	0.69	0.95	4.38	6.37
26	Mn_py_extract_spec	%	52	198.30	0.02	0.56	1.53	0.55
27	Mn_ox_extract_spec	%	52	113.99	0.06	0.75	2.03	2.42
28	Mn_dcb_extract_spec	%	52	134.49	0.02	0.87	2.84	2.68
29	Ca_ICPOES_spec	mg kg^{-1}	71	155.75	0.10	0.72	1.90	1.56
30	K_ICPOES_spec	mg kg^{-1}	71	82.46	0.04	0.64	1.67	0.93
31	Mg_ICPOES_spec	mg kg^{-1}	71	100.45	0.09	0.73	1.94	1.80
32	Na_ICPOES_spec	mg kg^{-1}	71	88.91	0.01	0.84	2.49	3.34
33	P_ICPOES_spec	mg kg^{-1}	71	91.20	0.06	0.74	1.99	1.71
34	Ti_ICPOES_spec	mg kg^{-1}	71	84.08	0.06	0.92	3.60	5.97
35	AL_ICPOES_spec	g kg^{-1}	71	64.69	0.82	0.93	3.77	6.68
36	Fe_ICPOES_spec	g kg^{-1}	71	46.13	0.45	0.98	7.87	12.91
37	Mn_ICPOES_spec	mg kg^{-1}	71	96.91	0.09	0.73	1.92	2.60
38	Si_XRF_spec	%	58	25.66	0.37	0.99	11.09	13.44
39	S_XRF_spec	µg g^{-1}	58	72.17	198.00	0.77	2.11	1.77
40	Zr_XRF_spec	µg g^{-1}	58	24.71	46.66	0.74	1.98	1.90
41	Ti_XRF_spec	%	58	43.95	0.06	0.98	7.02	12.72

42	TI_ZR_ratio_spec	-	58	48.95	0.00	0.99	8.31	14.01
43	clay_SI_ratio_spec	-	52	51.10	0.56	0.86	2.72	4.24
44	Fe_tot_Fe_dcb_ratio_spec	-	52	47.16	0.38	0.78	2.15	1.07
45	delta_14C_CO2_spec	‰	148	-352.82	44.07	0.68	1.76	1.73
46	delta_14C_soil_spec	‰	160	-119.35	103.54	0.69	1.79	2.02
47	FMC_CO2_spec	%	148	7.94	0.04	0.68	1.76	1.74
48	FMC_soil_spec	%	160	21.99	0.10	0.69	1.79	2.02
49	Cmass_bulk_spec	mg C g soil^{-1}	169	70.88	115.17	0.91	3.39	4.30
50	Cmass_CPOM_spec	mg C g soil^{-1}	166	244.41	59.34	0.33	1.23	0.41
51	Cmass_m_spec	mg C g soil^{-1}	166	90.11	77.90	0.88	2.85	2.68
52	Cmass_s+c_spec	mg C g soil^{-1}	166	69.30	87.65	0.72	1.90	2.22
53	Cconc_bulk_spec	%TC	169	73.34	0.49	0.94	4.05	4.64
54	Cconc_CPOM_spec	%TC	162	153.35	2.87	0.47	1.37	0.87
55	Cconc_m_spec	%TC	166	80.68	0.63	0.91	3.40	3.57
56	Cconc_s+c_spec	%TC	166	66.72	0.47	0.95	4.40	5.52
57	Nmass_bulk_spec	mg N g soil^{-1}	169	74.55	10.72	0.90	3.11	3.65
58	Nmass_CPOM_spec	mg N g soil^{-1}	166	270.86	3.05	0.29	1.19	0.32
59	Nmass_m_spec	mg N g soil^{-1}	166	91.48	7.26	0.85	2.61	2.66
60	Nmass_s+c_spec	mg N g soil^{-1}	166	70.43	7.92	0.69	1.80	2.06
61	Nconc_bulk_spec	%TN	169	77.13	0.05	0.90	3.23	3.34
62	Nconc_CPOM_spec	%TN	162	175.39	0.12	0.48	1.39	0.65
63	Nconc_m_spec	%TN	166	84.11	0.06	0.88	2.90	2.83
64	Nconc_s+c_spec	%TN	166	67.39	0.04	0.94	4.16	4.96

*) Note: The Unit [me 100 g^{-1}] is equal to [cmol(+) kg^{-1}]

Table 2. Statiscial performance evaluation for the validation dataset for all measured parameters (n= 26 - 482) described above for TropSOC's soil data and following the workflow suggested by Summerauer et al. (2021). Abbreviations: n= number of observations used; CV = Coefficient of variation; R^2 = explained variation; RMSE = root mean squared error; RPD = Ratio of standard error of performance to standard deviation; RPIQ= ratio of performance to interquartile distance (RPIQ).

Variable no.	Variable	Unit	n	CV	RMSE	R^2	RPD	RPIQ
3	BD_spec	g cm^{-3}	482	27.56	0.30	0.43	1.32	1.63
4	clay_spec	wt %	78	36.04	5.40	0.89	3.02	4.88
5	silt_spec	wt %	78	45.49	4.74	0.61	1.55	1.80
6	sand_spec	wt %	78	38.12	6.55	0.81	2.25	3.89
7	pH_KCL_spec	-	78	14.70	0.25	0.86	2.62	3.91
8	TC_spec	wt %	101	70.94	0.53	0.90	3.20	4.00
9	TN_spec	wt %	101	70.84	0.05	0.88	2.86	3.57
10	CN_spec	-	101	127.93	9.30	0.87	2.05	0.34
11	P_avail_spec	mg kg^{-1}	79	165.86	31.97	0.80	2.25	1.14
12	bases_in_CEC_spec	%	79	87.13	11.11	0.79	2.13	3.53
13	CEC_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	78	51.02	5.00	0.86	2.50	3.38
14	bases_in_ECEC_spec	%	79	57.82	14.87	0.80	2.24	4.46
15	ECEC_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]]	79	59.17	2.21	0.84	2.31	2.69
16	exch_bases_Mg_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	79	101.92	0.80	0.70	1.82	2.27
17	exch_bases_Ca_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	79	102.30	1.58	0.90	2.86	3.63
18	exch_bases_K_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	79	119.58	0.18	0.58	1.52	1.11
19	exch_acidity_Al_spec	meq 100 g^{-1} [= cmol(+) kg^{-1}]	79	121.69	0.90	0.75	1.96	3.10
20	Al_py_extract_spec	%	27	81.54	0.06	0.81	2.30	3.26
21	Al_ox_extract_spec	%	27	69.80	0.11	0.72	1.87	2.20
22	Al_dcb_extract_spec	%	27	57.03	0.70	0.72	1.61	1.99
23	Fe_py_extract_spec	%	27	117.14	0.16	0.59	1.43	1.60
24	Fe_ox_extract_spec	%	27	89.71	0.53	0.61	1.60	1.56
25	Fe_dcb_extract_spec	%	27	58.76	0.60	0.96	4.63	7.68
26	Mn_py_extract_spec	%	27	177.87	0.03	0.36	1.26	0.47
27	Mn_ox_extract_spec	%	27	107.84	0.08	0.59	1.57	1.76
28	Mn_dcb_extract_spec	%	27	106.21	0.02	0.79	2.12	1.94
29	Ca_ICPOES_spec	mg kg^{-1}	36	147.03	0.12	0.65	1.69	1.64
30	K_ICPOES_spec	mg kg^{-1}	36	88.39	0.05	0.69	1.70	0.91
31	Mg_ICPOES_spec	mg kg^{-1}	36	97.11	0.13	0.72	1.80	1.68
32	Na_ICPOES_spec	mg kg^{-1}	36	79.53	0.01	0.84	2.50	3.20
33	P_ICPOES_spec	mg kg^{-1}	36	72.36	0.06	0.76	1.75	2.41
34	Ti_ICPOES_spec	mg kg^{-1}	36	81.34	0.07	0.92	3.48	5.12
35	AL_ICPOES_spec	g kg^{-1}	36	59.89	1.17	0.89	2.75	4.93
36	Fe_ICPOES_spec	g kg^{-1}	36	40.27	0.84	0.94	3.87	6.72
37	Mn_ICPOES_spec	mg kg^{-1}	36	75.20	0.09	0.59	1.16	1.98
38	Si_XRF_spec	%	29	22.52	0.93	0.93	3.72	4.20
39	S_XRF_spec	µg g^{-1}	29	100.93	383.82	0.69	1.56	0.55
40	Zr_XRF_spec	µg g^{-1}	29	26.47	77.53	0.41	1.33	0.99
41	Ti_XRF_spec	%	29	39.24	0.11	0.92	3.36	5.41

42	TI_ZR_ratio_spec	-	29	50.52	0.00	0.92	3.49	6.08
43	clay_SI_ratio_spec	-	26	48.77	0.79	0.77	2.08	3.94
44	Fe_tot_Fe_dcb_ratio_spec	-	27	44.22	0.36	0.81	2.16	1.23
45	delta_14C_CO2_spec	‰	74	-414.37	40.41	0.34	1.15	1.49
46	delta_14C_soil_spec	‰	80	-111.32	124.71	0.50	1.40	1.80
47	FMC_CO2_spec	%	74	4.71	0.04	0.34	1.15	1.49
48	FMC_soil_spec	%	80	20.63	0.13	0.50	1.40	1.81
49	Cmass_bulk_spec	mg C g soil^{-1}	85	63.80	125.86	0.85	2.57	3.14
50	Cmass_CPOM_spec	mg C g soil^{-1}	84	189.40	44.84	0.26	1.13	0.58
51	Cmass_m_spec	mg C g soil^{-1}	84	80.71	92.06	0.76	2.06	2.13
52	Cmass_s+c_spec	mg C g soil^{-1}	84	59.65	78.17	0.69	1.81	2.44
53	Cconc_bulk_spec	%TC	85	64.68	0.57	0.86	2.66	3.31
54	Cconc_CPOM_spec	%TC	82	142.08	3.48	0.35	1.22	0.97
55	Cconc_m_spec	%TC	84	71.67	0.60	0.88	2.92	3.16
56	Cconc_s+c_spec	%TC	84	60.96	0.41	0.95	4.45	6.15
57	Nmass_bulk_spec	mg N g soil^{-1}	85	60.44	11.80	0.80	2.20	3.10
58	Nmass_CPOM_spec	mg N g soil^{-1}	84	181.25	1.77	0.27	1.10	0.57
59	Nmass_m_spec	mg N g soil^{-1}	84	78.18	7.18	0.78	2.13	2.38
60	Nmass_s+c_spec	mg N g soil^{-1}	84	57.58	7.20	0.65	1.66	2.50
61	Nconc_bulk_spec	%TN	85	61.81	0.05	0.82	2.29	2.81
62	Nconc_CPOM_spec	%TN	81	148.58	0.15	0.42	1.27	0.85
63	Nconc_m_spec	%TN	84	69.55	0.06	0.86	2.54	3.18
64	Nconc_s+c_spec	%TN	84	60.87	0.04	0.93	3.59	4.60

*) Note: The Unit [meq 100 g^{-1}] is equal to [cmol(+) kg^{-1}]

Acknowledgment

TropSOC was funded via the Emmy-Noether-Program of the German Research Foundation (project ID 387472333).

References

Baumann, P.: simplerspec: Soil and plant spectroscopic model building and prediction, https://github.com/philipp-baumann/simplerspec, R package version 0.1.0.9001, 2020.

Naes, T., Isaksson, T., Fearn, T., and Davies, T., A user friendly guide to multivariate calibration and classification. NIR Publications, Chichester, United Kingdom.

Kuhn, M. (2020). caret: Classification and Regression Training. R package version 6.0-86. https://CRAN.R-project.org/package=caret, 2002.

Stevens, A. and Ramirez-Lopez, L.: prospectr: Miscellaneous Functions for Processing and Sample Selection of Spectroscopic Data, https://CRAN.R-project.org/package=prospectr, R package version 0.2.0, 2020.

Summerauer, L., Baumann, P., Ramirez-Lopez, L., Barthel, M., Bauters, M., Bukombe, B., Reichenbach, M., Boeckx, P., Kearsley, E., Van Oost, K., Vanlauwe, B., Chiragaga, D., Heri-Kazi, A.B., Moonen, P., Sila, A., Shepherd, K., Mujinya, B.B., Van Ranst, E., Baert, G., Doetterl, S. and Six, J.: A spectral platform for soil properties of central Africa, SOIL DISCUSSIONS (in review)., 2021.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https: //www.R-project.org/, 2020.

Ramirez-Lopez, L.: resemble: Regression and Similarity Evaluation for Memory-Based Learning in Spectral Chemometrics, https://CRAN. R-project.org/package=resemble, R package version 2.1.1, 2020.