

Specific surface area of soils with different clay mineralogy can be estimated from a single hygroscopic water content

Fulai Yan^{a,b,1}, Markus Tuller^{c,2}, Lis W de Jonge^{a,3}, Per Moldrup^{d,4}, Emmanuel Arthur^{a,5,*}

^a Department of Agroecology, Aarhus University, Blichers Alle 20, DK-8830 Tjele, Denmark

^b Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semi-arid Areas of Ministry of Education, Northwest A&F University, Yangling 712100, China

^c Department of Environmental Science, The University of Arizona, 1177 E. 4th Street, Tucson, AZ 85721, USA

^d Department of the Built Environment, Aalborg University, Thomas Manns Vej 23, DK-9220, Aalborg, Denmark

ARTICLE INFO

Handling Editor: Morgan Cristine L.S.

Keywords:

Water sorption isotherms
Cation exchange capacity
Particle-size distribution
Regression models

ABSTRACT

The soil specific surface area (SSA) is an important variable for soil science and geoenvironmental engineering applications, but traditional measurement methods are difficult and time-consuming. Regression models or pedotransfer functions are often used to estimate SSA from other soil properties (e.g., clay content and cation exchange capacity), but these models do not consider the impact of clay mineralogy. Hygroscopic water content (w_h) is intimately linked to these soil properties, which suggests that w_h may be a better parameter for SSA estimation. This study (i) proposes regression models that estimate SSA from w_h at different relative humidity values (5 to 90%) for kaolinite-rich samples (KA), illite-rich or mixed clay samples (IL/MC), montmorillonite-rich samples (ML), and a combination of all samples (ALL) and (ii) compares the performance of the w_h models to other published models that comprise clay, silt and soil organic carbon contents and cation exchange capacity. We found that the sample-specific w_h regression models accurately estimated SSA for KA, IL/MC and ML samples. For KA and IL/MC samples, the performance of the KA model (e.g., for adsorption, average $RMSE = 10.5 \text{ m}^2/\text{g}$) and IL/MC model (average $RMSE = 21.3 \text{ m}^2/\text{g}$) were better than the ALL-calibration model (KA: average $RMSE = 18.7 \text{ m}^2/\text{g}$; ML: average $RMSE = 22.4 \text{ m}^2/\text{g}$). For ML samples, similar model performance between the ML-calibration model (average $RMSE = 41.4 \text{ m}^2/\text{g}$) and the ALL-calibration model (average $RMSE = 41.1 \text{ m}^2/\text{g}$) was observed. In addition, the model performance of regression models based on w_h was superior to models published in the literature that are based on clay, silt and soil organic carbon contents and cation exchange capacity. Overall, this study confirms that a single measure of w_h can provide reliable estimates of the SSA while revealing a significant impact of clay mineralogy on model performance.

1. Introduction

The soil specific surface area (SSA) is a fundamental soil property, which is commonly defined as the surface area per unit mass of soil solids (Pennell, 2002). The SSA plays a crucial role in numerous soil processes including soil aggregation and structure evolution (Utkaeva, 2007), water retention and movement (Arthur et al., 2018), microbial processes, and nutrient dynamics (Pennell, 2002). Thus, accurate measurement of SSA is essential for geoenvironmental engineering (Pennell,

2002; Maček et al., 2013) and agricultural applications (Bashina et al., 2019).

The SSA varies markedly among soils because of the differences in clay content and mineralogy, particle-size distribution, and organic matter content and composition (Carter et al., 1986; Yukselen-Aksoy and Kaya, 2010). Traditional methods for estimating SSA include direct and indirect measurement methods. The direct measurement methods encompass X-ray diffraction (Adamson and Gast, 1990), and electron or atomic force microscopy (Sposito, 1984; Macht et al., 2011)

* Corresponding author.

E-mail address: emmanuel.arthur@agro.au.dk (E. Arthur).

¹ <https://orcid.org/0000-0002-3144-2392>.

² <https://orcid.org/0000-0003-3659-2768>.

³ <https://orcid.org/0000-0003-2874-0644>.

⁴ <https://orcid.org/0000-0003-1619-1457>.

⁵ <https://orcid.org/0000-0002-0788-0712>.

<https://doi.org/10.1016/j.geoderma.2023.116614>

Received 24 March 2023; Received in revised form 2 July 2023; Accepted 13 July 2023

Available online 21 July 2023

0016-7061/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

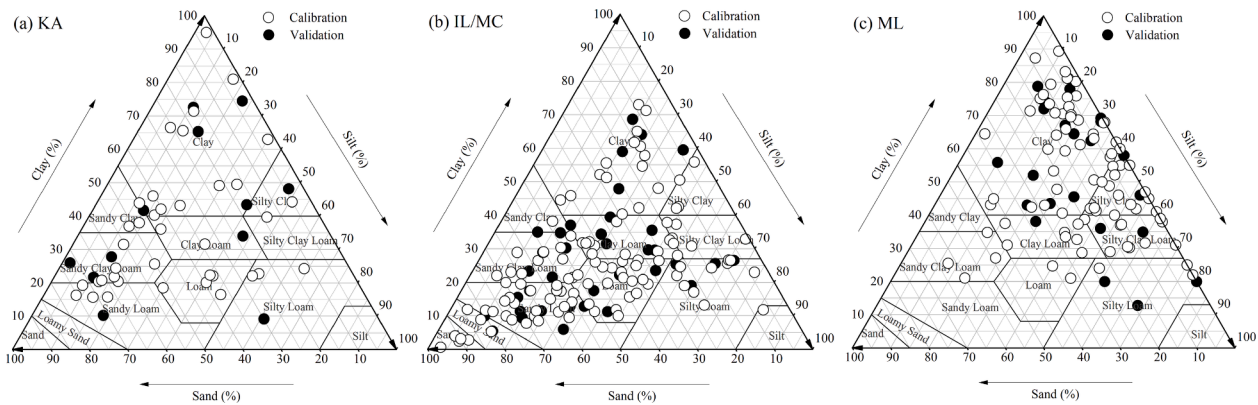


Fig. 1. Distribution of the investigated soil samples displayed on the USDA soil textural triangle. The samples are partitioned into calibration [white symbols] and validation [black symbols] datasets for the different sample groups. KA, kaolinite-rich samples, IL/MC, illite-rich or mixed clay samples; ML, montmorillonite-rich samples.

Table 1

Descriptive statistics for soil texture, organic carbon, cation exchange capacity and soil specific surface area of the investigated soil samples grouped based on dominant clay mineralogy.

Soil property	Units	Min	Max	Mean	CV (%)
Kaolinite-rich samples, $n = 39$ [12]					
Clay	%	15 [8]	86 [73]	36 [39]	52 [57]
Silt		3 [2]	62 [55]	24 [24]	62 [72]
Sand		2 [3]	73 [72]	37 [35]	57 [74]
SOC		0.0 [0.1]	4.9 [4.8]	1.0 [1.4]	111 [109]
CEC	$\text{cmol}_{(+)} \text{kg}^{-1}$	1 [3]	41 [34]	10 [11]	84 [78]
SSA	m^2/g	20 [13]	130 [100]	49 [48]	55 [51]
Illite-rich or mixed clay samples, $n = 114$ [36]					
Clay	%	1 [5]	71 [63]	25 [26]	59 [58]
Silt		3 [9]	78 [64]	30 [30]	55 [50]
Sand		1 [4]	96 [79]	42 [39]	55 [54]
SOC		0.1 [0.3]	4.0 [3.9]	1.6 [1.7]	67 [52]
CEC	$\text{cmol}_{(+)} \text{kg}^{-1}$	2 [6]	75 [51]	20 [20]	61 [56]
SSA	m^2/g	6 [16]	314 [293]	79 [64]	73 [77]
Montmorillonite-rich samples, $n = 84$ [26]					
Clay	%	20 [12]	89 [77]	50 [50]	36 [39]
Silt		2 [9]	72 [74]	33 [33]	48 [56]
Sand		0 [0]	62 [34]	14 [13]	102 [82]
SOC		0.1 [0.1]	6.8 [6.6]	1.3 [1.4]	95 [84]
CEC	$\text{cmol}_{(+)} \text{kg}^{-1}$	16 [11]	87 [83]	49 [47]	31 [34]
SSA	m^2/g	68 [40]	444 [406]	250 [245]	34 [40]
All samples, $n = 237$ [74]					
Clay	%	1 [5]	89 [77]	35 [37]	57 [57]
Silt		2 [2]	78 [74]	30 [30]	54 [56]
Sand		0 [0]	96 [79]	31 [30]	76 [77]
SOC		0.0 [0.1]	6.8 [6.6]	1.4 [1.6]	82 [72]
CEC	$\text{cmol}_{(+)} \text{kg}^{-1}$	1 [3]	87 [83]	29 [28]	71 [69]
SSA	m^2/g	6 [13]	444 [406]	135 [134]	80 [83]

Number of samples (n) are provided for the model calibration [validation] datasets. For each soil property, numbers outside parenthesis are for the calibration dataset and the numbers in parenthesis are for the validation dataset. CV, coefficient of variation; SOC, soil organic carbon; CEC, cation exchange capacity; SSA, soil specific surface area.

to determine the shapes and dimensions of soil particles, and then estimate the SSA. However, these direct measurement methods are rarely used in practical applications because of the large variation in particle dimensions and shapes (Chen et al., 2021). Thus, the SSA is usually estimated indirectly by the adsorption of molecules onto the particle surface. The molecule used may characterise either the external surface (the area of the solid-air interface of the soil), the internal surface (surfaces of clay interlayers or micropores of organic molecules) or the total surface area. Molecules such as N_2 (Heister, 2014) are used for the external SSA, whereas others such as CO_2 (Brunauer et al., 1938), methylene blue (MB) (Hang and Brindley, 1970), and ethylene glycol

monoethyl ether (EGME) (Akin and Likos, 2014) are used to quantify the total SSA. However, these indirect measurement methods also have some practical limitations, such as being laborious and relatively expensive (Rawlins et al., 2010). Therefore, a fast, simple, low-cost and reliable approach to estimating SSA is beneficial.

One approach with great potential is to use hygroscopic water content (w_h), either in the form of water vapor sorption isotherms or a single measurement point. Studies have shown that using water vapor sorption isotherms in combination with isotherm models provides reliable estimates of SSA comparable to that of EGME (Resurreccion et al., 2011; Arthur et al., 2013, Akin and Likos, 2014; Chen et al., 2021; Khorshidi and Lu, 2017). Other studies have also proposed estimating SSA from adsorbed w_h based on an assumed formation of a monolayer of water molecules on the soil sorption surfaces at a relative humidity (RH) of 47% (Newman, 1983) or 21% (Quirk and Murray, 1999). Our previous work showed that the RH at which the monolayer occurs can range from 15 to 34% depending on sorption direction (adsorption or desorption) and the dominant clay mineral in the sample (Arthur et al., 2018). To avoid these challenges, Chen et al. (2021) suggested using the change in w_h between an RH from 45 to 75%, rather than at a single RH point as a better SSA estimator. Recently, Ghanbarian et al. (2021) also combined the particle size distribution and the water retention curve to estimate SSA while considering the particle surface roughness. The above-mentioned methods for determining SSA serve as more viable alternatives, but they require data on either the full or partial water sorption isotherm that takes at least 1–3 days to measure for a sample, or a single w_h value at a fixed RH (e.g., 20%). This limits their application to the availability of sophisticated equipment such as vapor sorption analyzers or requires equilibration of samples to a fixed RH . Moreover, there is limited or no consideration of the impact of clay mineralogy on the above-mentioned methods.

Another alternative to laboratory measurements of SSA is the use of pedotransfer functions (PTFs) that are based on readily available soil variables such as clay content or cation exchange capacity (CEC). For example, Hepper et al. (2006) used data from 24 samples to develop a PTF for SSA based on clay and silt contents. Other studies based on similarly limited soil datasets indicated that SSA can be well estimated by CEC, clay and soil organic carbon (SOC) contents (Farrar and Coleman, 1967; Banin and Amiel, 1970). In addition, Ersahin et al. (2006) and Sepaskhah and Tafteh (2013) developed regression models that used the fractal dimension of the particle size distribution for estimating SSA values. Despite these PTFs being effective for estimating SSA for the dataset they were developed for, there is limited evidence on whether these PTFs can be generalized to soils with different clay mineralogy.

This study aims to extend previous work that proposes to use w_h as an estimator of SSA and to address some of the challenges of existing methods (e.g., limited datasets, non-inclusion of clay mineral data, and

Table 2

Intercept (a) and slope (b) of the SSA regression equations for adsorption soil water content (%) at different relative humidity levels (RH) for all calibration datasets based on clay mineralogy.

RH (%)	Kaolinite-rich samples n = 39			Illite-rich or mixed clay samples n = 114			Montmorillonite-rich samples n = 84			All samples n = 235		
	a	b	R ²	a	b	R ²	a	b	R ²	a	b	R ²
5	38.46	28.02	0.46	14.24	74.35	0.67	79.70	80.88	0.63	11.89	97.50	0.76
10	33.39	27.18	0.54	3.93	68.92	0.75	61.76	69.21	0.71	3.83	82.41	0.84
15	32.20	24.71	0.57	-1.04	61.38	0.79	45.28	60.90	0.75	0.53	69.36	0.87
20	30.26	23.32	0.62	-3.87	54.75	0.82	34.11	53.77	0.79	-0.91	59.39	0.90
25	28.34	22.33	0.65	-5.05	49.19	0.84	27.49	48.08	0.82	-2.01	52.28	0.91
30	26.90	21.45	0.68	-5.53	44.83	0.86	23.28	43.65	0.84	-2.28	46.93	0.93
35	25.28	20.77	0.70	-5.63	41.33	0.87	21.03	40.12	0.85	-2.54	42.93	0.93
40	23.72	20.12	0.73	-5.81	38.46	0.88	19.22	37.30	0.87	-3.00	39.78	0.94
45	22.40	19.45	0.75	-5.86	35.99	0.88	18.44	34.89	0.87	-3.38	37.19	0.94
50	21.04	18.82	0.77	-5.71	33.70	0.89	18.06	32.76	0.88	-3.82	34.94	0.95
55	19.80	18.16	0.79	-5.53	31.58	0.89	17.77	30.84	0.88	-4.29	32.92	0.95
60	18.40	17.47	0.81	-5.25	29.54	0.89	16.99	29.09	0.88	-4.94	31.03	0.95
65	16.71	16.74	0.83	-5.04	27.57	0.89	16.70	27.37	0.87	-5.84	29.24	0.95
70	15.13	15.92	0.85	-5.07	25.64	0.89	16.14	25.67	0.87	-6.91	27.48	0.95
75	13.28	14.87	0.87	-5.02	23.67	0.89	16.31	23.92	0.86	-8.36	25.70	0.94
80	11.57	13.55	0.90	-4.95	21.59	0.88	16.92	22.10	0.85	-10.14	23.88	0.94
85	9.79	11.72	0.91	-4.33	19.24	0.87	18.31	20.08	0.83	-12.50	21.89	0.93
90	9.20	8.90	0.89	-3.12	16.58	0.84	19.84	17.82	0.80	-15.70	19.53	0.90

The *p*-value for all regressions was <0.001.

Table 3

Intercept (a) and slope (b) of the SSA regression equations for desorption soil water content (%) at different relative humidity levels (RH) for all calibration datasets based on clay mineralogy.

RH (%)	Kaolinite-rich samples n = 39			Illite-rich or mixed clay samples n = 114			Montmorillonite-rich samples n = 84			All samples n = 235		
	a	b	R ²	a	b	R ²	a	b	R ²	a	b	R ²
5	37.52	25.26	0.47	11.17	68.85	0.69	65.14	78.62	0.72	7.87	90.63	0.80
10	34.15	22.24	0.54	4.87	55.04	0.77	47.11	59.69	0.79	2.89	66.90	0.87
15	32.12	20.61	0.59	1.20	47.71	0.80	30.68	51.95	0.84	0.16	55.67	0.90
20	30.25	19.48	0.62	-1.38	42.86	0.83	22.94	46.14	0.86	-2.08	48.75	0.92
25	28.64	18.66	0.64	-2.56	39.01	0.85	19.74	41.52	0.87	-3.02	43.69	0.93
30	26.99	18.05	0.67	-3.21	35.77	0.85	19.21	37.44	0.88	-3.32	39.53	0.93
35	25.49	17.40	0.69	-3.58	32.92	0.86	20.42	33.80	0.88	-3.25	35.95	0.94
40	24.10	16.85	0.71	-3.64	30.57	0.87	20.07	31.44	0.88	-3.84	33.49	0.94
45	22.95	16.22	0.73	-3.73	28.57	0.87	19.34	29.58	0.88	-4.61	31.50	0.94
50	21.79	15.56	0.75	-3.95	26.79	0.87	18.25	27.93	0.88	-5.55	29.71	0.94
55	20.59	14.82	0.77	-3.97	25.07	0.87	17.37	26.36	0.87	-6.52	28.03	0.94
60	19.29	14.09	0.79	-4.14	23.48	0.87	16.24	24.92	0.86	-7.70	26.48	0.94
65	17.63	13.39	0.81	-4.37	21.99	0.87	15.30	23.54	0.85	-9.15	25.04	0.93
70	15.51	12.71	0.84	-4.43	20.54	0.87	15.10	22.15	0.85	-10.72	23.65	0.93
75	13.06	11.97	0.87	-4.16	19.11	0.86	15.68	20.77	0.84	-12.36	22.32	0.93
80	10.57	10.86	0.90	-3.64	17.69	0.85	16.88	19.36	0.82	-14.24	20.98	0.92
85	9.02	9.32	0.90	-3.23	16.42	0.84	18.44	17.98	0.81	-16.15	19.63	0.91
90	9.86	7.49	0.86	-3.61	15.47	0.84	18.44	16.89	0.79	-17.99	18.41	0.89

The *p*-value for all regressions was <0.001.

the need for full water sorption isotherm data). The specific objectives were to: (i) develop regression-based models to estimate SSA from w_h , or particle-size distribution, SOC content, and CEC while considering the clay mineralogy of the samples; (ii) validate the models by comparing the estimated SSA to the SSA determined with EGME; and (iii) compare the performance of the developed models to previously published SSA PTFs.

2. Methodology

2.1. Investigated samples

A total of 311 soil samples with different clay mineralogy from different geographical regions were considered for this study (Fig. 1). The samples were grouped into 51 kaolinite-rich samples (KA), 150 illite-rich or mixed clay samples (IL/MC), and 110 montmorillonite-rich samples (ML). The soil samples are a subset of the samples used by Arthur et al. (2023) that had CEC data available in addition to SSA. An

overview of the soil texture, SOC content, CEC, and SSA of the investigated samples, categorised by dominant clay mineralogy, is presented in Table 1.

2.2. Laboratory measurements

2.2.1. Clay mineralogy, and particle size distribution

The soil samples were air-dried and passed through a 2-mm sieve before laboratory analyses. The clay mineralogy was quantified via X-ray Diffraction (XRD). Oriented clay glass slides were prepared for the XRD analyses using the filter peel method (Drever, 1973). To identify the major clay mineral groups based on the measured XRD diffraction patterns, the RockJock software package (Eberl, 2003), which performs a whole-pattern-modified Rietveld-type refinement, was applied. The clay, silt and sand contents of the soil samples were analysed by combining the hydrometer and wet sieving methods after the removal of organic matter (OM) and carbonates where necessary (Gee and Or, 2002).

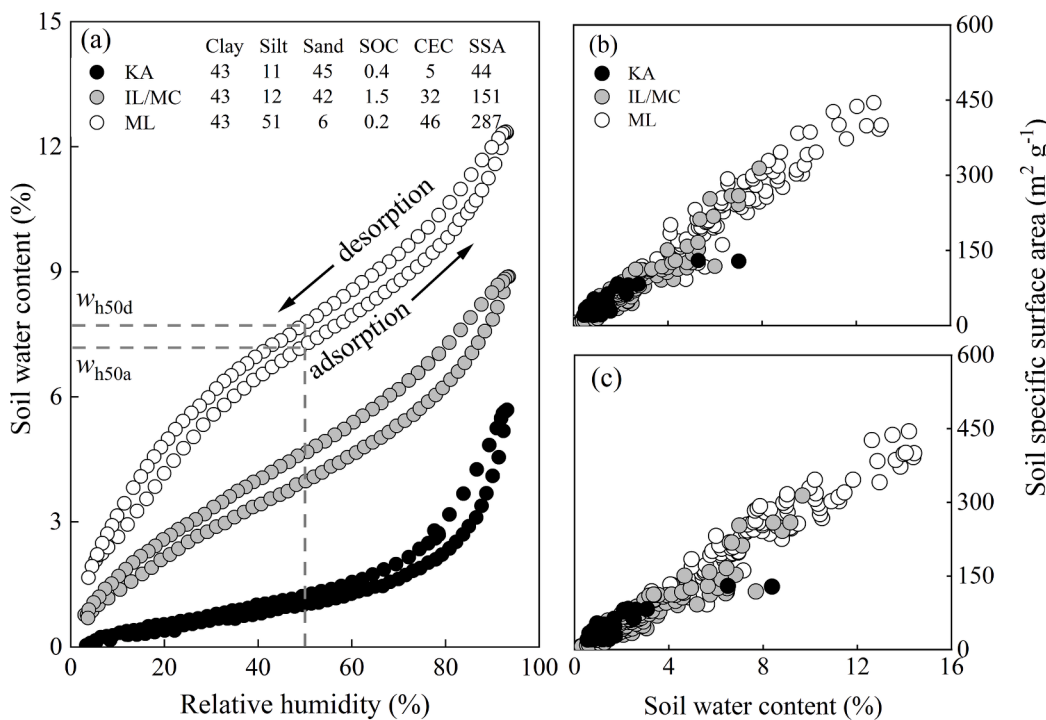


Fig. 2. (a) Examples of soil water vapor sorption isotherms for three soil samples that have identical clay content (43%), but different dominant clay minerals, silt, sand and SOC contents (%), CEC ($cmol_{(+) } kg^{-1}$) and SSA (m^2/g). The figure depicts hysteresis, and derivation of the w_h at relative humidity (RH) of 50% (w_{h50a} and w_{h50d}), (b) the relationship between w_{h50d} and SSA, and (c) w_{h50d} and SSA for the three sample groups. SOC, soil organic carbon; CEC, cation exchange capacity; SSA, soil specific surface area. KA, kaolinite-rich samples, IL/MC, illite-rich or mixed clay samples; ML, montmorillonite-rich samples.

2.2.2. Hygroscopic water content (w_h)

Soil water vapor sorption isotherms were measured with an automated vapor sorption analyzer (VSA) (METER Group Inc., Pullman, WA, USA) for soil samples that were air-dried to a stable relative humidity (RH) of $\sim 45\%$. The isotherms covered an RH range from 3% to 93% with a measurement resolution of 2% RH at a stable temperature of 25°C. After the measurements, the samples were oven-dried for 48 h to determine the water content at the respective RH values. Details about the measurement procedure are provided in Arthur et al. (2014).

For modeling, the w_h at 18 selected RH values (i.e., 5 to 90% with increments of 5%) were obtained directly from the adsorption and desorption data by linear interpolation between the two closest points when the w_h at the exact RH was not directly available from the data.

2.2.3. Soil organic carbon content, cation exchange capacity, and specific surface area

To determine the total carbon (TC) content, subsamples of the previously sieved samples were ball-milled before oxidation of the carbon at 950 °C using an elemental analyzer combined with a thermal conductivity detector (Thermo Fisher Scientific, Waltham, MA, USA). For samples containing calcium carbonate, SOC was calculated as the difference between TC and inorganic carbon estimated from the calcium carbonate percentage.

The CEC was measured with the ammonium acetate extraction method at pH 7.0 and pH 8.2 for 25 of the samples with high amounts of $CaCO_3$ (Sumner and Miller, 1996). The SSA was estimated with the retention of EGME at monolayer coverage (Pennell, 2002) without pretreatments (removal of OC or ion saturation). Briefly, the 2-mm sieved sample was mixed with EGME, and the mixture was sealed in an equilibration chamber under a high vacuum. After an equilibration period between 12 and 72 h, depending on sample SSA, the samples were repeatedly weighed in sealed weighing bottles until the mass remained constant. The average SSA was calculated from the mass of EGME retained by three replicate samples for each soil.

2.3. Modeling

2.3.1. Modeling rationale, and data exploration

The use of w_h as a predictor of SSA is possible because the amount of hygroscopic water sorbed onto the external and inter-crystalline surfaces of clay minerals is strongly dependent on the external and internal surfaces available for sorption, as well as the interlayer cations (Lu and Zhang, 2020). In addition to the type of clay mineral, increasing clay and organic matter contents are often correlated to increasing SSA, and an associated increase in w_h . Consequently, clay and organic matter contents, and CEC are often used as variables in PTFs to estimate the SSA values (e.g., Farrar and Coleman, 1967; Banin and Amiel, 1970; de Jong, 1999).

2.3.2. Data partitioning for modeling

For modelling purposes, the three sample groups (KA, IL/MC, and ML) were each partitioned into model calibration (75%) and validation (25%) subsets. The partitioning was done in R v4.2.2 (R Core Team, 2022) using the “createDataPartition” data splitting function in the “caret” package (Kuhn, 2019). The function creates balanced splits of the data while preserving their overall class distribution. The divisions were done based on the SSA data. Consequently, the model calibration datasets for the KA, IL/MC, and ML contained 39, 114, and 84 samples, respectively. The model validation datasets for KA, IL/MC, and ML contained 12, 36, and 26 samples, respectively (Table 1). To verify the impact of clay mineralogy on model calibration and validation, a fourth sample group denoted ALL was created by combining the datasets of the three sample groups. The ALL group calibration dataset consisted of 237 samples and a validation dataset of 74 samples.

2.3.3. Model development

Before the model development, we checked the suitability of the data (clay, silt and SOC contents, w_h , CEC and SSA) for the regression analyses. The normality was determined by the Shapiro-Wilk test and the homogeneity of variance by the Brown-Forsythe test. It was found that the data satisfied both criteria.

2.3.3.1. Hygroscopic water content (w_h). For the four calibration

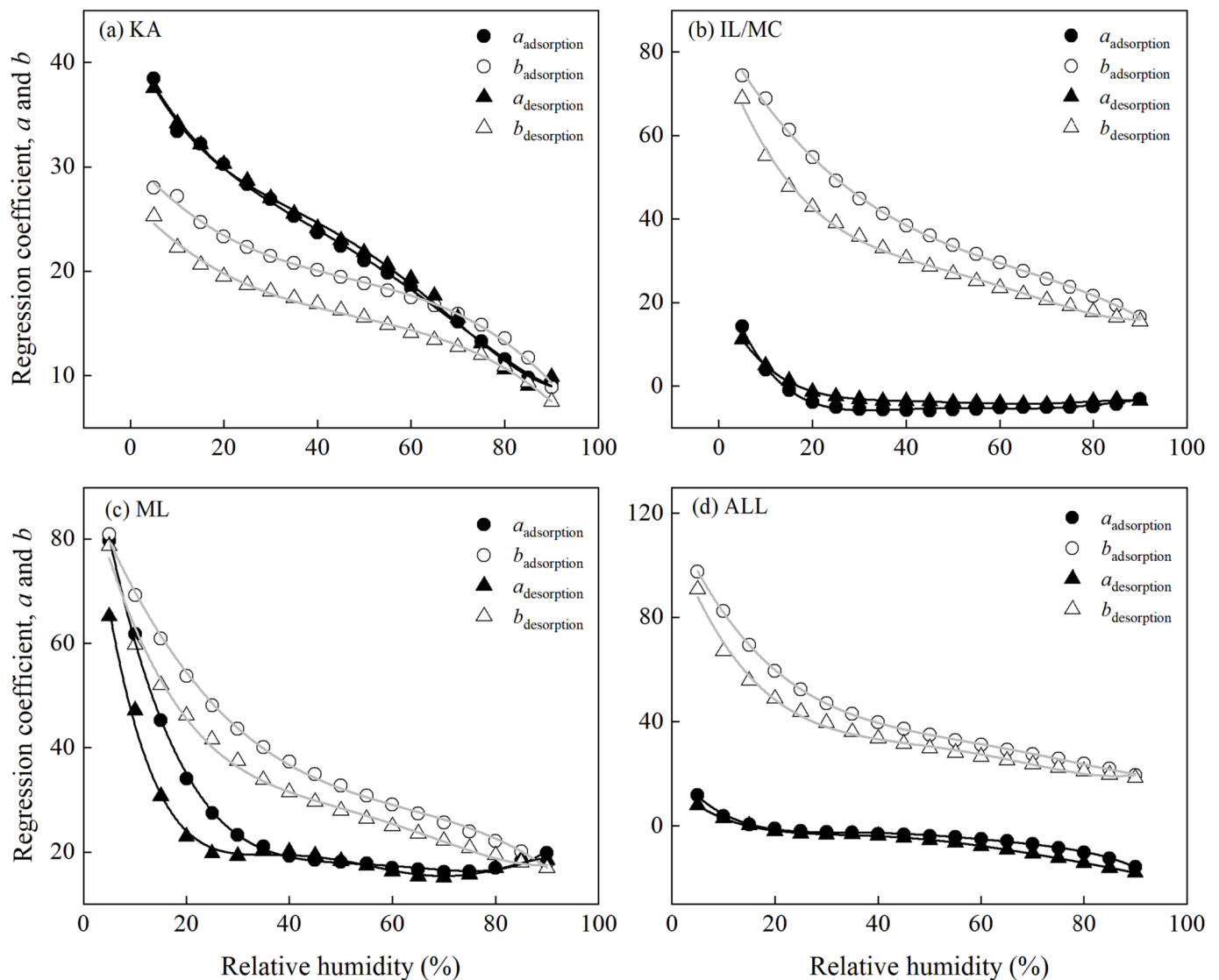


Fig. 3. (a) Relationship between relative humidity and the regression coefficients (a and b) for the calibration datasets of the three sample groups and for all samples combined. The equations for the polynomial regressions are provided in Table 2. KA, kaolinite-rich samples; IL/MC, illite-rich or mixed clay samples; ML, montmorillonite-rich samples; ALL, all samples combined.

datasets, the measured SSA was regressed individually against the w_h as the predictor variable (i.e., 18 w_h values between 5% and 90% RH with increments of 5% for adsorption and desorption). The analysis was conducted using the “lm” function in R v4.2.2 to fit a linear model to the data. The general form of the regression model was:

$$SSA = a + b \times w_h \tag{1}$$

where SSA is in m^2/g , w_h is the hygroscopic water content at a given RH (%), and a and b are the intercept and slope of the regression model, respectively.

After establishing the 18 SSA equations each for adsorption and desorption, we found that a and b both varied markedly with increasing RH (Tables 2 and 3), so a polynomial function was fitted to the RH- a and RH- b data to allow estimation of the two parameters at an arbitrary RH value other than the 18 used.

2.3.3.2. Pedotransfer functions (PTFs) from other soil properties. The four calibration datasets were also used to develop SSA PTFs based on either a combination of clay, silt and SOC contents (denoted as CSS) or CEC. The general forms of the models were:

$$SSA = a + b \times Clay + c \times Silt + d \times SOC \tag{2}$$

$$SSA = a + b \times CEC \tag{3}$$

where SSA is in m^2/g , a , b , c , and d are the model parameters. The clay, silt and SOC contents are in %, and CEC is in $cmol_{(+)} kg^{-1}$. For Eq. (2), clay, silt or SOC contents were excluded if they did not significantly contribute (i.e., $p < 0.05$) to the improvement of model performance.

2.3.4. Evaluation of model performance

The SSA of the samples for the four validation datasets was estimated from the models developed from w_h and PTFs. To compare the performance of the developed models, the root mean square error (RMSE) and the Nash-Sutcliffe coefficient (NSE) were used and are given as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (SSA_m - SSA_e)^2}{n}} \tag{4}$$

$$NSE = 1 - \frac{\sum_{i=1}^n (SSA_m - SSA_e)^2}{\sum_{i=1}^n (SSA_e - \overline{SSA_m})^2} \tag{5}$$

Table 4

Polynomial functions to estimate the value of the regression coefficients (a & b) for adsorption and desorption at arbitrary relative humidity (x) for the calibration datasets of the three sample groups and for all groups together. The regression equation for soil specific surface area (SSA) is of the form $SSA = a + b \times w_h$. The a and b values for 18 specific relative humidity levels are provided in Table 2 and 3.

Sample group	S_d	Regression coefficient	Equation (x = relative humidity)	R_{adj}^2	
KA	adsorption	a	$42.00 - 0.92x + 0.02x^2 - 2.78 \times 10^{-4}x^3 + 1.28 \times 10^{-6}x^4$	0.997	
		b	$30.89 - 0.52x + 8.86 \times 10^{-3}x^2 - 6.33 \times 10^{-5}x^3$	0.997	
	desorption	a	$42.08 - 1.00x + 0.03x^2 - 3.80 \times 10^{-4}x^3 + 1.83 \times 10^{-6}x^4$	0.995	
		b	$26.85 - 0.49x + 8.02 \times 10^{-3}x^2 - 5.49 \times 10^{-5}x^3$	0.995	
IL/MC	adsorption	a	$28.06 - 3.45x + 0.13x^2 - 2.44 \times 10^{-3}x^3 + 2.13 \times 10^{-5}x^4 - 7.05 \times 10^{-8}x^5$	0.993	
		b	$84.36 - 1.91x + 0.02x^2 - 1.23 \times 10^{-4}x^3$	0.999	
	desorption	a	$20.77 - 2.31x + 0.09x^2 - 1.65 \times 10^{-3}x^3 + 1.50 \times 10^{-5}x^4 - 5.24 \times 10^{-8}x^5$	0.999	
		b	$81.16 - 3.10x + 0.08x^2 - 8.82 \times 10^{-4}x^3 + 3.80 \times 10^{-6}x^4$	0.997	
	ML	adsorption	a	$106.95 - 6.05x + 0.15x^2 - 1.76 \times 10^{-3}x^3 + 7.43 \times 10^{-6}x^4$	0.999
			b	$91.07 - 2.45x + 0.03x^2 - 1.84 \times 10^{-4}x^3$	0.999
desorption		a	$101.39 - 8.55x + 0.34x^2 - 6.57 \times 10^{-3}x^3 + 5.93 \times 10^{-5}x^4 - 2.03 \times 10^{-7}x^5$	0.993	
		b	$93.60 - 3.90x + 0.10x^2 - 1.09 \times 10^{-3}x^3 + 4.58 \times 10^{-6}x^4$	0.994	
ALL	adsorption	a	$22.69 - 2.77x + 0.12x^2 - 2.40 \times 10^{-3}x^3 + 2.32 \times 10^{-5}x^4 - 8.73 \times 10^{-8}x^5$	0.997	
		b	$117.37 - 4.39x + 0.09x^2 - 9.32 \times 10^{-4}x^3 + 3.44 \times 10^{-6}x^4$	0.999	
	desorption	a	$15.67 - 1.89x + 0.07x^2 - 1.37 \times 10^{-3}x^3 + 1.16 \times 10^{-5}x^4 - 3.72 \times 10^{-8}x^5$	0.999	
		b	$110.94 - 5.27x + 0.14x^2 - 1.62 \times 10^{-3}x^3 + 6.97 \times 10^{-6}x^4$	0.992	

The p -value for all polynomial regressions was <0.001 . KA, kaolinite-rich samples; IL/MC, illite-rich or mixed clay samples; ML, montmorillonite-rich samples; ALL, all samples combined; R_{adj}^2 , r-squared adjusted for the number of model variables;

where n is the number of samples in the validation dataset, SSA_m is the measured SSA (m^2/g), SSA_e is the estimated SSA (m^2/g), \overline{SSA}_m is the mean of measured SSA (m^2/g). The lower the $RMSE$ value, the better the model performance. The NSE can range from $-\infty$ to 1, with 1 indicating perfect agreement and a number <0 indicating a model with less predictive power than the mean of the measured values.

3. Results

3.1. Calibration and validation datasets

The investigated samples exhibited a large variability in SSA and other soil properties (Table 1). The particle size distribution of the samples covered all USDA soil textural classes (Fig. 1). The KA, IL/MC and ML samples had SOC contents ranging from 0 to 4.9%, 0.1 to 4.0%, and 0.1 to 6.8%, respectively, and the CEC ranged 1 to 41 $cmol_{(+)} kg^{-1}$, 2 to 75 $cmol_{(+)} kg^{-1}$, and 11 to 87 $cmol_{(+)} kg^{-1}$, respectively. Similarly, The KA, IL/MC and ML samples had SSA ranging from 13 to 130 m^2/g , 6 to 314 m^2/g , and 40 to 440 m^2/g , respectively. The partitioning of the dataset into the calibration and validation data sets yielded evenly distributed groups (Fig. 1), which is essential to demonstrate the broad

applicability of the developed models.

3.2. Models based on hygroscopic water content

3.2.1. Relationship between SSA and w_h

The strong link between SSA and w_h is exemplified by the water vapor sorption isotherms of samples that have identical clay content but significantly differ in clay mineralogy (Fig. 2a). For example, the ML samples exhibited significantly higher SSA (287 m^2/g) than the KA (44 m^2/g) and IL/MC (151 m^2/g) samples at identical clay contents (43%) (Fig. 2a). Fig. 2b and 2c, depict the relationship between the SSA and w_h at 50% RH for adsorption (w_{h50a}) and desorption (w_{h50d}). For all the calibration samples, the SSA increased linearly with both w_{h50a} and w_{h50d} . The same trend was also observed for the SSA- w_h relationship at other RH values (Figures S1 and S2).

3.2.2. Developed models considering sorption hysteresis

The developed calibration models for the four sample groups (KA, IL/MC, ML, ALL) based on adsorption and desorption water contents are presented in Tables 2 and 3. For all models, the coefficient of determination (R^2) of the relationship between SSA and w_h ranged from 0.67 to 0.95. It first increased for RH from 5% to 35% and then stabilized for RH from 35% to 90%. For the two sorption directions, the average R^2 for the entire RH range for adsorption for the KA, IL/MC, ML and ALL calibration datasets was 0.74, 0.85, 0.83 and 0.92, respectively (Table 2). Similarly, the average R^2 for the entire RH range for desorption for the KA, IL/MC, ML and ALL calibration datasets was 0.73, 0.84, 0.84 and 0.92, respectively (Table 3).

For the calibration models of each sample group, we found that the intercept a and slope b parameters both changed markedly with increasing RH. Thus, we used a polynomial function to fit the a and b values for each sample group (Fig. 3). The polynomial functions accurately fitted the a -RH and b -RH data for all sample groups. For adsorption the R_{adj}^2 values ranged from 0.993 to 0.999, and for desorption R_{adj}^2 ranged from 0.992 and 0.999 with $p < 0.001$ (Table 4).

3.2.3. Model validation

The reason for combining all data (i.e., the ALL sample group) for model development was to investigate if the impact of clay mineralogy on model performance was significant enough to warrant developing separate models for different clay mineral groups. Therefore, for each sample group (KA, IL/MC, ML), the validation of the respective calibration models was also compared to the ALL model validation (Fig. 4). For example, in Fig. 4a, the SSA of the 12 KA validation samples were estimated from the KA calibration model as well as from the ALL-calibration model. A comparison of the measured SSA and estimated SSA from adsorption and desorption w_h measured at RH of 50% and 70%, respectively, is presented in Fig. 4, and for all RH levels and sorption directions in Fig. S3. In general, there was a slightly better estimation of SSA based on the w_h at 70% RH when compared to 50% RH for all models (except for IL/MC-desorption and ML-desorption). When the whole RH range was considered, we found that the $RMSE$ for SSA decreased and then increased with increasing RH (Fig. S3). The optimum RH for SSA estimation was approximately 70% for the KA and IL/MC samples, and 40 to 50% for the ML samples.

For the KA and IL/MC samples, the performance of the KA-calibration model (average $RMSE = 10.5 m^2/g$ and $11.1 m^2/g$, respectively, for adsorption and desorption) and the IL/MC-calibration model (average $RMSE = 21.3 m^2/g$ and $22.0 m^2/g$, respectively) were better than the ALL-calibration model (KA: average $RMSE = 18.7 m^2/g$ and $22.0 m^2/g$, respectively; ML: average $RMSE = 22.4 m^2/g$ and $23.3 m^2/g$, respectively) (Fig. S3). In addition, similar model performance was observed for the ML samples, regardless of whether the validation was done using the ML-calibration model (average $RMSE = 41.4 m^2/g$ and $38.9 m^2/g$, respectively) or the ALL-calibration model (average $RMSE =$

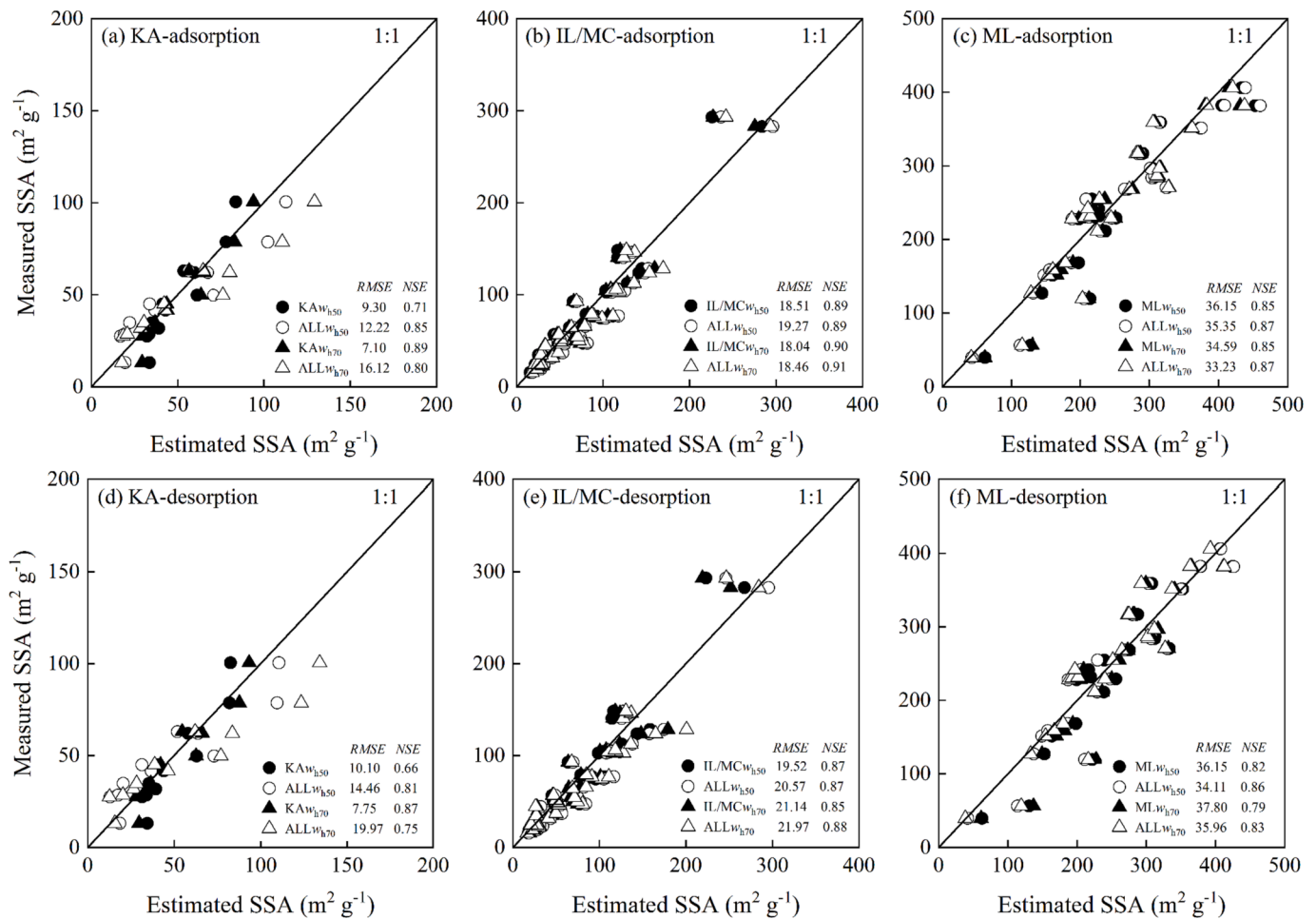


Fig. 4. Measured versus estimated soil specific surface area (SSA) for the three sample groups based on regression models of adsorption (top row) and desorption (bottom row) water contents at 50% (w_{h50}) and 70% (w_{h70}). RMSE, root mean square error (m^2/g); NSE, Nash-Sutcliffe coefficient. KA, kaolinite-rich samples, IL/MC, illite-rich or mixed clay samples; ML, montmorillonite-rich samples; ALL, all samples combined.

Table 5

Models for soil specific surface area (SSA) based on clay, silt and soil organic carbon contents (SOC) (CSS) and based on cation exchange capacity (CEC) for the calibration datasets of the three sample groups and for all samples combined.

Sample group	Equation	R_{adj}^2	RMSE
KA	$SSA = 4.11 + 1.24 \times Clay^{***}$	0.75	13.51
	$SSA = 28.87 + 2.00 \times CEC^{***}$	0.38	21.19
IL/MC	$SSA = -12.53 + 3.20 \times Clay^{***} + 6.76 \times SOC^*$	0.74	29.51
	$SSA = 6.29 + 3.68 \times CEC^{***}$	0.59	37.30
ML	$SSA = 16.85 + 4.19 \times Clay^{***} + 1.49 Silt^{**} - 19.18 \times SOC^{***}$	0.64	51.75
	$SSA = 28.19 + 4.54 \times CEC^{***}$	0.66	50.15
ALL	$SSA = -45.22 + 4.09 \times Clay^{***} + 1.02 Silt^{***}$	0.56	70.59
	$SSA = -5.76 + 4.97 \times CEC^{***}$	0.81	45.93

Clay, silt and SOC are in %; CEC is in $cmol_{(+) } kg^{-1}$; SSA is in m^2/g . All regression models exhibited an overall $p < 0.001$. *** ($p < 0.001$), ** ($p < 0.01$), and * ($p < 0.05$) represent the statistical significance of SOC, clay, silt, and CEC in multiple linear regression, respectively. R_{adj}^2 , r-squared adjusted for the number of model variables; RMSE, root mean square error (m^2/g). The P-value for all regressions was < 0.001 . KA, kaolinite-rich samples, IL/MC, illite-rich or mixed clay samples; ML, montmorillonite-rich samples; ALL, all samples combined.

41.1 m^2/g and 36.8 m^2/g , respectively) (Fig. S3).

3.3. Models based on other soil properties

In Table 5, the PTFs for SSA based on clay, silt and SOC contents (denoted as CSS), and based on CEC are presented. Clay content was a significant estimator for SSA for all sample groups ($p < 0.001$). For the KA and ALL sample groups, SOC and or silt were excluded from the PTFs because they did not contribute to an increase in model accuracy ($p > 0.05$). The accuracy of the PTFs, quantified by the R_{adj}^2 and RMSE was less accurate for the ALL group when compared to the other three groups (Table 5). Visually, there seemed to be little difference between the validation of SSA either based on the sample-specific PTFs or the ALL-sample PTFs (Fig. 5). Despite this, for validation, the ALL-sample PTFs were less accurate (RMSE from 41.9 to 49.8 m^2/g) than the sample-specific PTFs (Fig. 5; RMSE from 19.1 to 49.8 m^2/g).

In general, higher CEC also was linked to higher SSA in all sample groups. The PTFs developed based on the ALL sample group indicated a strong relationship ($R_{adj}^2 = 0.83$) between CEC and SSA, while for the other sample groups (KA, IL/MC, ML), this relationship was weaker ($R_{adj}^2 \leq 0.66$). However, the CEC was more strongly related to SSA than to CSS for the ALL group. For validation, the PTFs based on CEC for the KA and ALL sample groups were less accurate for KA-rich samples ($NSE \leq 0.36$) (Fig. 5a). For IL/MC samples, the estimation accuracy of the PTFs based on CEC of the ALL sample group was better than that for the IL/MC samples (Fig. 5b). For the ML samples, however, the estimation

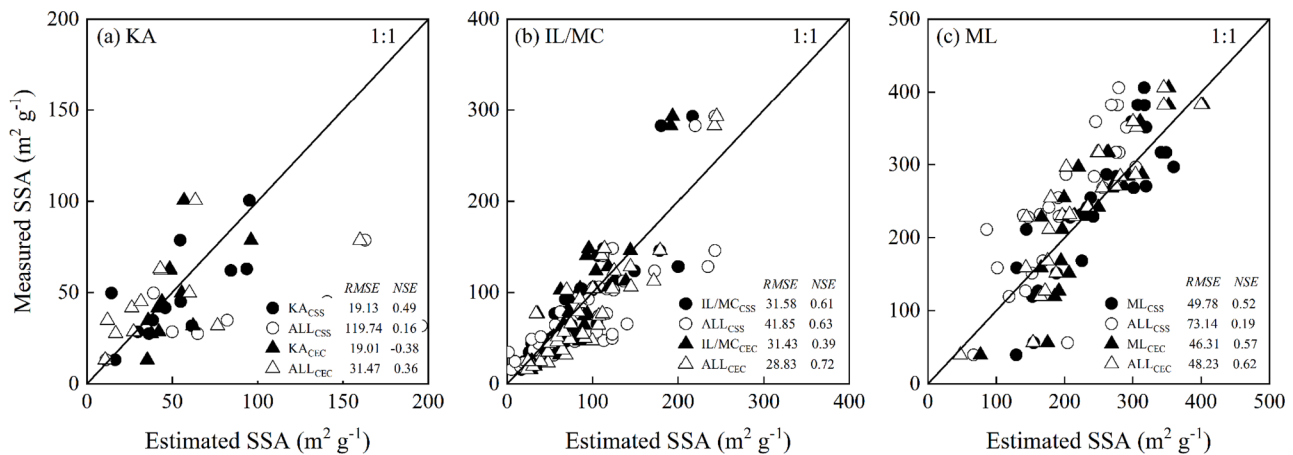


Fig. 5. Measured versus estimated soil specific surface area (SSA) for the three sample groups based on regression models derived from a combination of clay, silt and soil organic carbon contents (CSS), and cation exchange capacity (CEC). RMSE, root mean square error (m^2/g); NSE, Nash-Sutcliffe coefficient. KA, kaolinite-rich samples, IL/MC, illite-rich or mixed clay samples; ML, montmorillonite-rich samples; ALL, all samples combined.

Table 6

Models from literature for estimating the EGME soil specific surface area (SSA) from soil clay, silt and sand contents (%).

Equation	R^2	Reference
$SSA = 5.78 \times \text{Clay} + 15.054$	0.90	Banin and Amiel (1970)
$SSA = 0.40 \times \text{Clay} + 86$	0.86	de Jong (1999)
$SSA = 0.042 + 4.23 \times \text{Clay} + 1.12 \times \text{Silt} - 1.16 \times \text{Sand}$	0.56	Ersahin et al. (2006)

accuracy of the PTFs based on CEC was slightly better than for the ALL samples group (Fig. 5c).

A comparison of the w_h -based models (Fig. 4) and the models based on CSS and CEC (Fig. 5) showed that the models for SSA based on CSS and CEC were less accurate than the w_h -based models. This was the case for all sample groups, regardless of which w_h -based model was used (sample-specific or ALL sample models).

3.4. Evaluation of models from the literature

Existing PTFs for estimating the EGME SSA from three studies were

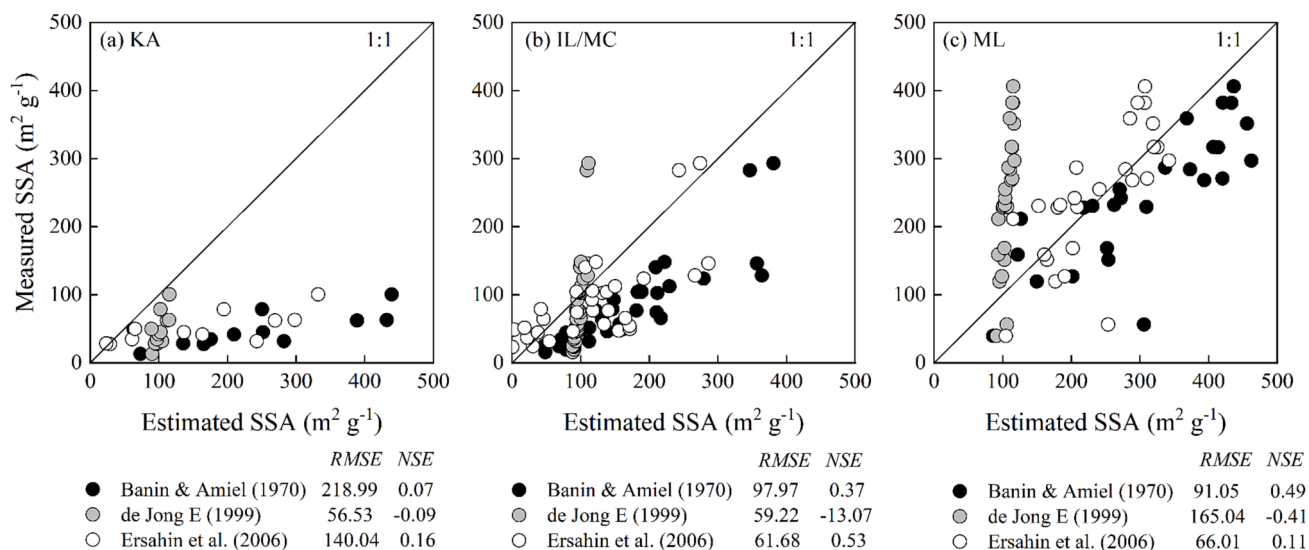


Fig. 6. Performance of SSA pedotransfer functions from literature (Table 6) for the validation datasets for the three sample groups. RMSE, root mean square error (m^2/g); NSE, Nash-Sutcliffe coefficient. KA, kaolinite-rich samples, IL/MC, illite-rich or mixed clay samples; ML, montmorillonite-rich samples.

evaluated and their performance was compared with the models developed in this study (Table 6). The three literature models estimated the SSA for different clay mineralogy with less accuracy and lower NSE values (Fig. 6) than the models based on w_h . Although the CSS models presented here used the same soil properties as used in the models from literature, the performance of the proposed CSS models in this study to estimate SSA for different clay mineralogy (Fig. 5) was better than for the models from the literature.

4. Discussion

In this study, it was demonstrated that w_h can be used to estimate SSA, as the w_h is closely related to clay and SOC contents (Wuddivira et al., 2012) and CEC (Arthur, 2017). These two soil properties are strong determinants of SSA for samples that are of similar clay mineralogy (Farrar and Coleman, 1967; Sepaskhah and Tafteh, 2013). In addition, we also found that the developed w_h -based SSA models for different clay mineralogy all have good estimation accuracy (Fig. 4). Thus, to apply the w_h -based models under RH conditions that do not correspond to the 18 values used in this study, the polynomial functions presented in Fig. 3 can be utilized to obtain the model parameters (a and

b) at an arbitrary RH and sorption direction. The purpose was that this effectively overcomes the limitation of wetting or drying soils to pre-determined RH before applying the w_h -based models, and similar approach have been reported in other studies for different soil properties (Arthur et al., 2021; Yan et al., 2022). Importantly, this approach w_h is very cheap and easy to apply because the SSA estimation models in Tables 2,3 and 4 require only a regular drying oven and a relative humidity meter. When the CSS and CEC of the investigated samples are already available, the developed models based on CSS and CEC (Table 5) can also be applied for estimating SSA. However, when new soil samples are obtained, the CSS and CEC still need to be measured, making it less cost-effective than directly measuring SSA. Thus, the w_h -based models constitute a much faster SSA estimation approach.

Our findings showed that clay mineralogy has a significant impact on the accuracy of SSA estimates, which suggests that separate models should be developed for each soil sample group. The poor performance of the w_h -based models from the ALL-calibration dataset for KA and IL/MC samples may be due to several reasons. First, the investigated samples differed in the shape of their isotherms, soil water content, and the magnitude of hysteresis. For example, the KA sample isotherms can be classified as Type III, while the IL/MC and ML samples exhibited Type II and Type IV isotherm shapes (Fig. 2; Brunauer et al., 1940). In addition, the KA sample isotherm is essentially non-hysteretic, whereas IL/MC and ML samples show hysteresis across the full RH range (Arthur et al., 2019). Although the w_h -based models from the ALL-calibration group dataset captured a wide range in soil textures and clay mineralogy, the differences in the shape and hysteresis of the isotherms for the investigated samples may cause the variations in the model performance among the KA, IL/MC and ML samples. Second, the distribution of calibration datasets from the three sample groups (KA, IL/MC and ML) varied in the ALL-calibration models. For instance, the developed w_h -based models from the ALL calibration dataset were dominated by soils with 2:1 clay minerals that tend to swell (e.g., IL/MC and ML samples) (Kumari and Mohan, 2021) that account for 84% of the total investigated samples (Table 1). Thus, the developed model may not accurately estimate the SSA for soils with 1:1 non-swelling clay minerals (e.g., KA samples), but additional studies are required from this standpoint. Previous studies have demonstrated the significance of incorporating clay mineralogy in soil property estimation models based on w_h (Wuddivira et al., 2012; Arthur et al., 2015; Arthur et al., 2019). Therefore, it is expected that we need to develop separate models for different clay mineralogy to enhance the accuracy of SSA estimation.

We found that both clay content and CEC were significant variables in estimating SSA for all sample groups (Table 5), which is in agreement with previous studies (Farrar and Coleman, 1967; Ersahin et al., 2006). In this study, the sand content of the investigated samples was excluded from the analyses because it was collinear with clay content. The developed PTFs based on CSS and CEC estimated SSA with varying accuracy for the samples with different mineralogy. One plausible explanation is that the Pearson correlation coefficients between the clay, silt, and SOC contents, CEC and SSA are different for each sample group (Fig. S4). In general, it is also difficult to develop a single model based on CSS or CEC to estimate SSA for different clay mineralogy because of the large SSA range. This is supported by the weak model performance observed when a PTF developed disregarding clay mineralogy (ALL) is used to estimate SSA for the sample groups with distinct clay mineralogy (KA, IL/MC or ML). This is because swelling clays (e.g., found in the IL/MC and ML samples) have SSA up to 810 m^2/g for pure montmorillonite, whereas non-swelling clay minerals (e.g., found in the KA samples) typically have SSA ranging from 10 to 40 m^2/g (Garzón and Sánchez-Soto, 2015). However, the particle-size distribution and CEC do not well reflect these differences in SSA among different clay mineralogy. Nevertheless, all the developed PTFs based on CSS and CEC showed weaker performance compared to w_h -based models. It may be that w_h better explains the variability in the SSA, compared to particle-size distribution, SOC content and CEC. Thus, this result provides a

guideline for future studies that is to prioritize the use of w_h -based models to improve the estimation accuracy of SSA values for different clay mineralogy.

Despite that the PTFs from the literature utilized the same soil properties as in the present study (Banin and Amiel, 1970; de Jong, 1999; Ersahin et al., 2006), the performance of the developed CSS-based PTFs (Fig. 5) was better than the models from literature (Fig. 6). One plausible explanation would be that the large variability in the samples used for developing the models in this study. Furthermore, when the models developed from one clay mineralogy are used to estimate the SSA of samples from a different clay mineralogy, none of the three literature models captured the wide range of soil textures and clay mineralogy. Thus, caution is required when applying the PTFs for samples that are significantly different from those used in the development of the models.

5. Conclusions

We proposed regression models to estimate the soil specific surface area (SSA) from hygroscopic water content (w_h) for three soil sample groups with different dominant clay minerals: kaolinite-rich samples (KA), illite-rich or mixed clay samples (IL/MC), and montmorillonite-rich samples (ML). Furthermore, we compared the developed w_h models with models based on clay, silt and soil organic carbon contents, and cation exchange capacity. The w_h -based models accurately estimated SSA for the KA, IL/MC and ML samples regardless of the RH at which w_h was measured or the sorption direction. However, disregarding clay mineralogy in the model development (ALL-calibration models) led to less SSA estimation accuracy. Nevertheless, the w_h -based models for estimating SSA based on differences in clay mineralogy were superior to clay, silt, soil organic carbon, and cation exchange capacity-based pedotransfer functions developed in the study or extracted from literature. Thus, for studies that require a large number of SSA measurements, the w_h -based models proved useful as they only require a relative humidity meter and a drying oven.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Co-author Markus Tuller gratefully acknowledges support from the USDA NIFA Hatch/Multi-State project # ARZT-1370600-R21-189. We also acknowledge the China Scholarship Council's support to Fulai Yan (No. 202006300074) for his study at Aarhus University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.geoderma.2023.116614>.

References

- Adamson, A.W., Gast, A.P., 1990. *Physical Chemistry Of Surfaces*. John Wiley and Sons Inc., New York.
- Akin, I.D., Likos, W.J., 2014. Specific surface area of clay using water vapor and EGME sorption methods. *Geotech. Test. J.* 37 (6), 1016–1027.
- Arthur, E., 2017. Rapid estimation of cation exchange capacity from soil water content. *Eur. J. Soil Sci.* 68 (3), 365–373.

- Arthur, E., Tuller, M., Moldrup, P., Resurreccion, A.C., Meding, M.S., Kawamoto, K., Komatsu, T., de Jonge, L.W., 2013. Soil specific surface area and non-singularity of soil-water retention at low saturations. *Soil Sci. Soc. Am. J.* 77 (1), 43–53.
- Arthur, E., Tuller, M., Moldrup, P., de Jonge, L.W., 2014. Evaluation of a fully automated analyzer for rapid measurement of water vapour sorption isotherms for applications in soil science. *Soil Sci. Soc. Am. J.* 78 (3), 754–760.
- Arthur, E., Tuller, M., Moldrup, P., Jensen, D.K., De Jonge, L.W., 2015. Prediction of clay content from water vapour sorption isotherms considering hysteresis and soil organic matter content. *Eur. J. Soil Sci.* 66 (1), 206–217.
- Arthur, E., Tuller, M., Moldrup, P., Greve, M.H., Knadel, M., de Jonge, L.W., 2018. Applicability of the Guggenheim-Anderson-Boer water vapour sorption model for estimation of soil specific surface area. *Eur. J. Soil Sci.* 69 (2), 245–255.
- Arthur, E., Tuller, M., Norgaard, T., Moldrup, P., de Jonge, L.W., 2019. Improved estimation of clay content from water content for soils rich in smectite and kaolinite. *Geoderma* 350, 40–45.
- Arthur, E., Rehman, H.U., Tuller, M., Pouladi, N., Norgaard, T., Moldrup, P., de Jonge, L.W., 2021. Estimating Atterberg limits of soils from hygroscopic water content. *Geoderma* 381, 114698.
- Arthur, E., Tuller, M., Norgaard, T., Moldrup, P., Chen, C., Ur Rehman, H., Weber, P.L., Knadel, M., Wollesen de Jonge, L., 2023. Contribution of organic carbon to the total specific surface area of soils with varying clay mineralogy. *Geoderma* 430, 116314.
- Banin, A., Amiel, A., 1970. A correlative study of the chemical and physical properties of a group of natural soils of Israel. *Geoderma* 3 (3), 185–198.
- Bashina, A.S., Smagin, A.V., Sadovnikova, N.B., Belyaeva, E.A., Korchagina, K.V., 2019. Specific surface area of strongly swelling polymer hydrogels for agriculture. In *IOP Conference Series: Earth and Environmental Science* (Vol. 368, No. 1, p. 012006). IOP Publishing. <https://doi.org/10.1088/1755-1315/368/1/012006>.
- Brunauer, S., Emmett, P.H., Teller, E., 1938. Adsorption of gases in multimolecular layers. *J. Am. Chem. Soc.* 60 (2), 309–319.
- Brunauer, S., Deming, L.S., Deming, W.E., Teller, E., 1940. On a theory of the van der Waals adsorption of gases. *J. Am. Chem. Soc.* 62 (7), 1723–1732.
- Carter, D.L., Mortland, M.M., Kemper, W.D., 1986. Specific surface. *Methods of Soil Analysis: Part 1 Physical and Mineralogical. Methods* 5, 413–423.
- Chen, C., Arthur, E., Tuller, M., Zhou, H., Wang, X., Shang, J., Hu, K., Ren, T., 2021. Estimation of soil specific surface area from adsorbed soil water content. *Eur. J. Soil Sci.* 72 (4), 1718–1725.
- de Jong, E., 1999. Comparison of three methods of measuring surface area of soils. *Can. J. Soil Sci.* 79 (2), 345–351.
- Drever, J.I., 1973. The preparation of oriented clay mineral specimens for X-ray diffraction analysis by a filter-membrane peel technique. *Am. Mineral.* 58 (5–6), 553–554.
- Eberl, D.D., 2003. User guide to RockJock-A program for determining quantitative mineralogy from X-ray diffraction data (No. 2003 – 78). US Geological Survey.
- Ersahin, S., Gunal, H., Kutlu, T., Yetgin, B., Coban, S., 2006. Estimating specific surface area and cation exchange capacity in soils using fractal dimension of particle-size distribution. *Geoderma* 136 (3–4), 588–597.
- Farrar, D.M., Coleman, J.D., 1967. The correlation of surface area with other properties of nineteen British clay soils. *Eur. J. Soil Sci.* 18 (1), 118–124.
- Garzón, E., Sánchez-Soto, P.J., 2015. An improved method for determining the external specific surface area and the plasticity index of clayey samples based on a simplified method for non-swelling fine-grained soils. *Appl. Clay Sci.* 115, 97–107.
- Gee, G.W., Or, D., 2002. Particle-size analysis. In: J.H. Dane, G.C. Topp (Eds.), *Methods of soil analysis. Part 4. SSSA Book Series No. 5. SSSA, Madison, WI*, pp. 255 – 293.
- Ghanbarian, B., Hunt, A.G., Bittelli, M., Tuller, M., Arthur, E., 2021. Estimating specific surface area: Incorporating the effect of surface roughness and probing molecule size. *Soil Sci. Soc. Am. J.* 85 (3), 534–545.
- Hang, P.T., Brindley, G.W., 1970. Methylene blue absorption by clay minerals. Determination of surface areas and cation exchange capacities (clay-organic studies XVIII). *Clay Clay Miner.* 18, 203–212.
- Heister, K., 2014. The measurement of the specific surface area of soils by gas and polar liquid adsorption methods – Limitations and potentials. *Geoderma* 216, 75–87.
- Hepper, E.N., Buschiazzi, D.E., Hevia, G.G., Urioste, A., Antón, L., 2006. Clay mineralogy, cation exchange capacity and specific surface area of loess soils with different volcanic ash contents. *Geoderma* 135, 216–223.
- Khorshidi, M.S., Lu, N., 2017. Intrinsic relation between soil water retention and cation exchange capacity. *J. Geotech. Geoenviron.* 143, 04016119.
- Kuhn, M., 2019. caret: Classification and Regression Training. R package version 6.0 – 84. <https://CRAN.R-project.org/package=caret>.
- Kumari, N., Mohan, C., 2021. Basics of clay minerals and their characteristic properties. *Clay Clay Miner.* 24, 1–29.
- Lu, N., Zhang, C., 2020. Separating external and internal surface areas of soil particles. *J. Geotech. Geoenviron.* 146 (2), 04019126.
- Maček, M., Mauko, A., Mladenović, A., Majes, B., Petković, A., 2013. A comparison of methods used to characterize the soil specific surface area of clays. *Appl. Clay Sci.* 83, 144–152.
- Macht, F., Eusterhues, K., Pronk, G.J., Totsche, K.U., 2011. Specific surface area of clay minerals: Comparison between atomic force microscopy measurements and bulk-gas (N₂) and-liquid (EGME) adsorption methods. *Appl. Clay Sci.* 53 (1), 20–26.
- Newman, A.C.D., 1983. The specific surface of soils determined by water sorption. *J. Soil Sci.* 34, 23–32.
- Pennell, K.D., 2002. Specific surface area. In: Dane, J.H., Topp, G.C. (Eds.), *Methods of Soil Analysis, Part 4: Physical Methods. Soil Science Society of America, Madison, WI*, pp. 295–315.
- Quirk, J.P., Murray, R.S., 1999. Appraisal of the ethylene glycol monoethyl ether method for measuring hydratable surface area of clays and soils. *Soil Sci. Soc. Am. J.* 63 (4), 839–849.
- R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/>.
- Rawlins, B.G., Turner, G., Mountney, I., Wildman, G., 2010. Estimating specific surface area of fine stream bed sediments from geochemistry. *Appl. Geochem.* 25 (9), 1291–1300.
- Resurreccion, A.C., Moldrup, P., Tuller, M., Ferre, T.P.A., Kawamoto, K., Komatsu, T., De Jonge, L.W., 2011. Relationship between specific surface area and the dry end of the water retention curve for soils with varying clay and organic carbon contents. *Water Resour. Res.* 47, W06522.
- Sepaskhah, A.R., Tafteh, A., 2013. Pedotransfer function for estimation of soil-specific surface area using soil fractal dimension of improved particle-size distribution. *Arch. Agron. Soil Sci.* 59 (1), 93–103.
- Sposito, G., 1984. The surface chemistry of soil. Oxford University Press, New York, pp. 23–25.
- Sumner, M.E., Miller, W.P., 1996. Cation exchange capacity and exchange coefficients. *Methods of soil analysis: Part 3. Chemical methods* 5, 1201–1229.
- Utkaeva, V.F., 2007. Specific surface area and wetting heat of different soil types in European Russia. *Eurasian Soil Sci.* 40 (11), 1193–1202.
- Wuddivira, M.N., Robinson, D.A., Lebron, I., Bréchet, L., Atwell, M., De Caires, S., Oatham, M., Jones, S.B., Abdu, H., Verma, A.K., Tuller, M., 2012. Estimation of soil clay content from hygroscopic water content measurements. *Soil Sci. Soc. Am. J.* 76 (5), 1529–1535.
- Yan, F., Fu, Y., Tall, A., Zhang, F., Arthur, E., 2022. Coefficient of linear extensibility of soil can be estimated from hygroscopic water content or clay and organic carbon contents. *Eur. J. Soil Sci.* 73 (5), e13298.
- Yukselen-Aksoy, Y., Kaya, A., 2010. Method dependency of relationships between specific surface area and soil physicochemical properties. *Appl. Clay Sci.* 50 (2), 182–190.