

Supplement of Biogeosciences, 16, 2857–2871, 2019  
<https://doi.org/10.5194/bg-16-2857-2019-supplement>  
© Author(s) 2019. This work is distributed under  
the Creative Commons Attribution 4.0 License.



*Supplement of*

## **Global soil–climate–biome diagram: linking surface soil properties to climate and biota**

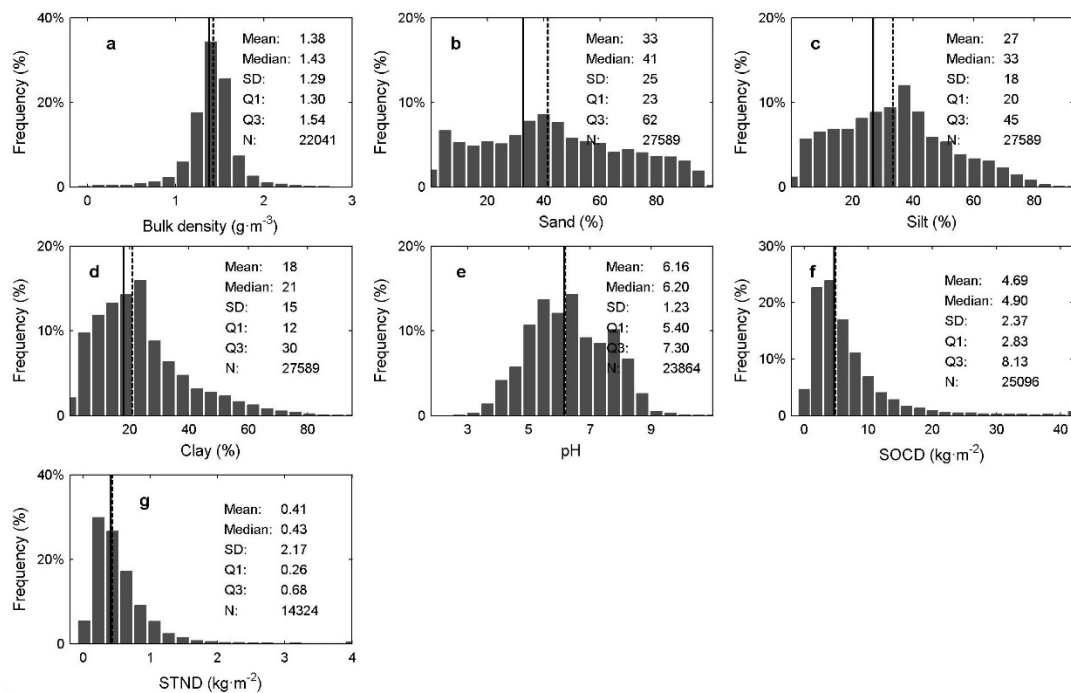
**Xia Zhao et al.**

*Correspondence to:* Jingyun Fang (jyfang@urban.pku.edu.cn)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.

21 **Figure S1.** Frequency distribution of soil properties in the surface soil layer (0-30 cm).  
 22 Mean is geometric mean; SD is standard deviation; Q1 and Q3 is quartile in 25% and  
 23 75%, respectively; N is the number of observations; The black line and dotted line  
 24 indicate mean and median. a: Bulk density ( $\text{g}\cdot\text{m}^{-3}$ ); b: Sand (%); c: Silt (%); d: Clay  
 25 (%); e: pH; f: SOCD ( $\text{kg}\cdot\text{m}^{-2}$ ); g: STND ( $\text{kg}\cdot\text{m}^{-2}$ ).

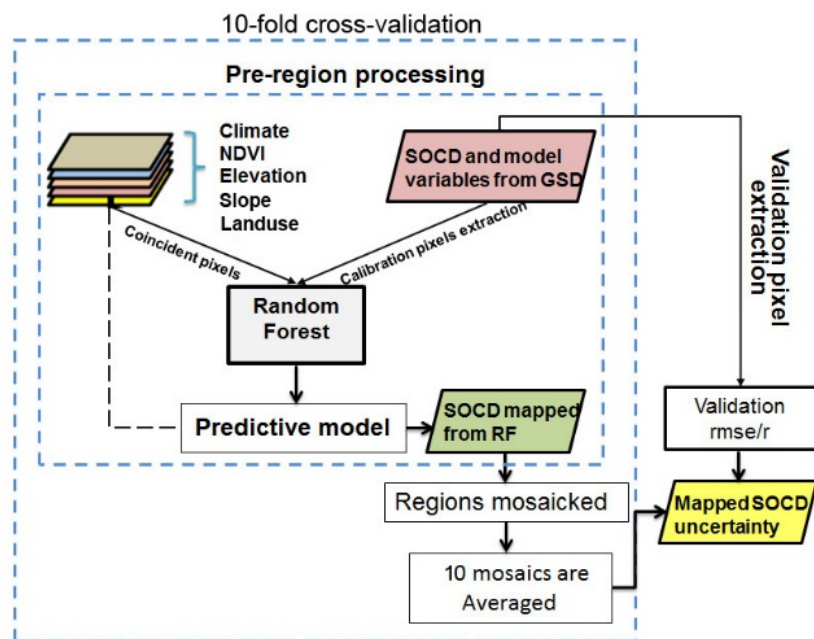
26



27

28 **Figure S2.** The work flow to map soil properties (using soil organic carbon as an  
 29 example). The mapping approach is based on the regional specified Random Forest  
 30 Machine Learning algorithm. The world is divided into 11 regions: Africa, Australia,  
 31 Canada and Alaska, East Asia, Europe, Mexico, Russia, South America, Tropical  
 32 Asia, continental USA, and West Asia. In each region, a diverse set of compiled  
 33 predictors are combined with regional soil organic carbon samples to train the  
 34 regional specified models. Predictions were conducted by using essential surrogate  
 35 variables of climate (e.g. MAT, MAP, temperature seasonality, precipitation  
 36 seasonality), vegetation activity (i.e., NDVI), topography (e.g., elevation, slope) and  
 37 land cover (i.e., land use type) (see method section).

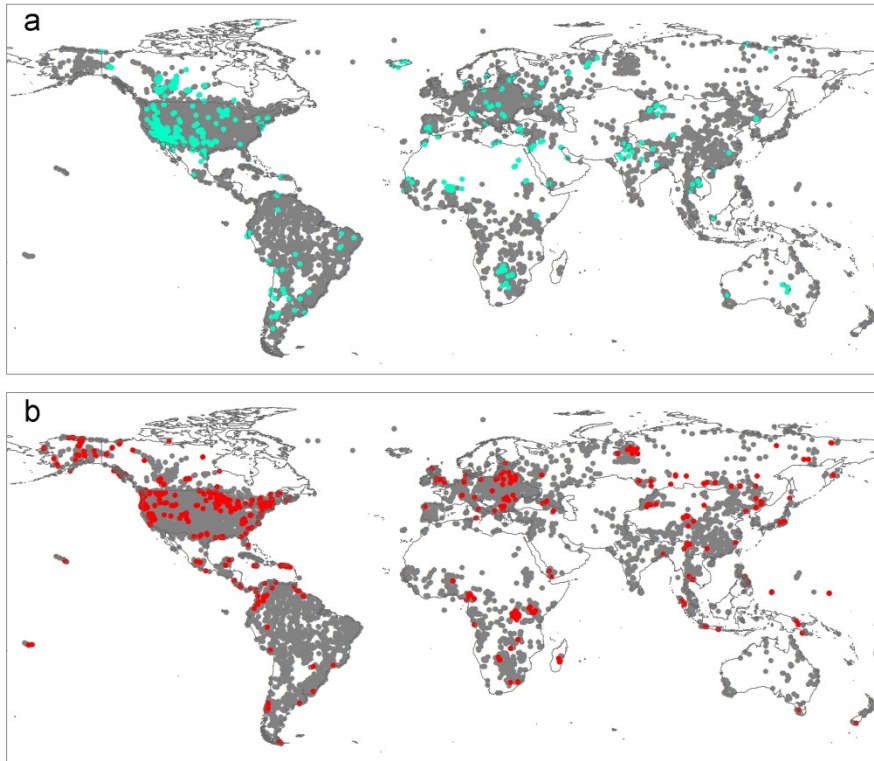
38



39

40 **Figure S3.** Spatial distribution of excluded outliers (taking SOC density as an  
41 example). (a) Data on SOC density below 2.5% quantile, (b) data above 97.5%  
42 quantile. Grey points indicate all sites included in the GSD database.

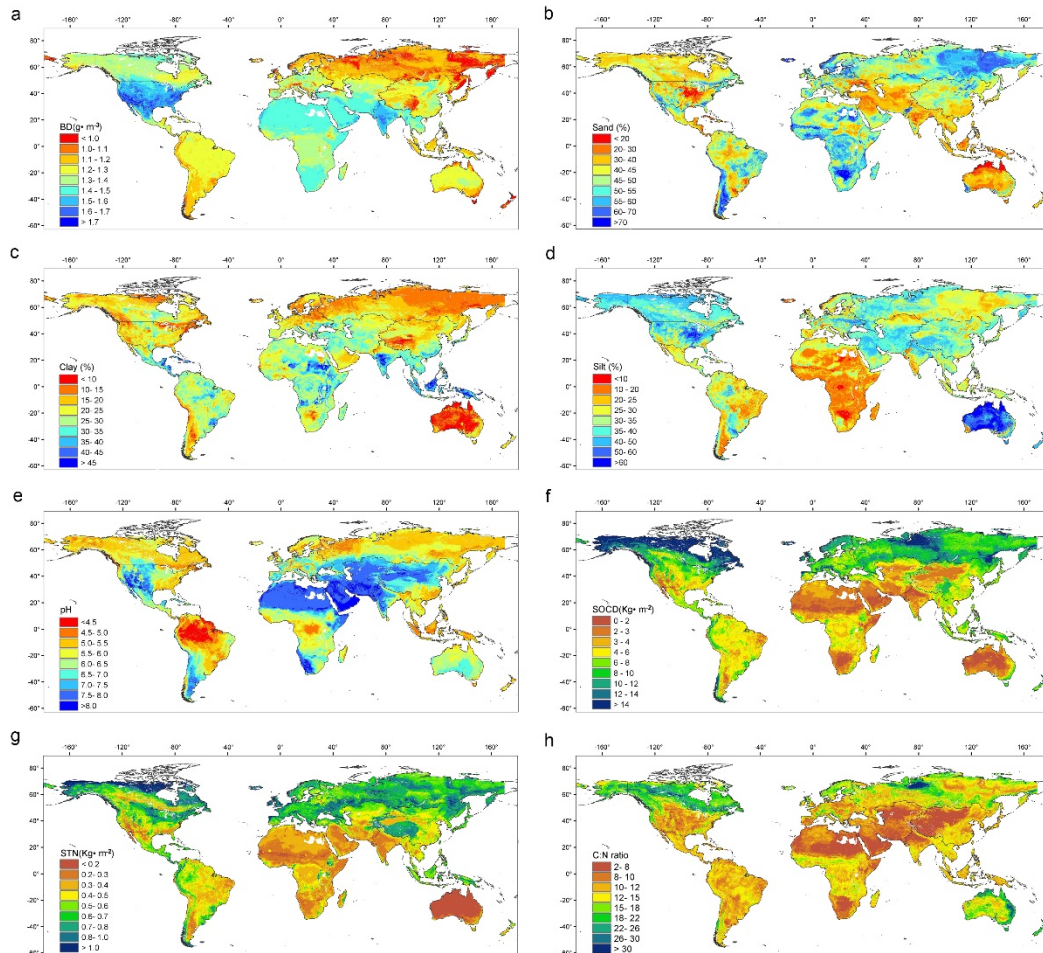
43



44

45

46 **Figure S4.** Maps of worldwide soil properties in the surface soil layer (0–30 cm)  
 47 based on an analysis using whole datasets. a, BD (bulk density,  $\text{g cm}^{-3}$ ); b, Sand  
 48 fraction (%); c, Silt fraction (%); d, Clay fraction (%); e, pH; f, SOCD (soil organic  
 49 carbon density,  $\text{kg m}^{-2}$ ); g, STND (soil total nitrogen density,  $\text{kg m}^{-2}$ ); and h, C:N  
 50 ratio.

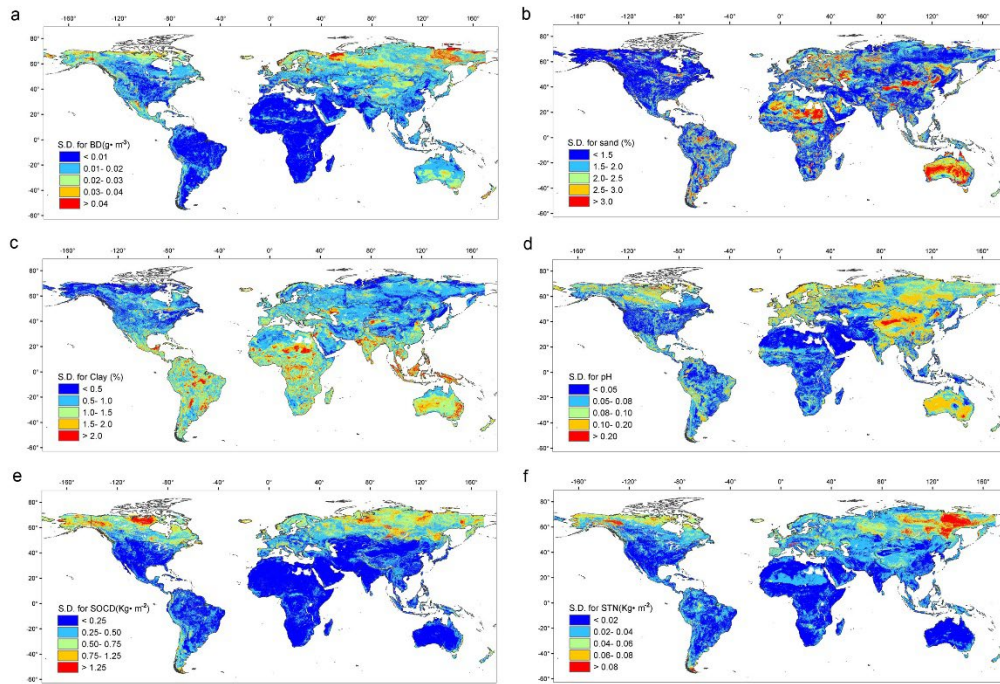


51

52

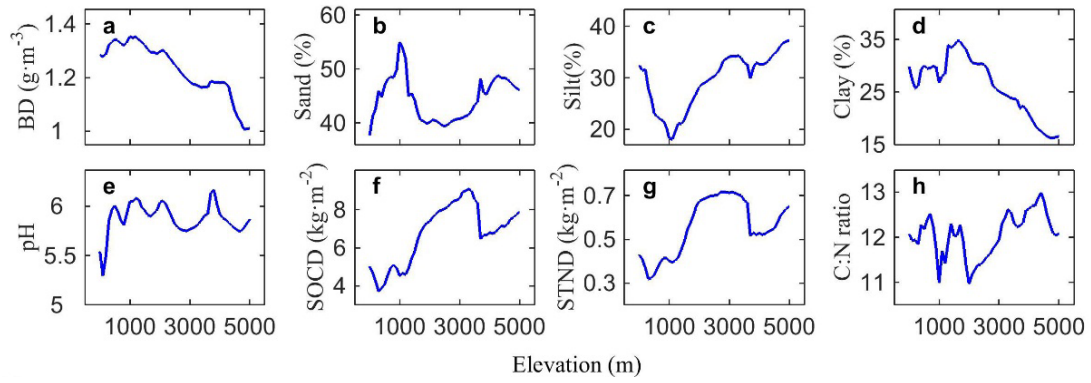
53 **Figure S5.** Spatial pattern of uncertainties (standard deviation, SD) of surface soil  
54 properties estimated by 10-fold cross-validation. a: Bulk density ( $\text{g}\cdot\text{m}^{-3}$ ); b: Sand (%);  
55 c: Clay (%); d: pH e: SOCD ( $\text{kg}\cdot\text{m}^{-2}$ ); f: STND ( $\text{kg}\cdot\text{m}^{-2}$ ).

56



57

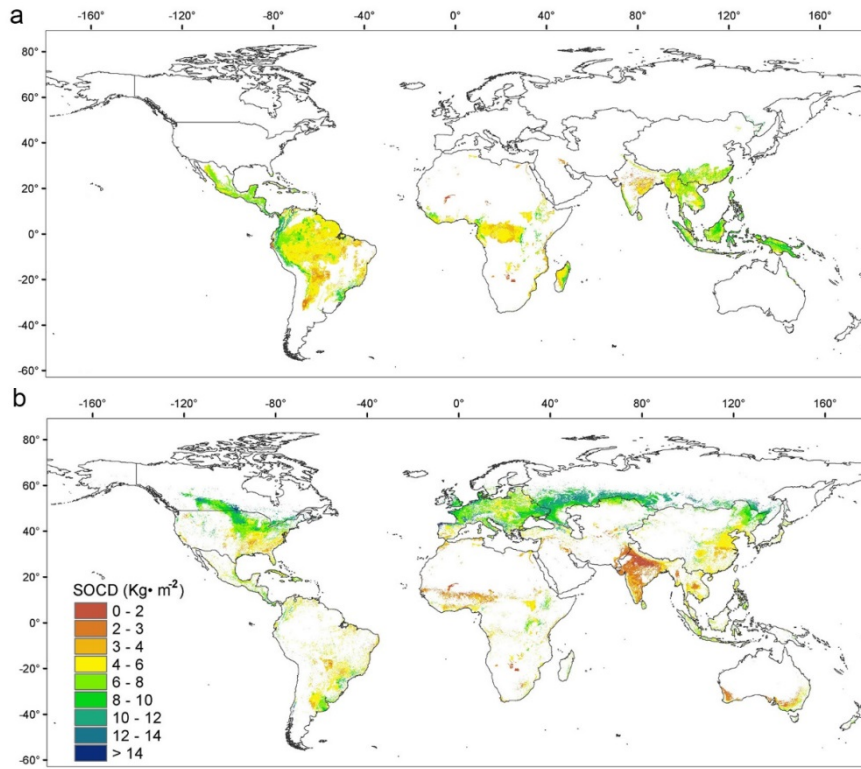
58 **Figure S6.** Changes in surface soil properties with elevation in tropical regions. a:  
59 Bulk density ( $\text{g}\cdot\text{m}^{-3}$ ); b: Sand (%); c: Silt (%); d: Clay (%); e: pH; f: SOCD ( $\text{kg C}\cdot\text{m}^{-2}$ );  
60 g: STND ( $\text{kg N}\cdot\text{m}^{-2}$ ); h: C:N ratio.  
61



62  
63



64 **Figure S7.** Spatial variations of SOCD in (a) tropical forests, and (b) croplands.



65

66



67 **Table S1.** Data sources for the Global Soil Database (GSD).

68

Data source	Country/region	Number of profiles used in GSD	Reference
ISRIC-WISE3.2	World	10184	Batjes (2009)
Canada Soil Reference Profiles	Canada	706	Pan et al. (2011)
Land Resources of Russia	Russia	254	IIASA, <a href="http://nsidc.org/data/ggd601.html">http://nsidc.org/data/ggd601.html</a>
International Soil Carbon Network	USA	13191	ISCN(2012) ( <a href="http://www.fluxdata.org/nscn/Data/AccessData/SitePages/Carbonto1M.aspx">http://www.fluxdata.org/nscn/Data/AccessData/SitePages/Carbonto1M.aspx</a> )
Soil Profile Analytical Database of Europe	Europe	495	SPADE
Northern Circumpolar Soil Carbon Database	Northern Circumpolar	519	NCSCD, Tarnocai et al. (2009)
Second State Soil Survey of China	China	2157	National Soil Survey Office (1998)
Literature retrieval on China's forests and the field campaign data in northern China (from our team)	China	668	Yang et al. (2008, 2010, 2014)
Field survey data in Australia	Australia	48	Wynn et al. (2006)

69

70 **Table S2.** Sample numbers for each region in the GSD.

71

Region	Study area (10 <sup>6</sup> ha)	Number of data				
		Bulk density (g·m <sup>-3</sup> )	Sand, silt and clay fractions (%)	pH	SOCD (kg·m <sup>-2</sup> )	STND (kg·m <sup>-2</sup> )
Tropical Asia	819	860	885	879	855	528
Mexico	264	316	315	312	313	261
Africa	2976	3740	3938	3870	3706	2305
Continental US	765	9322	13119	12005	11496	5326
Canada	852	790	903	223	889	596
Russia	1636	386	589	372	607	350
South America	1751	1764	1831	1737	1763	1427
Europe	558	1527	1588	1250	1509	790
East Asia	1136	2762	3120	2115	3063	2191
Australia	799	162	212	214	162	137
West Asia	1140	333	585	582	297	185
Alaska	133	79	497	295	428	228
Total	12829	22041	27582	23854	25088	14324

72 Note: SOCD stands for soil organic carbon density; STND stands for soil total nitrogen density.

73 **Table S3.** Empirical models for relationship between bulk density (BD, g m<sup>-3</sup>) and soil organic carbon content (SOC, %) for each region.

Region	Model	R <sup>2</sup>	Num
Tropical Asia	$BD = 1.336e^{-0.054 \text{ SOC}}$	0.26	765
Mexico	$BD = 1.380e^{-0.061 \text{ SOC}}$	0.63	1243
Africa	$BD = 1.480e^{-0.073 \text{ SOC}}$	0.30	3770
Continental US	$BD = -0.173\ln(\text{SOC}) + 1.382$	0.45	1239
Canada	$BD = 1.507e^{-0.027 \text{ SOC}}$	0.20	163
Russia	$BD = -0.222\ln(\text{SOC}) + 1.287$	0.59	777
South America	$BD = -0.07\ln(\text{SOC}) + 1.233$	0.15	2105
Europe	$BD = 1.4661e^{-0.041 \text{ SOC}}$	0.60	2391
East Asia	$BD = 1.4719e^{-0.08 \text{ SOC}}$	0.35	634
Australia	$BD = 1.3319e^{-0.062 \text{ SOC}}$	0.74	167

74

75

76 **Table S4.** Proxies used in global mapping of surface soil properties.

77

Variables	Description	Source
MAT	Mean annual temperature (°C)	WorldClim database, 1*1 km, 1950-2000 ( <a href="http://www.worldclim.org">http://www.worldclim.org</a> )
MAP	Mean value of total annual precipitation (mm)	WorldClim database, 1*1 km, 1950-2000 ( <a href="http://www.worldclim.org">http://www.worldclim.org</a> )
TS	The seasonality of temperature	Calculated from the monthly climate data*
PS	The seasonality of precipitation	Calculated from the monthly climate data*
NDVI	Mean annual NDVI	GIMMS NDVI3g dataset (8*8 km, 1982-2011), MODIS NDVI product (MOD13A2, 1*1 km, 2001-2005)
Elevation	Gobal DEM (m)	GTOPO30 (1*1 km, <a href="http://edcdaac.usgs.gov/gtopo30/README.asp">http://edcdaac.usgs.gov/gtopo30/README.asp</a> )
Slope	Slope	Calculated from DEM
Land use	Gobal land use map	The MODIS Global Land Cover Product (MCD12Q1, 500*500 m, 2006)

78 \* Seasonality =  $100 \times \frac{SD_{\text{monthly}}}{\text{Mean}_{\text{monthly}}}$  (Xu & Hutchinson, 2011)

79 **Table S5.** Coefficient of determination ( $R^2$ ) and root mean square error (RMSE) for the regional Random Forest models.  
 80

Region	$R^2$						RMSE					
	Bulk density	Content of sand	Content of clay	pH	SOCD	STND	Bulk density ( $\text{g}\cdot\text{m}^{-3}$ )	Content of sand (%)	Content of clay (%)	pH	SOCD ( $\text{kg}\cdot\text{m}^{-2}$ )	STND ( $\text{kg}\cdot\text{m}^{-2}$ )
Tropical Asia	0.56	0.27	0.28	0.67	0.37	0.33	0.12	20	15	0.73	2.35	0.21
Mexico	0.55	0.39	0.41	0.63	0.55	0.49	0.15	18	12	0.64	1.96	0.19
Africa	0.50	0.44	0.41	0.64	0.51	0.50	0.09	18	13	0.69	2.00	0.17
Continental US	0.27	0.50	0.44	0.62	0.50	0.46	0.14	16	9	0.66	2.76	0.23
Canada & Alaska	0.33	0.43	0.49	0.56	0.46	0.40	0.16	15	7	0.69	4.61	0.30
Russia	0.28	0.21	0.39	0.48	0.29	0.10	0.22	21	8	0.81	6.12	0.41
South America	0.23	0.24	0.16	0.57	0.32	0.24/	0.11	21	14	0.75	3.45	0.26
Europe	0.29	0.20	0.36	0.49	0.32	0.20	0.19	24	11	0.90	4.55	0.35
East Asia	0.51	0.18	0.39	0.51	0.47	0.28	0.14	16	8	0.79	2.50	0.21
Australia	0.46	0.65	0.47	0.28	0.47	0.31	0.17	18	12	0.81	2.41	0.18
West Asia	0.55	0.31	0.41	0.62	0.58	0.36	0.13	17	10	0.52	2.58	0.28

81 Notes: Values are the averaged  $R^2$  and RMSE from test dataset of 10-fold cross-validation.  
 82

83 **Table S6** Coefficient of determination ( $R^2$ ) of the regional Random Forest models by using all data samples.

84

Region	ALL DATA					
	Bulk density	Content of sand	Content of clay	pH	SOCD	STND
Tropical Asia	0.56	0.33	0.31	0.71	0.42	0.32
Mexico	0.52	0.41	0.45	0.67	0.52	0.48
Africa	0.44	0.49	0.43	0.64	0.58	0.43
Continental US	0.28	0.54	0.46	0.66	0.39	0.43
Canada & Alaska	0.37	0.48	0.53	0.59	0.41	0.39
Russia	0.24	0.25	0.44	0.47	<b>0.16</b>	<b>0.06</b>
South America	0.28	0.22	0.19	0.6	0.3	0.17
Europe	0.23	0.25	0.37	0.53	<b>0.15</b>	<b>0.08</b>
East Asia	0.54	0.25	0.33	0.54	0.4	0.18
Australia	0.47	0.67	0.5	0.31	0.51	0.32
West Asia	0.49	0.3	0.37	0.66	0.51	0.41

85 Notes: Values are the averaged  $R^2$  from test dataset of 10-fold cross-validation.