

1 Supporting Information for

2

3 **Global soil-climate-biome diagram: linking surface soil properties to climate and**
4 **biota**

5

6 Xia Zhao¹, Yuanhe Yang¹, Haihua Shen¹, Xiaoqing Geng¹, Jingyun Fang^{1,2*}

7

8 1 State Key Laboratory of Vegetation and Environmental Change, Institute of Botany,
9 Chinese Academy of Sciences, Beijing, China 100093

10 2 College of Urban and Environmental Sciences, and Key Laboratory for Earth
11 Surface Processes of the Ministry of Education, Peking University, Beijing, China
12 100871

13

14 ***Corresponding author:** Jingyun Fang, Tel/Fax: +86-10-6275 6560, E-mail:

15 jyfang@urban.pku.edu.cn

16

17

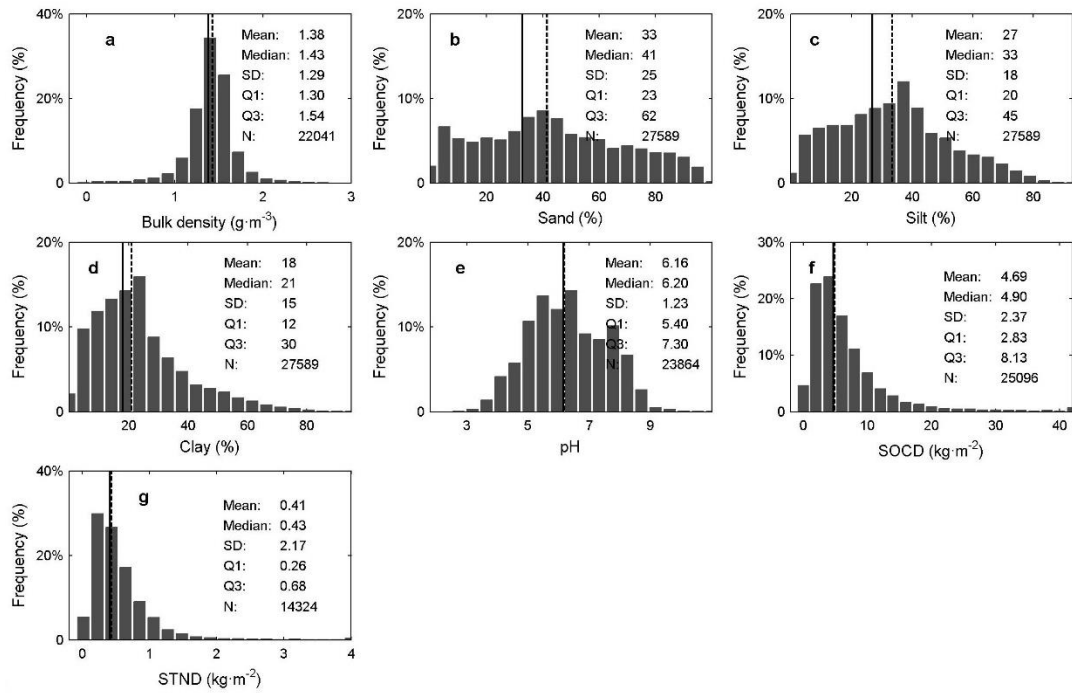
18 **Contents of this file**

19 Figures S1 to S4; Tables S1 to S4

20

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

27

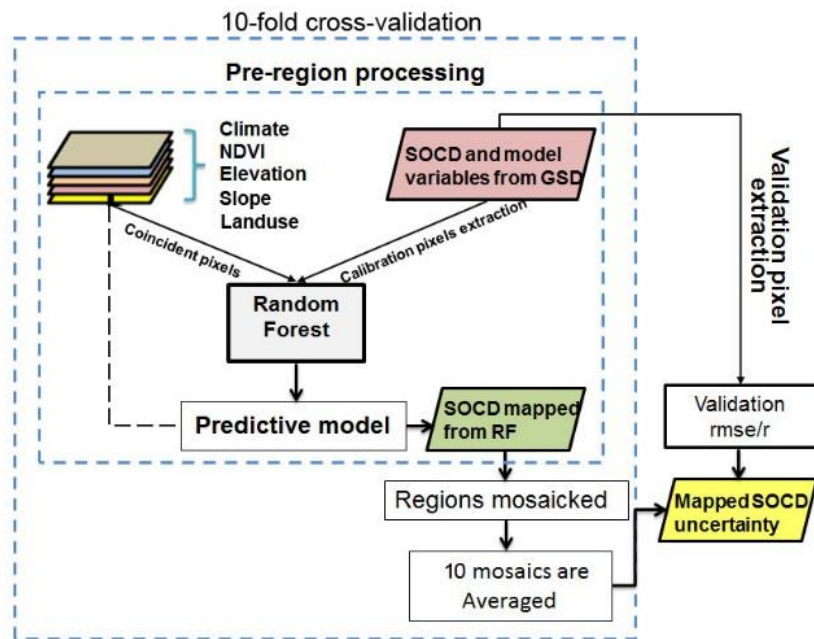


28

29

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

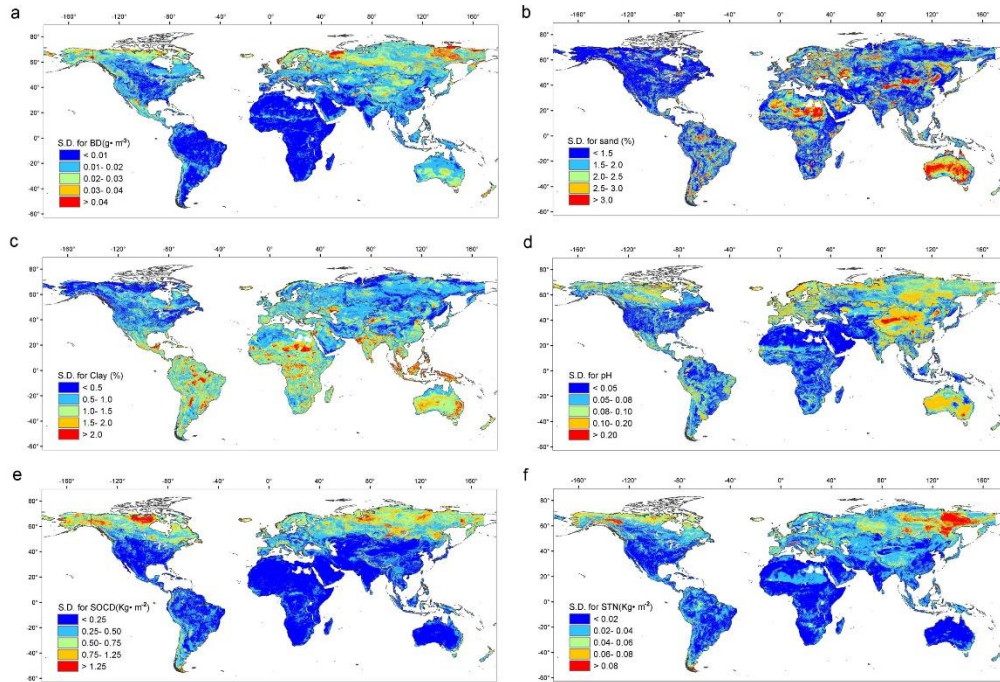
40



41

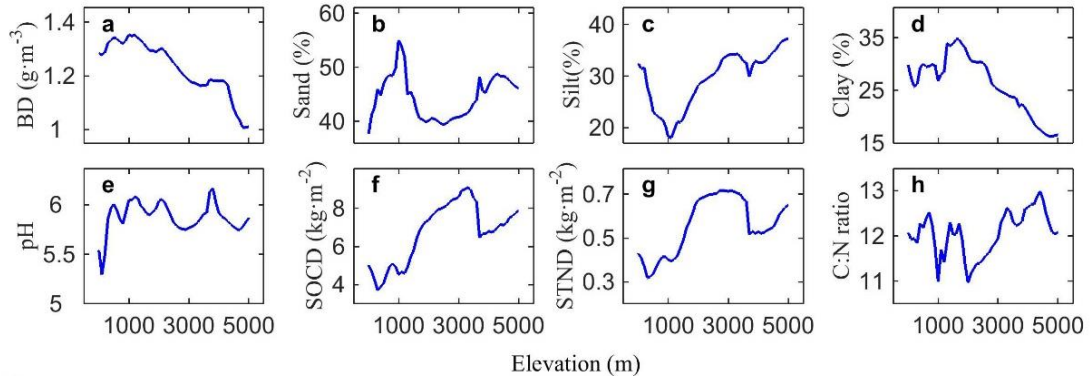
42 **Figure S3** Spatial pattern of uncertainties (standard deviation, SD, n=10) of soil
 43 properties in the upper 30-cm layer estimated by 10-fold cross-validation. a: Bulk
 44 density (g m^{-3}); b: Sand fraction (%); c: Clay fraction (%); d: pH; e: SOC density
 45 (SOCD, kg m^{-2}); f: STN density (STND, kg m^{-2}).

46



47

48 **Figure S4** Changes in surface soil properties with elevation in tropical regions. a:
49 Bulk density (g m^{-3}); b: Sand fraction (%); c: Silt fraction (%); d: Clay fraction (%); e:
50 pH; f: SOC density (SOCD, kg C m^{-2}); g: STN density (STND, kg N m^{-2}); h: C:N
51 ratio.
52



53

54 **Table S1** Data sources for the soil profiles constituting the database of the Global Soil Database (GSD).

55

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, http://nsidc.org/data/ggd601.html
International Soil Carbon Network	USA	13191	ISCN(2012) (http://www.fluxdata.org/nscn/Data/AccessData/SitePages/Carbonto1M.aspx)
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)

56

57 **Table S2** Number of valid soil attributes data for each region in the global soil database.

Region	Study area (10 ⁶ ha)	Number of data				
		Bulk density (g m ⁻³)	Sand, silt and clay fractions (%)	pH	SOCD (kg m ⁻²)	STND (kg m ⁻²)
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

58 Note: "-" indicate data are not available. SOCD stands for soil organic carbon density; STND stands for soil total nitrogen density.

59 **Table S3** Proxies used in spatial soil modelling.

60

Variables	Description	Source
MAT	Mean annual temperature (°C)	WorldClim database, 1*1 km, 1950-2000 (http://www.worldclim.org)
MAP	Mean value of total annual precipitation (mm)	WorldClim database, 1*1 km, 1950-2000 (http://www.worldclim.org)
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, http://edcdaac.usgs.gov/gtopo30/README.asp)
Slope	Slope	Calculated from DEM
Land use	Gobal land use map	The MODIS Global Land Cover Product (MCD12Q1, 500*500 m, 2006)

61 * Seasonality = $100 \times \frac{SD_{monthly}}{Mean_{monthly}}$ (Xu & Hutchinson, 2011)

62 **Table S4** Coefficient of determination (R^2) and root mean square error (RMSE) of the Random Forest models.

63

Region	R^2						RMSE					
	Bulk density	Content of sand	Content of clay	pH	SOCD	STND	Bulk density (g m^{-3})	Content of sand (%)	Content of clay (%)	pH	SOCD (kg m^{-2})	STND (kg 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	0.512	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

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