Supplement materials of

A data-driven topsoil δ13C dataset and the drivers of spatial variability across the Tibetan Plateau

Yunsen Lai [1], Shaoda Li [1, *], Yuehong Shi [1], Xinrui Luo [1], Liang Liu [1], Peng Yu [1], Guo Chen [2], Longxi Cao [2], Chunju Cai [3], Jian Sun [4], Shaohui Chen [5], Houyuan Lu [6, 7, 8], Xuanlong Ma [9], Xiaolu Tang [2, 10, *]

[1] College of Earth Science, Chengdu University of Technology, Chengdu 610059, Chengdu, China

[2] College of Ecology and Environment, Chengdu University of Technology, Chengdu 610059, Chengdu, China

[3] Key Laboratory of Bamboo and Rattan Science and Technology of National Forestry and Grassland Administration, International Centre for Bamboo and Rattan, Beijing 100102, China

[4] Key Laboratory of Alpine Ecology, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China

[5] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China

[6] Center for Excellence in Tibetan Plateau Earth Science, Chinese Academy of Sciences, 100101, Beijing, China

[7] University of Chinese Academy of Sciences, 100049, Beijing, China

[8] Key Laboratory of Cenozoic Geology and Environment, Institute of Geology and Geophysics, Chinese Academy of Sciences, 100029, Beijing, China

[9] College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, 730000, China

[10] State Environmental Protection Key Laboratory of Synergetic Control and Joint Remediation for Soil & Water Pollution, Chengdu University of Technology, Chengdu 610059, Chengdu, China

*Shaoda Li, corresponding author at College of Earth Science, Chengdu University of Technology, Chengdu, China. Email: 1661348392@qq.com.

*Xiaolu Tang, corresponding author at College of Ecology and Environment, Chengdu University of Technology, Chengdu, China. Email: lxtt2010@163.com.

33

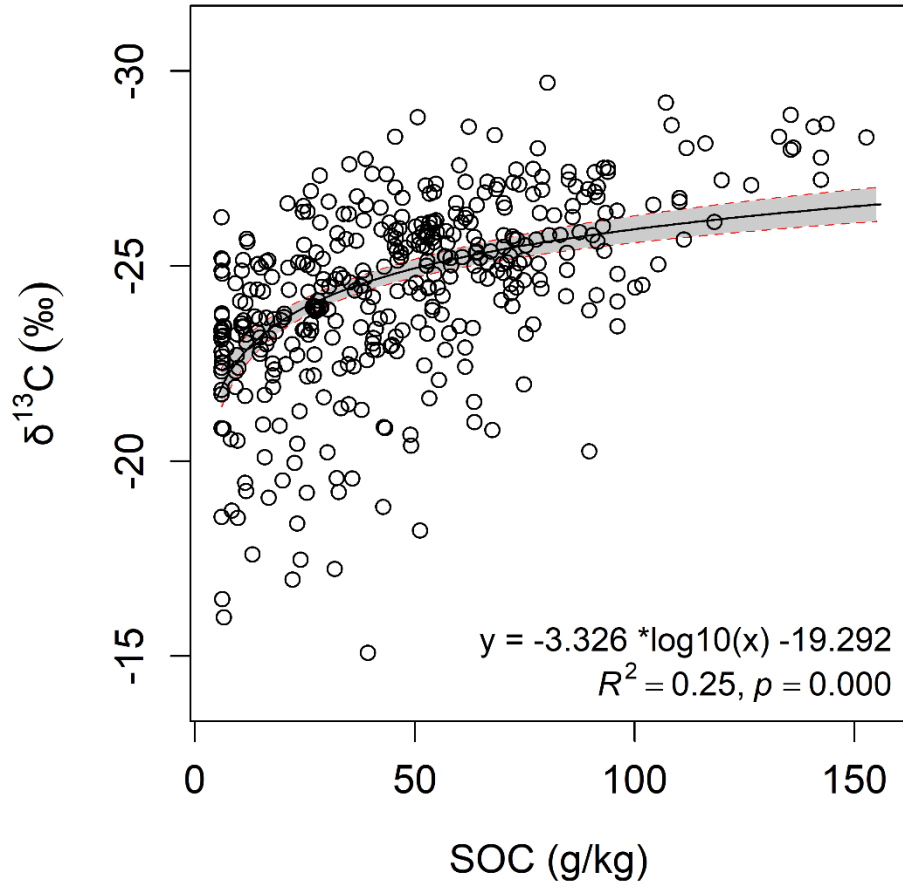



Figure S1. Relationships of soil δ¹³C with SOC content in 0-5 c

35

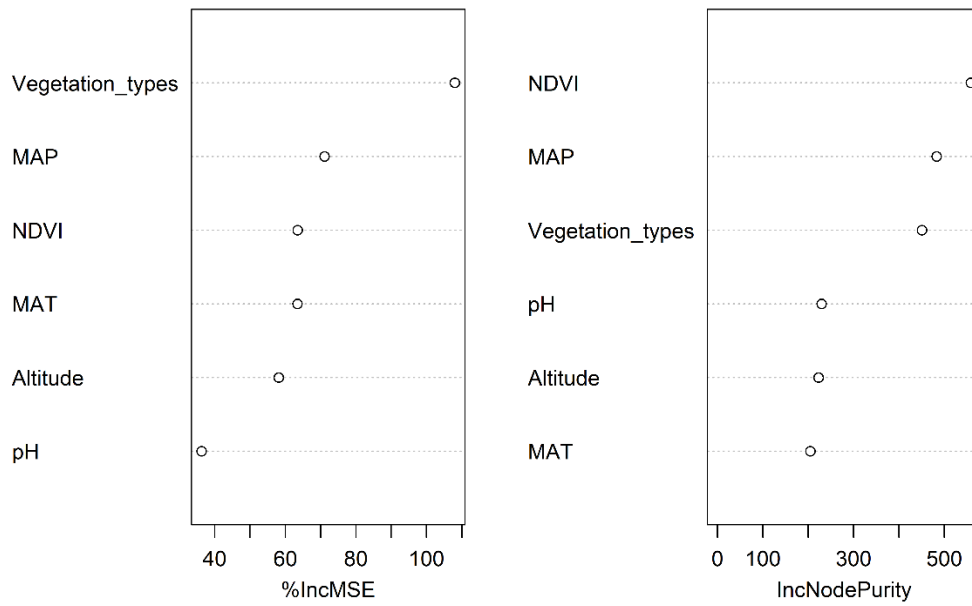



Figure S2. Importance of variables, denoted by percentage increase of mean-squared error

(%IncMSE) (left) and the increase in node purity (IncNodePurity) (right), for topsoil $\delta^{13}C$ estimation

RF model built from training dataset without considering plot location uncertainty. Vegetation_types,

MAP, NDVI, MAT, Altitude, pH, represent the ecosystem types, annual mean precipitation,

normalized difference vegetation index, elevation, soil pH, respectively.

41

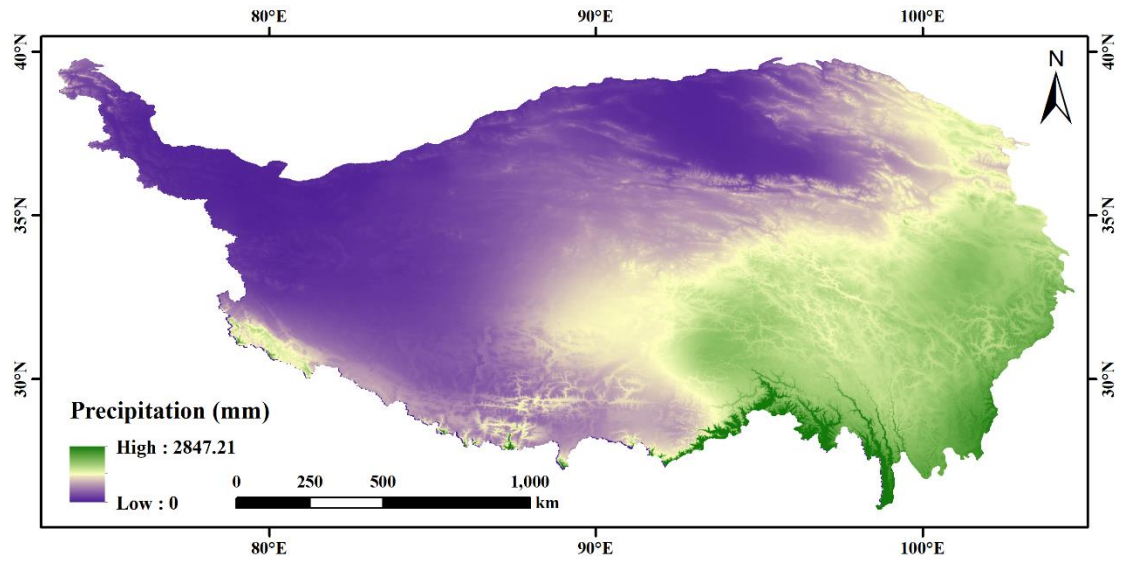
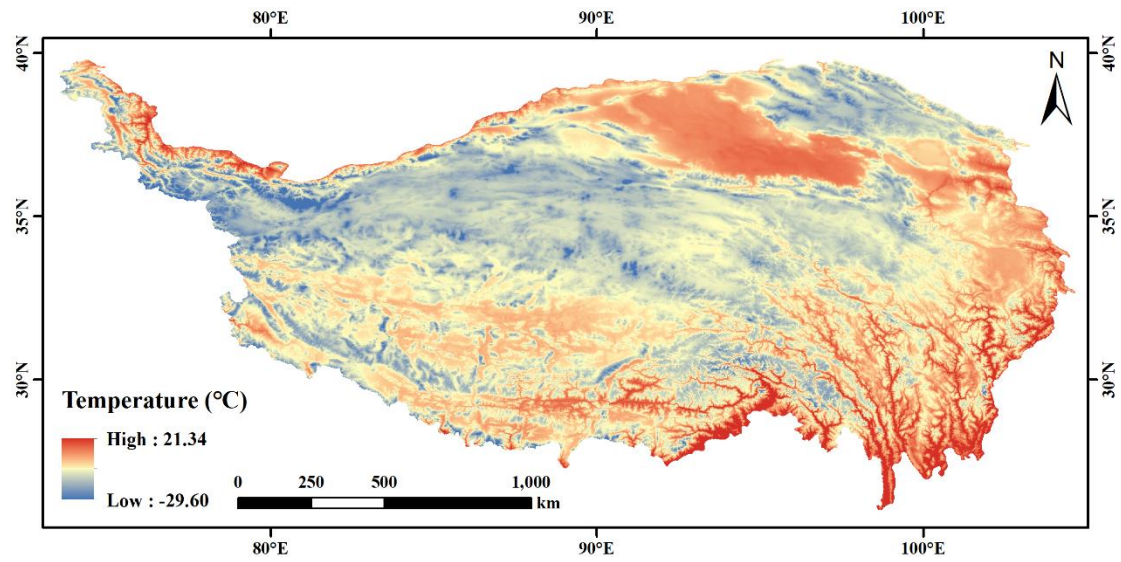
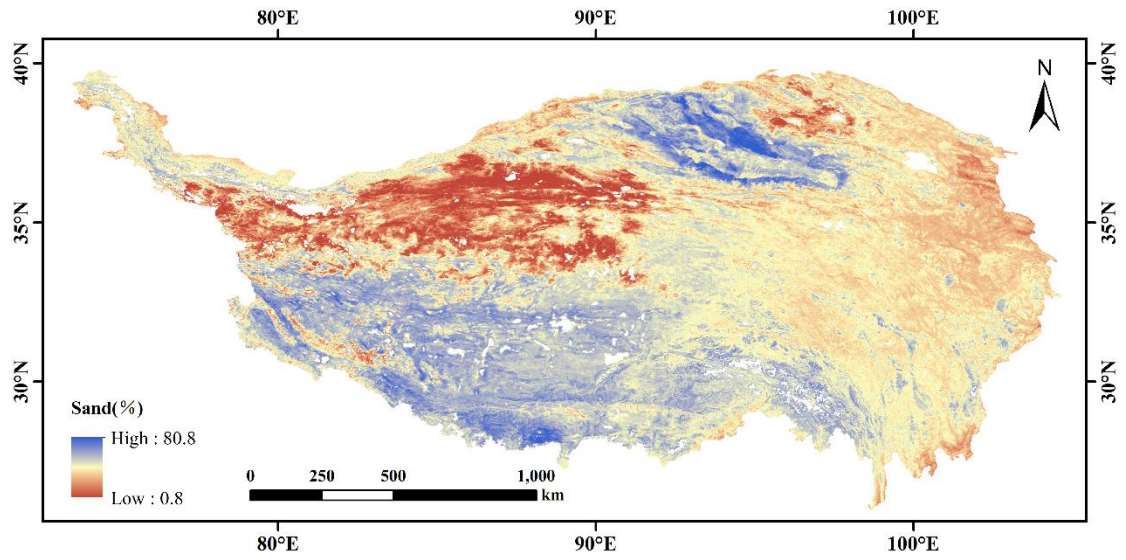
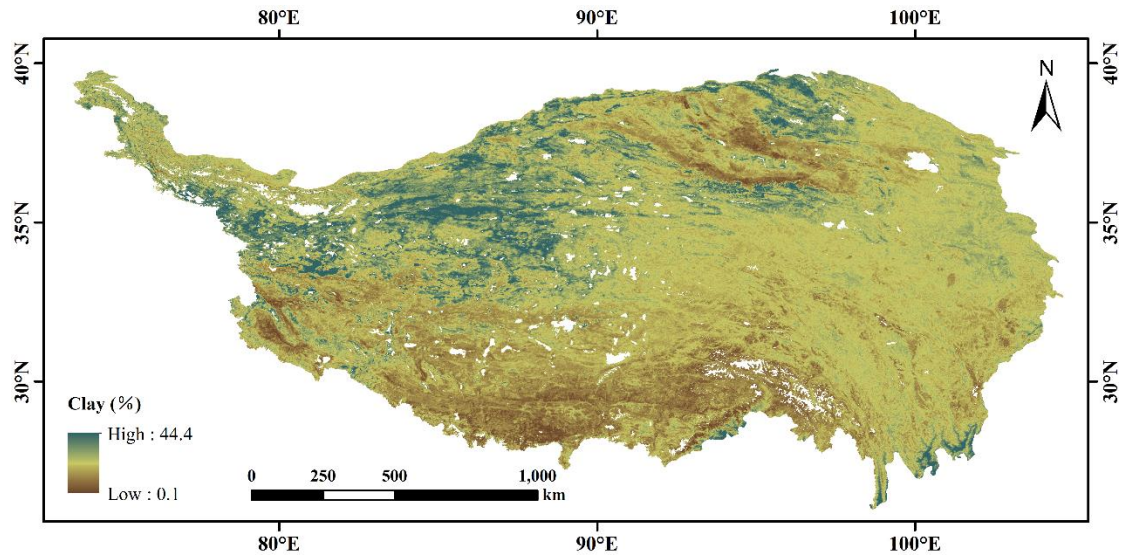



Figure S3. The distribution of precipitation during 2000-2010.

42

43

44   **Figure S4.** The distribution of temperature during 2000-2010.

45

46   **Figure S5.** The distribution of soil sand content on the Tibetan Plateau.

47

48   **Figure S6.** The distribution of soil clay content on the Tibetan Plateau.

        Table S1. General descriptions of the dataset used in this study.

Ecosystem types	N	Mean	Standard Deviation	Minimum	Maximum
Forests	103	-26.3	1.60	-29.7	-18.8
Shrublands	59	-24.3	2.00	-28.8	-18.2
Grasslands	218	-23.9	1.84	-27.6	-18.5
Deserts	16	-18.9	2.37	-22.5	-15.1

50   N are the number of soil samples.