Journal of Geophysical Sciences - Biogeosciences

Supporting Information for

Assessing soil carbon vulnerability in the Western USA by geo-spatial modeling of pyrogenic and particulate carbon stocks

Zia U. Ahmed[1,2], Peter B. Woodbury[2], Jonathan Sanderman[3, 4] , Bruce Hawke[3] , Verena Jauss [3] , Dawit Solomon[2], Johannes Lehmann[2, 5]

[1]International Maize and Wheat Improvement Center (CIMMYT), Dhaka, Bangladesh

Email:z.ahmed@cgiar.org

Tel: +88017466667416

[2]Crop and Soil Sciences Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14850, USA

[3]CSIRO Land and Water, Glen Osmond, Australia

[4]Woods Hole Research Center, Falmouth, MA 0240 USA

[5]Atkinson Center for a Sustainable Future, Cornell University, Ithaca, NY 14853, USA

Contents of this file

Figure S1. Histogram and spatial distributionsoil total organic carbon (SOC) and other organic soil C (calculated as SOC-POC-PyC).

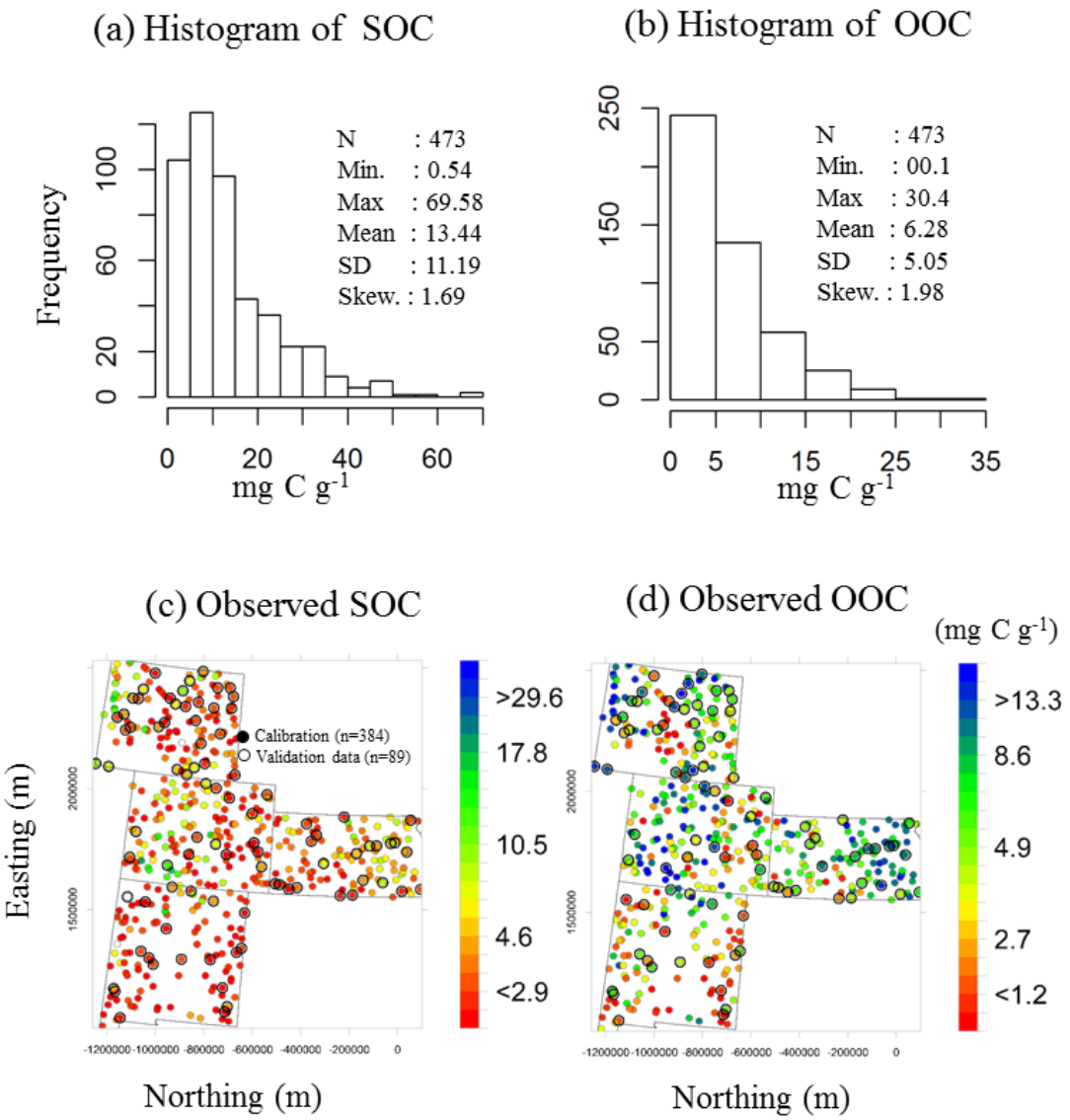

Figure S2. Importance plots for soil organic (SOC) (upper panel) and other organic C (OOC) (lower panel) derived from intial (left panel) and final (right panel) random forest regression models. Bar-plots in center panel show the percentage of the explained variance and Mean Square Error (MSE) of RF mdels with different numbers of predictors. (ELEV= elevation, G= slope gradient, SL= slope length, A = aspect, TPI = topographic position index, TSP= topographic slope position; C_v = vertical curvature, C_h = horizontal curvature, MAT = mean annual temperature, MAP= mean annual precipitation, NDVI= normalized vegetation index, NLCD = national land cover data, FRG = Fire regime group, K_f= K-factor)

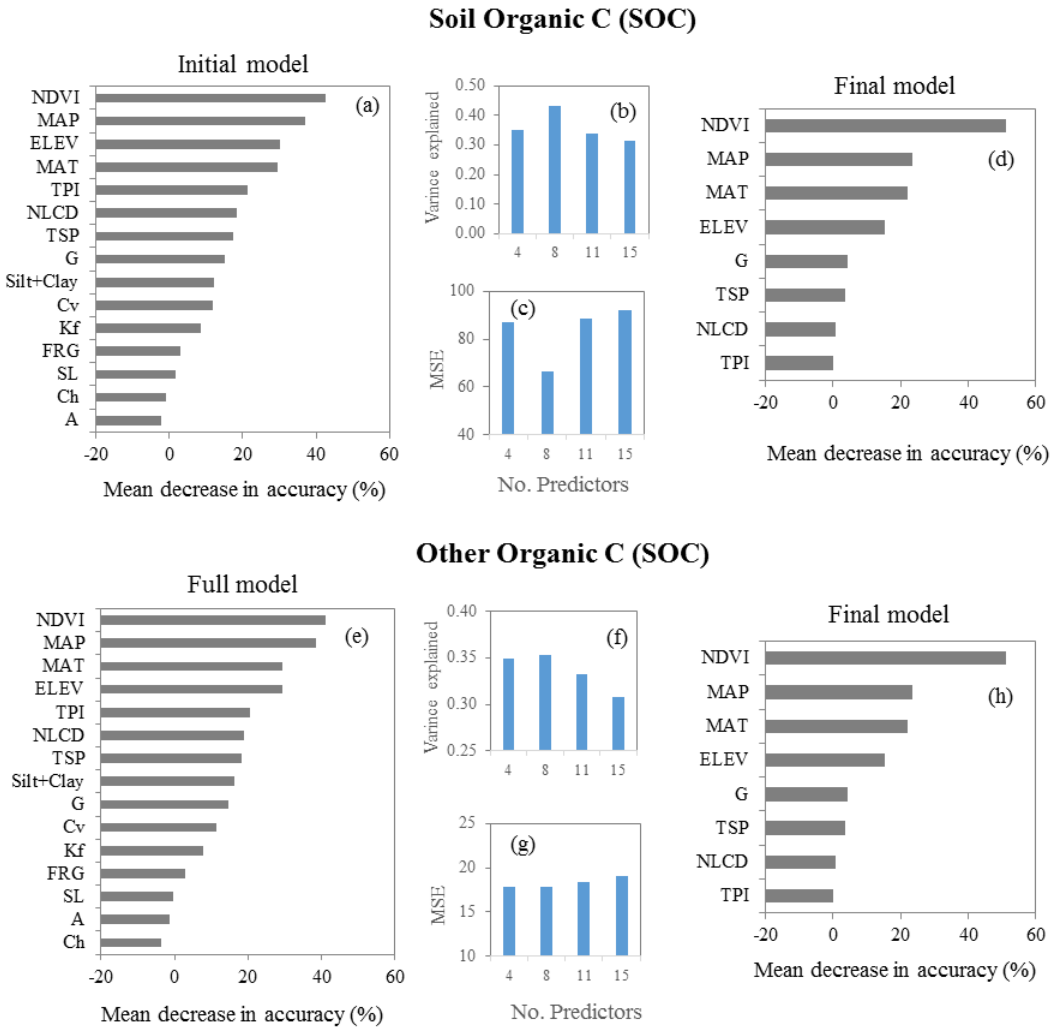

Figure S3. Variograms, observed vs predicted values, and ordinary (OK) and random forest regression kriging (RFK) predicted maps of soil organic (SOC) and other organic soil C (calculated as SOC-POC-PyC).

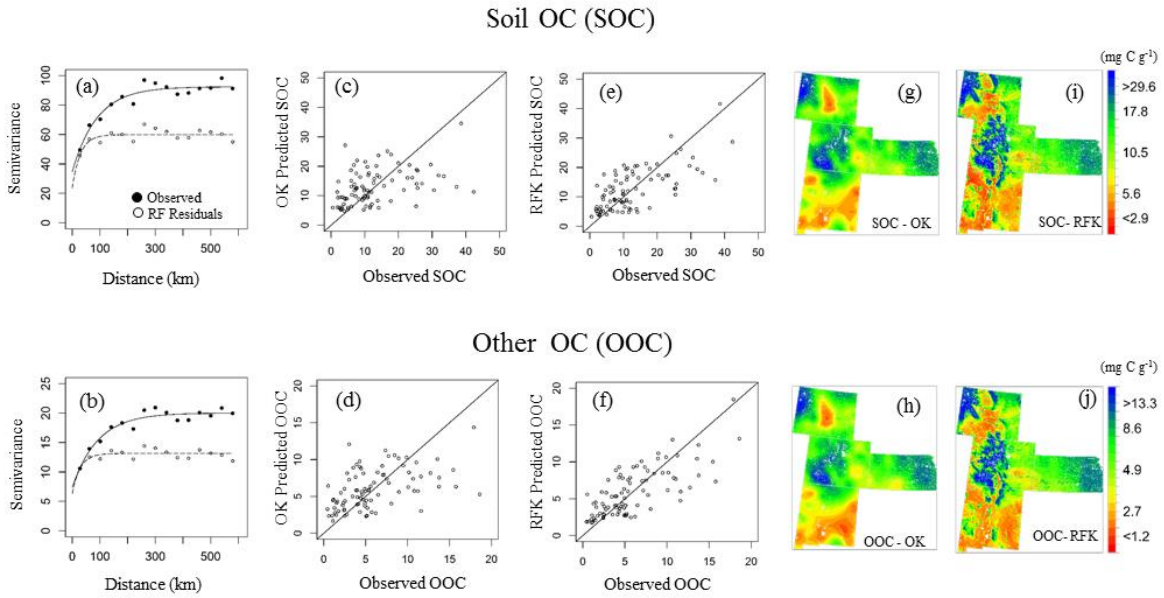

Table S1 Descriptive statistics of continuous environmental variables

Variables	Mean	Std.Dev	Max	Min	Median
Elevation (m)	1610.2	682.3	3346.0	249.0	1594.5
Slope gradient (^{o})	3.5	4.54	44.67	0.35	2.06
Slope length (m)	4.2	12.92	185.61	0.00	0.61
Slope aspect (^{o})	166.2	98.57	357.27	-1.00	161.57
Vertical curvature (m^{-1})	0.011	0.08	0.73	-0.31	0.00
Horizontal curvature (m^{-1})	0.002	0.06	0.38	-0.54	0.001
Topographic position index					
TPI_{250}	-0.232	3.42	14.44	-23.11	0.111
TPI_{500}	-0.843	7.85	29.28	-46.97	0.278
TPI_{1000}	-2.047	13.29	44.27	-90.66	-0.157
TPI_{1500}	-3.672	20.20	78.40	-128.20	-0.332
TPI_{2000}	-9.964	43.22	128.49	-285.7	-1.199
Temperature (°C)	9.32	4.04	16.83	-0.42	9.78
Precipitation (mm)	463.05	197.13	1128.91	180.56	407.69
NDVI	0.393	0.174	0.841	0.121	0.366
Silt+clay (%)	54.09	17.84	90.90	1.00	53.82
K-factor	0.26	0.10	0.55	0.05	0.29

Table S2 Pearson correlation coefficients (*r*) between environmental variables and soil total organic C (SOC) and other organic C (calculated as SOC-POC-PyC) for 473 soil samples.

Variables	***r*-value**	
	OOC	**SOC**
Elevation *(ELEV)*	0.164***	0.198***
Slope gradient (*G*)	0.298***	0.278***
Slope length (*SL*)	0.181***	0.209***
Slope aspect (*A*)	0.037^{NS}	0.038^{NS}
Topographic position index (*TPI*)		
TPI_{250}	-0.045^{NS}	-0.040^{NS}
TPI_{500}	-0.162***	-0.170***
TPI_{1000}	-0.187***	-0.192***
TPI_{1500}	-0.198***	-0.204***
TPI_{2000}	-0.202***	-0.228***
Vertical curvature (K_v)	0.053^{NS}	0.073^{NS}
Horizontal curvature (K_h)	0.034^{NS}	0.024^{NS}
Temperature (*MAT*)	-0.359***	-0.380***
Precipitation (*MAP)*	0.515**	0.481**
NDVI	0.612***	0.603***
K-factor (K_f)	-0.110*	-0.110*
Silt+Clay	0.225**	0.173**

NS = not significant; *, ** and *** indicate significance at $p < 0.05$, < 0.01 and < 0.001, respectively

Table S3. Properties of fitted variogram models and validation results

Variables	Variogram properties					Validation results		
	Model	*Co*	*C+Co*	*A*	A_e	MAE	RMSE	RI
Organic Carbon						Ordinary Kriging		
OOC	Exp	7.39	19.99	92.1	276.3	2.72	3.62	
SOC	Exp.	34.7	92.3	89.1	267.3	6.48	8.64	
		Residuals				Random Forest Kriging		
OOC	Exp.	6.31	13.14	28.6	85.8	2.14	2.77	23.7
SOC	Exp.	23.72	59.82	29.3	87.9	5.26	6.74	22.0

SOC= total organic carbon; OOC = other organic carbon (calculated as SOC- POC-PyC).
Exp. = Exponential; *Co*=nugget*; C+Co* = sill*; A*=range (km); A_e= effective range (km), range x 3 for exponential model; *MAE* = mean absolute error*; RMSE* = root mean square error*; RI* = relative improvement (%) of RMSE of random forest regression kriging over ordinary kriging.