Journal of Advances in Modeling Earth Systems

Supporting Information for

Improving soil carbon estimates by linking conceptual pools against measurable carbon fractions in the DAYCENT Model Version 4.5

Shree R.S. Dangal[1,*], Christopher Schwalm[1], Michel A. Cavigelli[2], Hero T. Gollany[3], Virginia L. Jin[4] & Jonathan Sanderman[1]

[1]Woodwell Climate Research Center, 149 Woods Hole Road, Falmouth, MA 02540, USA
[2]US Department of Agriculture - Agricultural Research Service, Sustainable Agricultural Systems Laboratory, Beltsville Agricultural Research Center, Beltsville, MD 20705, USA
[3]US Department of Agriculture - Agriculture Research Service, Columbia Plateau Conservation Research Center, Pendleton, OR 97810, USA
[4]US Department of Agriculture - Agricultural Research Service, Agroecosystem Management Research Unit, University of Nebraska-Lincoln, NE 68583, USA

Correspondence to: Shree R.S. Dangal (shree.dangal@unl.edu)

[*]*Current Address*: School of Natural Resources, University of Nebraska-Lincoln, NE 68583

Contents of the file:

Figures S1-S9

Tables S1-S2

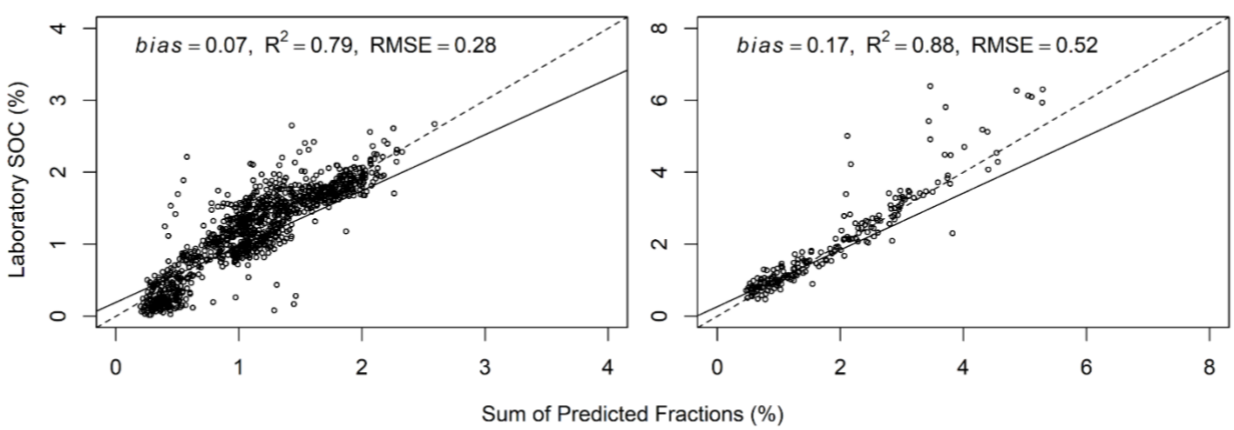

Fig S1. Comparison of machine learning based prediction of the sum of C fractions (POC, MAOC and PyC) against laboratory based total SOC for seven long term research sites in the continental US. The left panel figure represents croplands and the right panel figure represents grassland sites.

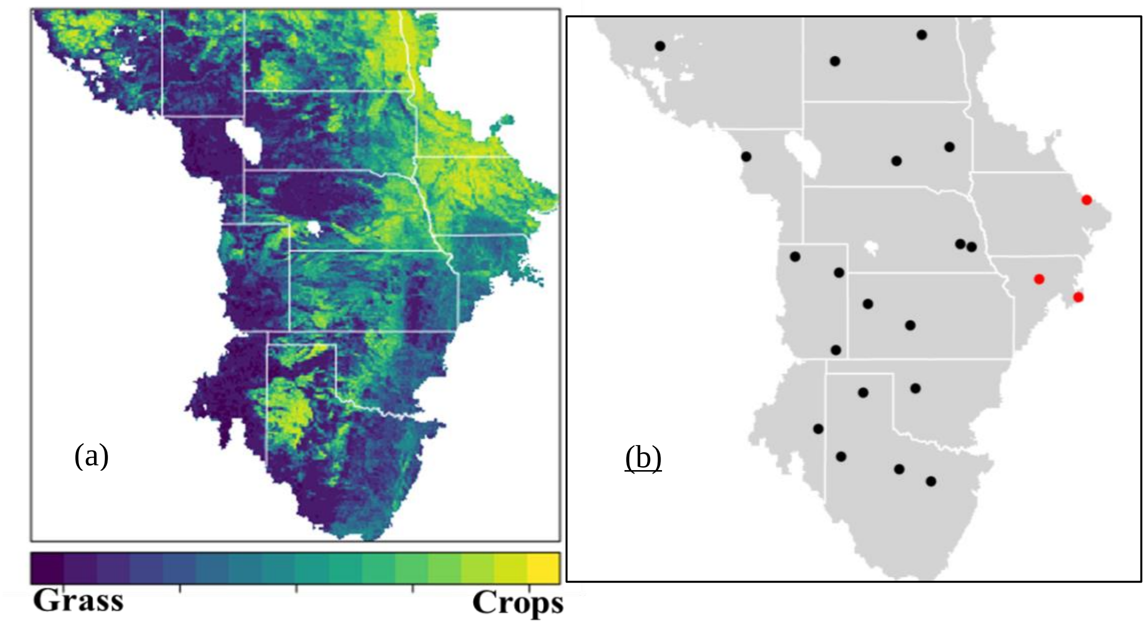

Fig S2. Cropland and grassland distribution (a) and distribution of the schedule files that represent different cropping systems (b) in the Great Plains region, US. The black dots in Fig. b represent 24 unique county level cropping systems and crop rotations, while the red dots represent new randomly selected grid points added to the clustering algorithm for building the unsupervised classification model.

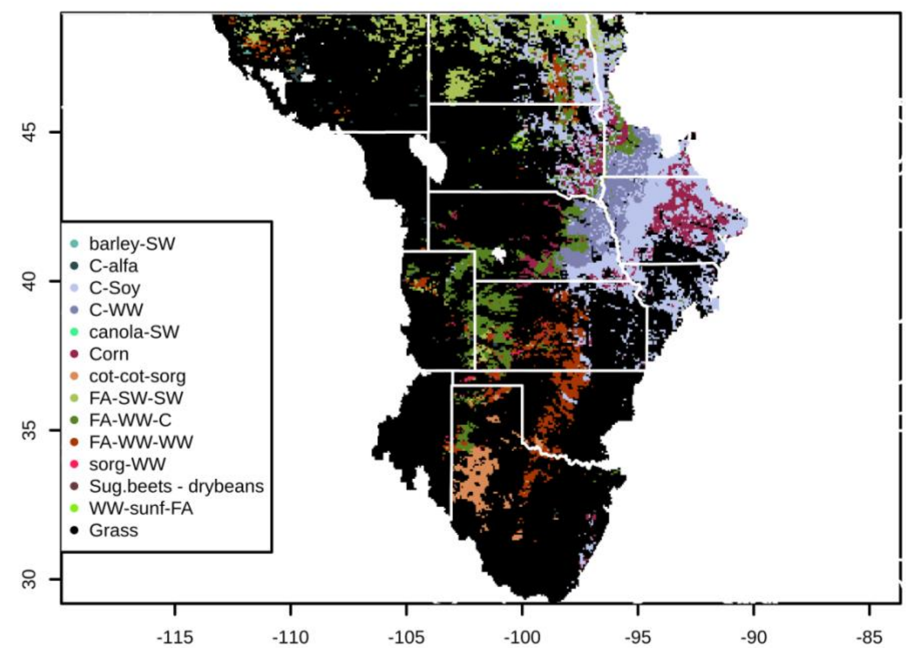

Fig S3. Crop rotation maps for the contemporary time period using the K-means unsupervised classification algorithm. The crop rotation map is used only when there is cropping in the given pixel. In the absence of cropping, the given pixel is assumed to be continuously grazed native grasslands.

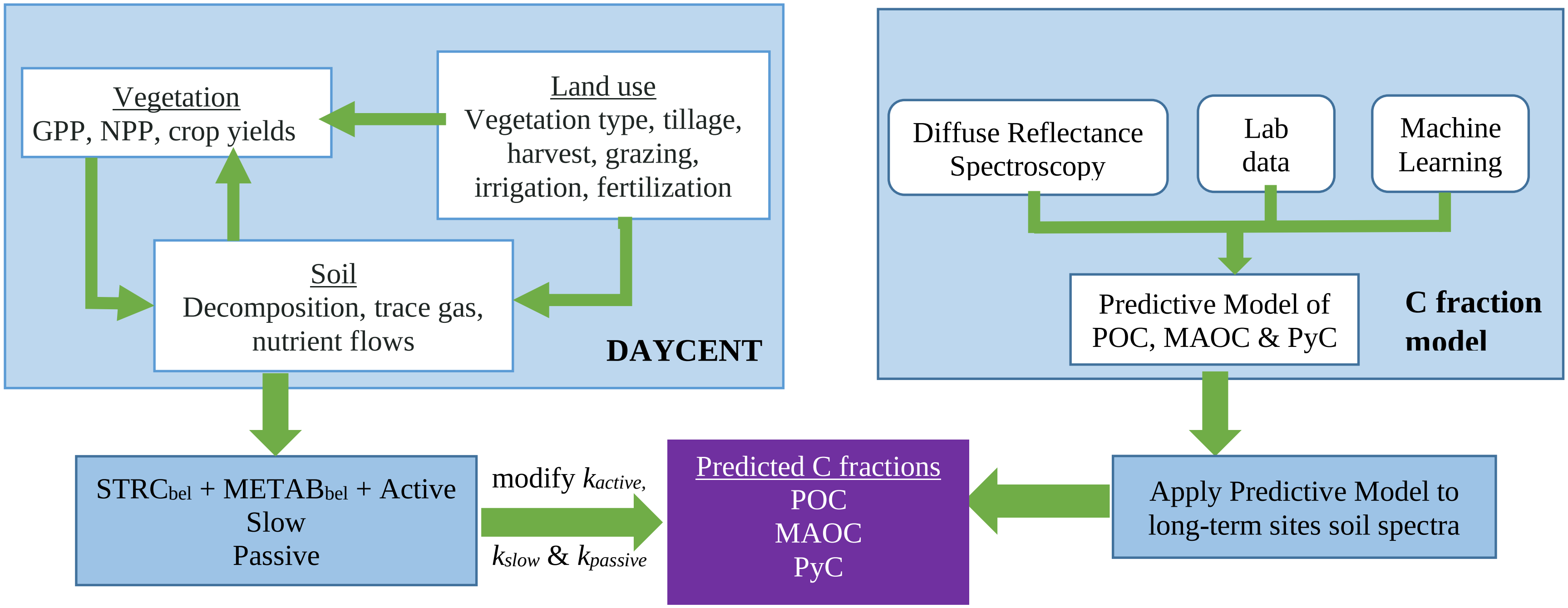

Fig. S4. Linking DAYCENT conceptual pools to C fraction data predicted using a combination of mid-infrared spectroscopy and a local memory-based learning approach, where $STRC_{bel}$ is structural, $METAB_{bel}$ is metabolic, Active, Slow and Passive are active, slow and passive soil C pools, and POC, MAOC and PyC are particulate, mineral associated and pyrogenic organic carbon.

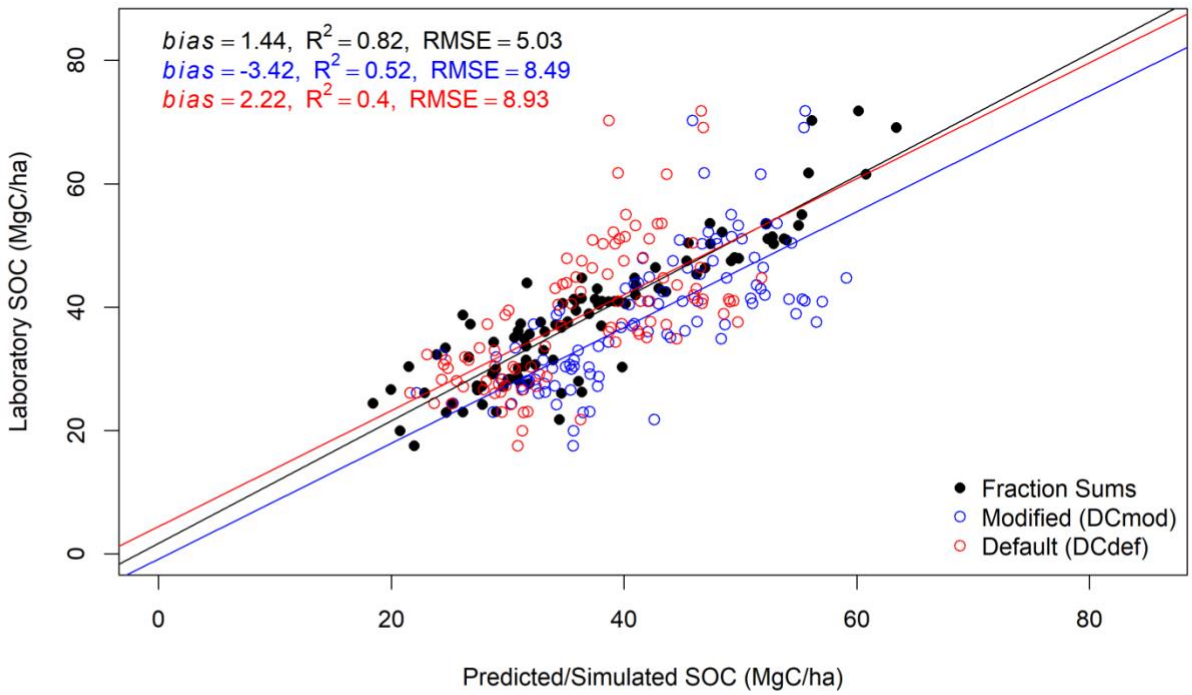

Fig. S5. Comparison of the sum of C fractions, DAYCENT simulated SOC using the default (DC_{def}) and the modified (DC_{mod}) models against laboratory based SOC estimates at the long-term research sites.

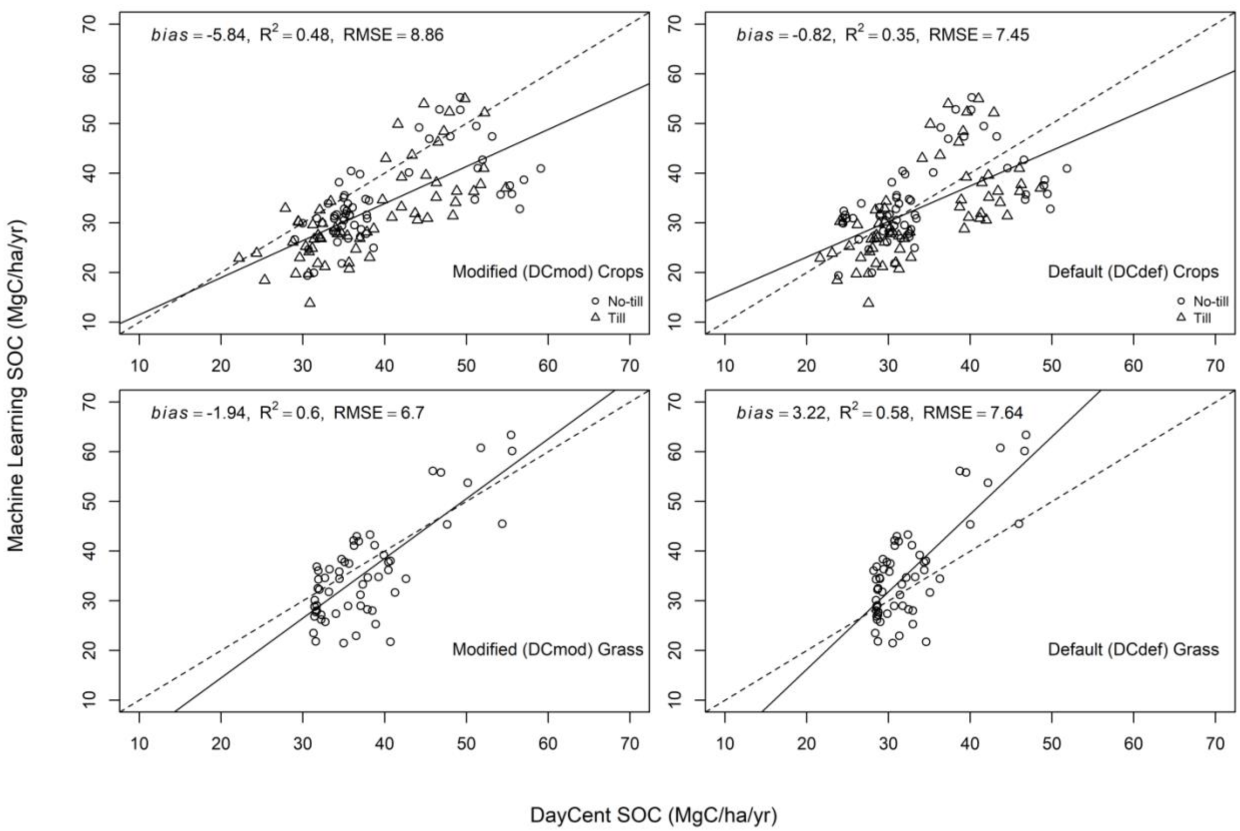

Fig S6 Scatterplots of the comparison of modified (DC_{mod}) and default (DC_{def}) simulation against data-driven estimates of total SOC at the long-term research sites. The top and bottom panels show the comparison for croplands and grasslands, respectively.

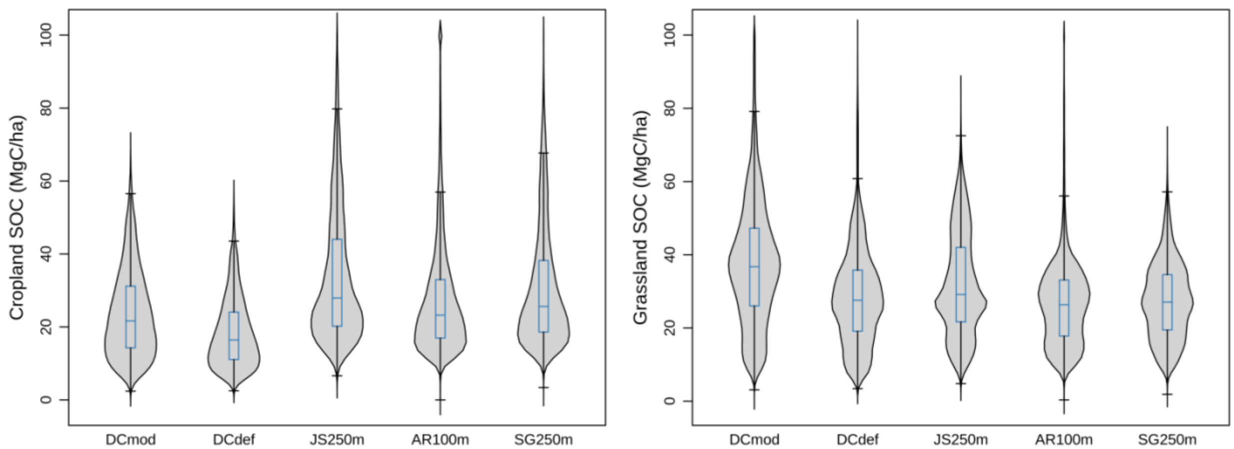

Fig S7. Comparison of total SOC (20 cm depth) between the DAYCENT and data driven modeling for the contemporary period. JS250, Sanderman et al. 2021; AR100m, Ramcharan et al. (2018); SG250m, Hengl et al. (2017).

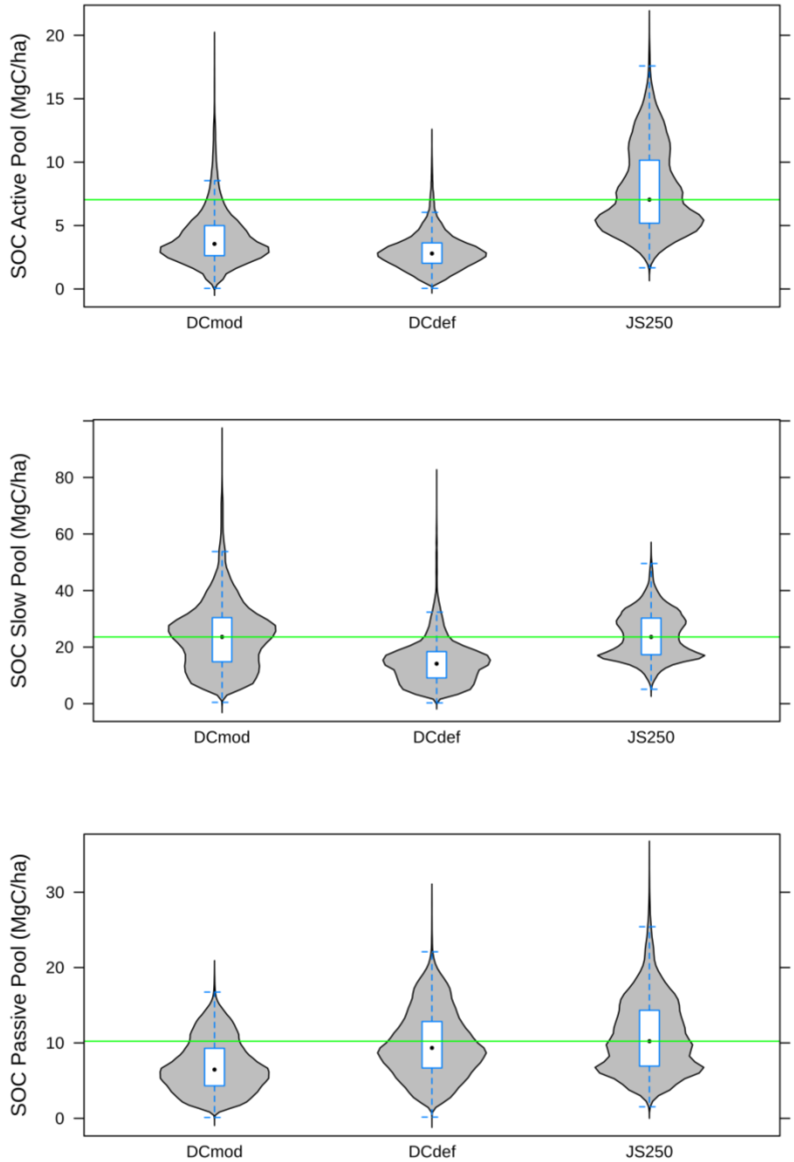

Fig S8. Comparison of the simulated active-, slow- and passive-SOC (20 cm depth) against Sanderman et al. (2020) for the US Great Plains Agricultural region during the contemporary period. The green line represents the median SOC values based on JS250 (Sanderman et al. 2021) C fraction predictions.

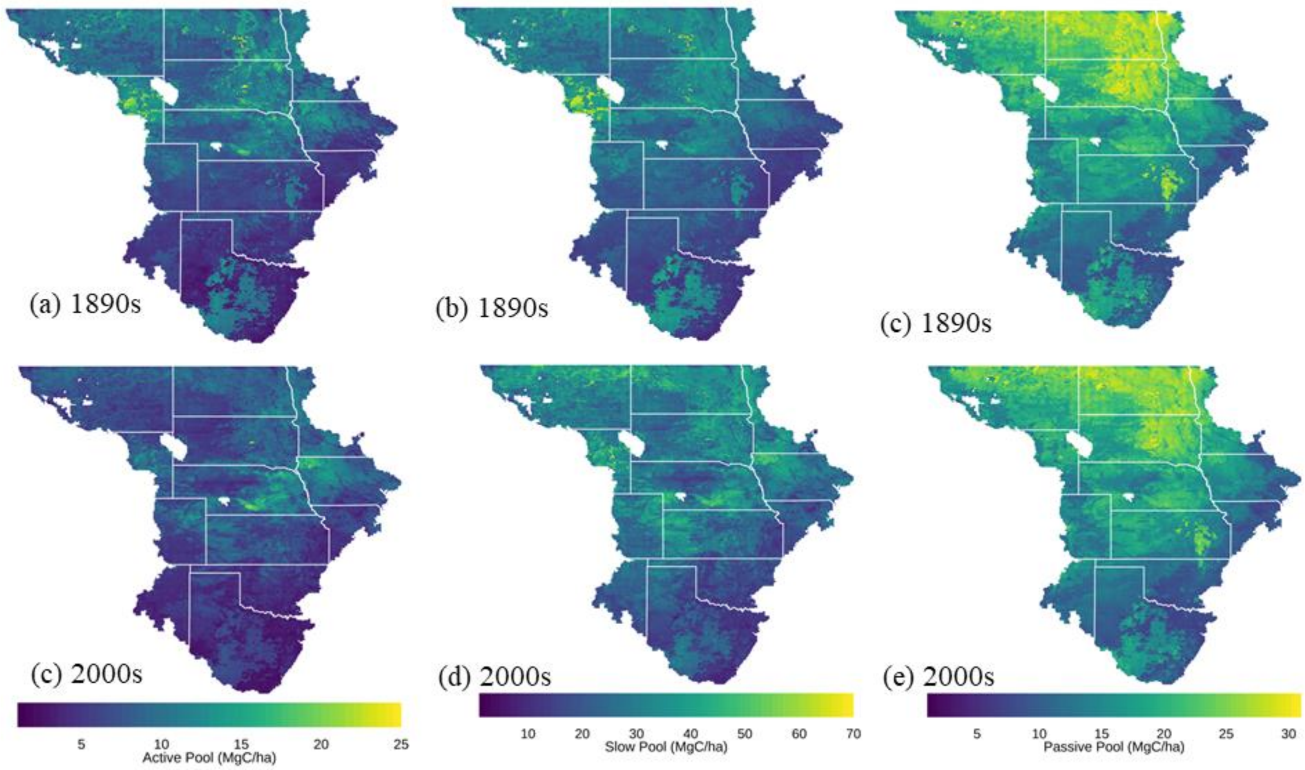

Fig S9. Active, slow and passive SOC pools at 20-cm depth based on the default (DC_{def}) model under native vegetation (1895-1899 average; top maps) and following land cover land use change (2001-2005 average; bottom maps).

Table S1. Predictive performance of US Samples using spectra acquired on Woodwell instrument with and without calibration transfer

	No calibration transfer[1]			After calibration transfer[1]		
	Bias	R^2	RMSE	Bias	R^2	RMSE
POC (g/kg)	0.65	0.50	4.93	1.04	0.70	4.39
MAOC (g/kg)	0.86	0.81	3.30	0.62	0.88	2.84
PyC (g/kg)	0.38	0.49	2.83	0.29	0.68	2.29

[1]Leave-one-out cross validation on the 99 GP samples

Table S2. Distribution of SOC across different pools by plant functional types (PFTs) when compared to C fractions predictions at the long-term research sites.

	Grasslands			Croplands		
	C fractions	DC_{mod}	DC_{def}	C fractions	DC_{mod}	DC_{def}
Active	0.20	0.13	0.08	0.14	0.14	0.08
Slow	0.56	0.63	0.49	0.57	0.56	0.39
Passive	0.24	0.24	0.43	0.29	0.30	0.53