Table 1: Modelling approaches

	Dataset 1:	Dataset 2: German Agricultural Soil Inventory +		
	German Agricultural Soll Inventory	LUCAS		
One-Model-Approach	AP1	AP1L		
Two-Model-Approach	AP2	AP2L		

Table	2:	Mean	of	error	metrics	of	the	three	models	for	each	approach.
Lable		1vi Cull	UL.	CI I UI	menico	OI.	unc	unice	moucis	101	cucii	appi oucin

Approach	Mean RMSE (g kg ⁻¹)	Mean MAE (g kg ⁻¹)	Mean MAPE (%)
AP1	32.6	12.3	49.0
AP1L	32.1	12.1	46.9
AP2	21.6	8.8	34.4
AP2L	21.3	8.7	34.3

Table S1: Predictive model performance of the models trained with different machine learning algorithms and datasets: A) built on the German Agricultural Soil Inventory, B) including LUCAS data in the training set. BRT = boosted regression trees, RF = random forest, and SVR = support vector regression.

	Algorithm	RMSE	MAE	%MAPE	%Bias	AIC	BIC	Approach
	BRT	32.9	12.4	50.9	-32	14865	14889	AP1
	RF	33.2	12.3	48.6	-30	14913	14919	AP1
	SVR	31.6	12.3	47.4	-20	14643	14661	AP1
	BRT	9.5	6.2	35.9	-20	7500	7524	Mineral
	RF	9.1	5.9	34	-20	7288	7294	Mineral
	SVR	9.2	5.8	31.8	-10	7331	7349	Mineral
А	BRT	107	90.4	48.5	-26	757	768	Organic
	RF	106.1	89.3	48.2	-28	750	753	Organic
	SVR	101.7	86.9	45.6	-22	746	754	Organic
	BRT	22	9.1	36.3	-20	12578	12602	AP2
	RF	21.7	8.8	34.5	-20	12496	12502	AP2
	SVR	21	8.6	32.3	-10	12310	12328	AP2
	Algorithm	RMSE	MAE	%MAPE	%Bias	AIC	BIC	Approach
	BRT	31.3	11.8	47.4	-30	14568	14592	AP1L
-	RF	32.5	12.1	46.8	-30	14754	14759	AP1L
	SVR	32.6	12.3	46.4	-20	14775	14792	AP1L
В	BRT	9.4	6.2	35.6	-20	7429	7453	Mineral
	RF	9.1	6	34.6	-20	7268	7274	Mineral
	SVR	9.1	5.8	31.7	-10	7275	7293	Mineral

BRT	105.4	88.4	45	-20	754	765	Organic
RF	104.1	86.2	43.5	-20	745	748	Organic
SVR	100.2	81.7	40.2	-12	741	749	Organic
BRT	21.7	9	36	-20	12486	12510	AP2L
RF	21.4	8.7	34.9	-20	12379	12385	AP2L
SVR	20.7	8.4	31.9	-10	12191	12209	AP2L

Table S2: Percent change in predictive model performance comparing models trained with different machine learning algorithms and data sets: A) and B) comparison of models trained by using data from the German Agricultural Soil Inventory, only. C) and D) comparison of models trained by using data from the German Agricultural Soil Inventory and LUCAS. A) and C) comparison with regards to the machine learning algorithms, B) and D) comparison of the one-model approach (AP1) to the two-model approach (AP2). E) comparison with regards to the machine learning algorithms trained by using data from the German Agricultural Soil Inventory and LUCAS. BRT = boosted regression trees, RF = random forest, and SVR = support vector regression.

	Algorithm	RMSE (%)	MAE (%)	MAPE (%)	Approach
	BRT to RF	0.9	-0.8	-4.5	AP1
	RF to SVR	-4.8	0.0	-2.5	AP1
	BRT to SVR	-4.0	-0.8	-6.9	AP1
	BRT to RF	-4.2	-4.8	-5.3	Mineral
	RF to SVR	1.1	-1.7	-6.5	Mineral
	BRT to SVR	-3.2	-6.5	-11.4	Mineral
А	BRT to RF	-0.8	-1.2	-0.6	Organic
	RF to SVR	-4.1	-2.7	-5.4	Organic
	BRT to SVR	-5.2	-4.0	-6.0	Organic
	BRT to RF	-1.4	-3.3	-5.0	AP2
	RF to SVR	-3.2	-2.3	-6.4	AP2
	BRT to SVR	-4.5	-5.5	-11.0	AP2
	Algorithm	RMSE	MAE	MAPE	Approach
	BRT	-33.1	-26.6	-28.7	AP1 to AP2
В	RF	-34.6	-28.5	-29.0	AP1 to AP2
	SVR	-33.5	-30.1	-31.9	AP1 to AP2
	Algorithm	RMSE	MAE	MAPE	Approach
	BRT to RF	3.8	2.5	-1.3	AP1L
	RF to SVR	0.3	1.7	-0.9	AP1L
	BRT to SVR	4.2	4.2	-2.1	AP1L
	BRT to RF	-3.2	-3.2	-2.8	Mineral
С	RF to SVR	0.0	-3.3	-8.4	Mineral
	BRT to SVR	-3.2	-6.5	-11.0	Mineral
	BRT to RF	-1.2	-2.5	-3.3	Organic
	RF to SVR	-3.7	-5.2	-7.6	Organic
	BRT to SVR	-5.2	-8.2	-10.7	Organic

	BRT to RF	-1.4	-3.3	-3.1	AP2L
	RF to SVR	-3.3	-3.4	-8.6	AP2L
	BRT to SVR	-4.6	-6.7	-11.4	AP2L
	Algorithm	RMSE	MAE	MAPE	Approach
	BRT	-30.7	-23.7	-24.1	AP1L to AP2L
D	RF	-34.2	-28.1	-25.4	AP1L to AP2L
	SVR	-36.5	-31.7	-31.3	AP1L to AP2L
	Algorithm	RMSE	MAE	MAPE	Approach
	BRT	-4.9	-4.8	-6.9	AP1 to AP1L
	RF	-2.1	-1.6	-3.7	AP1 to AP1L
	SVR	3.2	0.0	-2.1	AP1 to AP1L
	BRT	-1.1	0.0	-0.8	Mineral
	RF	0.0	1.7	1.8	Mineral
Б	SVR	-1.1	0.0	-0.3	Mineral
E	BRT	-1.5	-2.2	-7.2	Organic
	RF	-1.9	-3.5	-9.8	Organic
	SVR	-1.5	-6.0	-11.8	Organic
	BRT	-1.4	-1.1	-0.8	AP2 to AP2L
	RF	-1.4	-1.1	1.2	AP2 to AP2L
	SVR	-1.4	-2.3	-1.2	AP2 to AP2L



Figure 1: Performance indicators of the three algorithms. One-model approach (without LUCAS data AP1 and with LUCAS data AP1L) versus the two-model approach (AP2 and AP2L) for A) RMSE ($g kg^{-1}$), B) MAE ($g kg^{-1}$) and C) MAPE (%). The whiskers of boxplots show 1.5 times the interquartile range. Please note that the y-axis is shortened for better visibility and does not display a zero. BRT = boosted regression trees, RF = random forest, and SVR = support vector regression.



Figure S2: Regression plot for SOC depth extrapolation in A) Mineral soils, B) Organic soils, C) Cropland, D) Grassland.



Figure S4: Spatial distribution of relative residuals from the models trained with the different machine learning algorithms. A) AP1 approach, B) AP1L approach, C) AP2 approach and D) AP2L approach. BRT = boosted regression trees, RF = random forest, and SVR = support vector regression.



Figure S5: Spatial prediction of SOC content (g kg-1) of German agricultural soils based on the two-model approach for the three algorithms (BRT AP2L, RF AP2L, SVR AP2L). BRT = boosted regression trees, RF = random forest, and SVR = support vector regression.