Geoderma 370 (2020) 114365

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Soil organic carbon in peninsular Spain: Influence of environmental factors and spatial distribution

R. Calvo de Anta[a,*], E. Luís[a], M. Febrero-Bande[b], J. Galiñanes[a], F. Macías[a], R. Ortíz[c], F. Casás[a]

[a] Dept. of Edaphology and Agrícultural Chemistry, University of Santiago de Compostela, Campus Vida 15782, Santiago de Compostela, Spain
[b] Dept. of Statistics, Mathematical Analysis and Optimization, Faculty of Mathematics, University of Santiago de Compostela, Campus Vida 15782, Santiago de Compostela, Spain
[c] Dept. of Agricultural Chemistry, Geology and Pedology, University of Murcia, Campus of Espinardo 30100 Murcia, Spain

A R T I C L E I N F O

Keywords:
Soil organic carbon
Digital soil organic carbon map
Random forests
Peninsular Spain

A B S T R A C T

Soil organic carbon (SOC) stocks and their geographical distribution in peninsular Spain were estimated from a georeferenced database consisting of 12,724 surface samples (0–30 cm) and 3607 subsurface samples (30–50 cm), covering different climate, land use, elevation, parent material, soil type and soil pH. SOC density showed a high heterogeneity, with the lowest values in arid regions, where the average in topsoil ranged between 20 and 60 t C ha^{-1}, under woody crops and forest respectively. Carbon stocks gradually increases as precipitation increases, and its variability is also dependent of other factors, fundamentally the presence/absence of active lime or active Al. In semi-arid zones, calcaric soils (pH $\approx$ 8.3) have higher contents of SOC than neutral to weakly acidic soils from siliceous materials. However, in humid regions, calcareous materials have undergo total or partial decarbonatation in the upper layer (pH $<$ 4.0–7.5) and SOC stocks are markedly lower than in other materials. In forest soils it seems that a steady state (around 100–120 t C ha^{-1}) (0–30 cm) has been reached in a wide range of precipitation, between 900 and 1700 mm; most of this carbon (about 80%) is labile-C. Soils from granitic rocks are acidic (pH 4.5–5.5) (Al buffering) and the mean SOC stock in the indicated precipitation range is between 170 and 200 t ha^{-1} (it is estimated that approximately 60% is stabilized as metal-C or mineral-C complexes). The highest values (190–240 t ha^{-1}) are recorded in acidic soils derived from mafic rocks, which in these regions usually develop *andic* properties (around 73% is involved in stable metal-C or mineral C complexes). Finally, the SOC stored in neutral soils from serpentinized ultramafic rocks (without excess Ca or Al) is similar to that of the decarbonated soils derived from calcareous materials. In all regions, forest soils are a much more important SOC sink than live forest biomass (2–4 times higher in the upper 30 cm and 3–6 times greater in the upper 50 cm).

Random Forest regression was used as modeling tool and digital mapping. Mean annual precipitation was estimated to be the most important predictor variable, followed by land use, lithology/soil type and soil pH. Model performance was calibrated by the internal RF validation and through cross-validation, and the results were similar. In topsoil, the mean error, root mean square error and R^2 were -0.007% C, 1.48% C and 0.61, respectively. In the subsurface layer these indices were -0.020, 1.07 and 0.37, respectively. SOC stock for peninsular Spain was estimated at 3.33 Pg in the upper (0–30 cm) layer, and 0.85 Pg in the subsurface (30–50 cm) layer. Total SOC stock for 0–50 cm was 4.19 Pg, with a 95% confidence interval ranging between 3.33 and 5.03 Pg.

1. Introduction

Despite important advances in knowledge about the role of soil as a carbon sink, the accurate determination of global stocks and their distribution through the generation of high resolution maps remains a challenge. The total soil organic carbon (SOC) stock, excluding litter and charcoal, has been estimated to be around 1500 Pg in the upper 100 cm of soil (700 Pg in the upper 30 cm), almost three times higher than in the vegetation and approximately double than in the atmosphere (about 560 and 760 Pg respectively) (Eswaran et al., 1993; Batjes, 1996; Jobbágy and Jackson, 2000; Scharlemann et al., 2009; Hiederer and Köchy, 2011) and is only surpassed by the ocean

---

* Corresponding author.
E-mail address: rcalvo.deanta@usc.es (R. Calvo de Anta).

https://doi.org/10.1016/j.geoderma.2020.114365
Received 6 November 2019; Received in revised form 25 March 2020; Accepted 30 March 2020

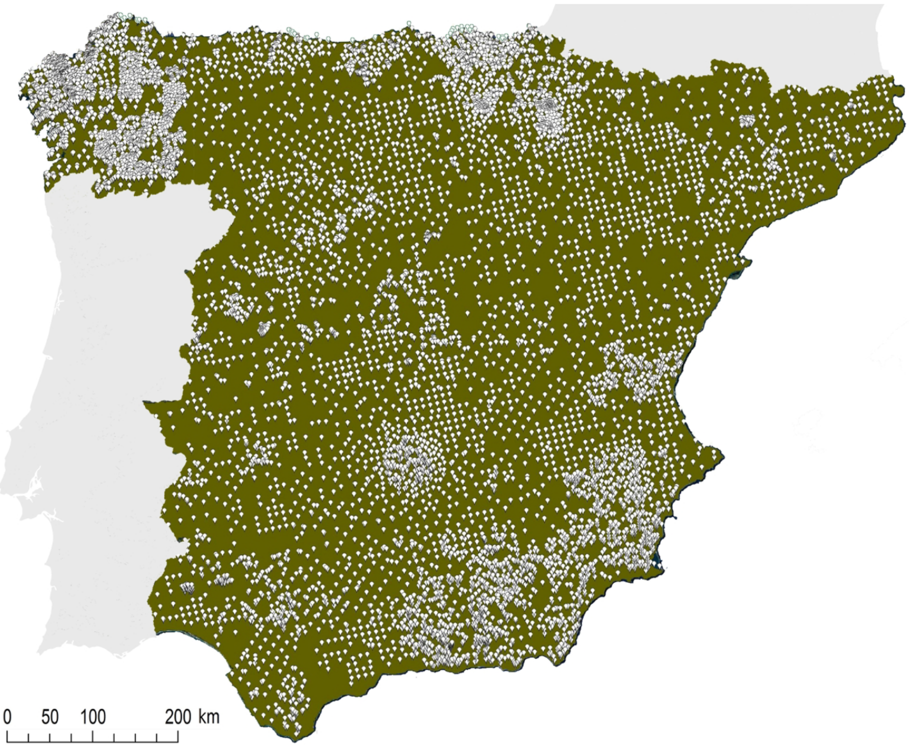

Fig. 1. Location of soil profiles that make up the database of the study.

Fig. 2. Principle systems making up the relief in pSpain (red lines: limits of the Autonomous Communities. 1- Galicia; 2- Asturias; 3- Cantabria; 4- País Vasco; 5- Navarra; 6- La Rioja; 7- Aragón; 8- Cataluña; 9- Castilla-León; 10- Madrid; 11- Extremadura; 12- Castilla-La Mancha; 13- Comunidad Valenciana; 14- Andalucía; 15- Murcia). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(~39000 Pg) (IPCC, 1990). In a recent study, Batjes (2016) estimated the global SOC stock up to 2 m depth in 2060 ± 215 Pg C. However, these global values are a rough calculation. Scharlemann et al. (2014) reviewed the estimates published in 27 studies from 1951 to 2014 and point out that, although the mean across all valuations is as indicated above (1460.5 Pg C, in the top 100 cm) there is a very considerable range of variation, between 504 and 3000 Pg. This significant uncertainty derives from the limitations or variability of the procedures used (disparity in the period, intensity and/or sampling method, poor spatial resolution of the soil databases, differences in approaches to calculations and lack of data for bulk density or gravel content, among others) (Stockmann et al., 2013; Scharlemann et al. 2014; Batjes, 2016). In conventional approaches, maps of SOC concentration and stock are based on coverage of single factors such as soil type, land cover,

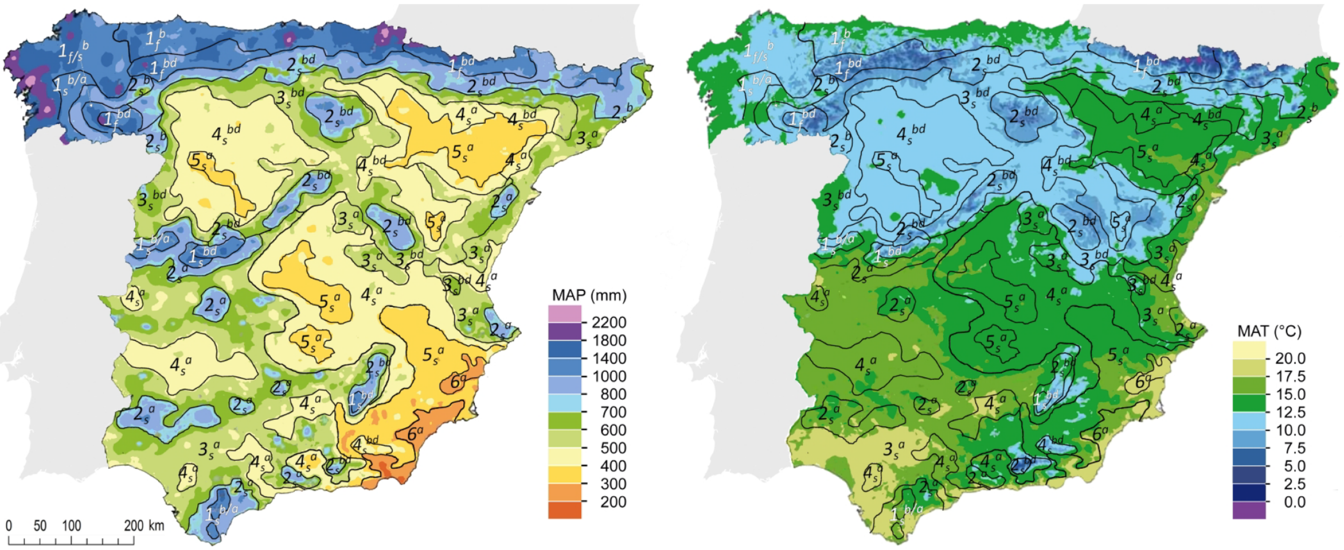

Fig. 3. Climatic regions of pSpain (superimposed on maps of mean annual precipitation, MAP, and mean annual temperature, MAT) (AEMET, 2011). Own elaboration from different digitized coverages provided by AEMET. Criteria used: *1*- Humid (MAP > 1000 mm); *2*- Subhumid (MAP: 700–1000 mm); *3*- Dry subhumid (MAP: 500–700 mm; *4*- Semi-arid (MAP: 400–500 mm); *5*- Arid (MAP: 300–400 mm); *6*- Desert (MAP < 300 mm); *f*- without dry season; *s*- dry summer (Mediterranean character); *b*- warm summer (mean T of the hottest month < 22.5 °C); *a*- hot summer (mean T of the hottest month > 22.5 °C); *d*- cold winter (mean T of the coldest month < 5 °C).

combinations of soil-land cover, or climate-soil for a particular type of land use. Lately, the accuracy of the estimates has improved markedly as a result of the development of digital soil mapping (DSM) technologies (McBratney et al., 2003) and by a wide range of statistical methods that are used as modeling tools of SOC stocks (Minasny et al., 2013; Todd-Brown et al., 2013).

The first estimate of SOC content (in topsoil 0–25 cm) at the European level was carried out by Van Ranst et al. (1995), who applied pedotransfer rules to a database elaborated in order to interpret the 1:1,000,000 EU Soil Map. Later, Jones et al. (2005) generated the first map of the SOC content (0–30 cm). They applied a refined pedotransfer rule to a data set derived from the European Soil Database. From this map the total SOC stock in Europe was estimated to be equal to 73 to 79 Pg (Schils et al., 2008). Nonetheless, the authors of these studies also recognize serious uncertainties due to the lack of comprehensive geo-referenced and harmonized data (Morvan et al., 2008). In 2009, the European Commission implemented the European harmonized geo-referenced topsoil database survey (LUCAS) (approximately 20,000 samples) and statistics were generated for the SOC concentration (0–20 cm) (Tóh et al., 2013**).** Application of a generalized additive model (GAM) and DSM techniques to the LUCAS database led to generation of the first digital map of SOC content (0–20 cm) in Europe (de Brogniez et al., 2015). Later, de Brogniez (2015) estimated the SOC stocks in the upper 20 cm of 23 EU countries at 38.3 Pg with a 95% confidence interval ranging between 34.9 and 42.4 Pg.

In order to improve continental and global estimates, regional and national scale studies become relevant. In recent decades there has been a notable increase in information in different regions of the world. Particularly in Europe, there are important contributions, e.g. in the Netherlands- in agricultural soils- (Kuikman et al., 2003), Denmark (Krogh et al., 2003), United Kingdom (Bradley et al., 2005), Ireland (Tomlinson, 2005), Austria (Gingrich et al., 2007), Switzerland (Bolliger et al., 2008), France (Arrouays et al., 2001; Martin et al., 2011; Mulder et al., 2016), Denmark (Adhikari et al., 2014), Scotland (Poggio and Gimona, 2014). The objectives of this study were to create a comprehensive soil database in peninsular Spain (pSpain) and to identify high-resolution environmental controllers to estimate SOC stocks at 0–30 and 0–50 cm depth and their spatial distribution, by applying Random Forest regression model.

2. Materials and methods

2.1. Database

The database built for this study consists of 12,724 georeferenced soil profiles, with data for 0–30 and 30–50 cm depth (3607 samples for the subsurface layer) (Fig. 1). The information included derives from three main sources: (1) 30% of the data were taken from published literature, either isolated articles or within the framework of inventory projects or soil mapping (ICONA, 1986–1992; lniguez, 1982-1992; Montoya and López-Arias, 1997). Most of these studies provide information on location, parent material, land use, soil classification, soil thickness, rockiness, stoniness, bulk density (BD) (not in all samples) and general physicochemical parameters in the fine earth fraction, such as organic-C and carbonate-C concentration, pH in H_2O and cation exchange capacity, among others. In these works organic-C refers to oxidizable carbon (Walkley-Black method), so the total organic-C content was estimated by applying an conversion factor (oxidizable C/total organic-C) obtained -from the other databases- for soils of the same area and land use. (2) A second sub-database (38% of the samples) was constructed from previous studies of the authors of the present manuscript, publications or reports of funded projects (Calvo de Anta, 2000–2004). In addition to the general information indicated above, data are available for oxidizable and total organic-C (by automated analyzer) and, for a selection of 361 forest soils on different lithology and soil type, there are data for selective extractions of C, Al and Fe by pyrophosphate solution (Bascomb, 1968), dithionite-citrate-bicarbonate (Mehra and Jackson, 1960) and ammonium oxalate (Blakemore et al., 1987), as well as phosphate retention (Blakemore et al., 1987) and pH in NaF (Fieldes and Perrott, 1966). (3) The third sub-database (32%) was specifically prepared to address the aims of this study. It consists of 3040 topsoil samples taken in the Cantabrian Cornice and Galicia in the period 2009–2013 (project funded by the MICINN) (Calvo de Anta et al., 2015); and 950 samples collected in the rest of pSpain (between 2014 and 2016) with the aim of covering areas where the density of sampling points was low, or to complete gaps in information, especially for BD and volume of coarse fragments. The air-dried soil samples were analyzed for: total organic-C; oxidizable-C (in 213 representative samples of different climate, lithology and land use

Table 1
General characteristics of the environment in the autonomous communities (ACs) in peninsular Spain: climate, lithology, soil type, soil pH, land use and dominant forest species (ACs ordered from NW to SE). Climate (*H:* Humid*; HO:* humid-oceanic; *SH:* subhumid; *DSH*: dry subhumid; *SA*: semi-arid; *A*: arid; *D*: desert; *M*: Mediterranean; *hs*: hot summer; *ws:* warm summer; *cw*: cold winter) (see criteria in Fig. 3); lithology (A_C: acid rocks; B_A: basic rocks; U_B: ultrabasic rocks; C_A: calcareous materials; Am: amphibolites; Cg: conglomerates; Dt: dolostones; G: granitic rocks; Gb: gabbros; Lt: limestones; Ma: marls; Mb: marble; Q: quartzites; Sch: schists; Sd: Tertiary & Quaternary sediments, calcareous or siliceous; Se: serpentinites; Sl: slates; St: sandstones); RSG qualifiers (*aa*: aluandic; *an*: andic; *ca*: calcaric; *cm*: cambic; *dy*: dystric; *eu*: eutric; *ha*: haplic; *li*: lithic; *le*: leptic; *pt*: petric; *rz*: rendzic; *vr*: vertic); land use (SH: Shrubs and scrubs; FOR: natural or plantation forests; GR: grasslands and pastures; CUL: arable crops; VOA: woody crops -vineyards, olive groves, almond trees and fruit trees-; R(es): rocky outcrops and eroded soils with scarce herbaceous coverage).

ACs (km^2)	Climate	Lithology (predominant)	Soil type RSG[a] (qualifiers)	pH$_{H2O}$[b] (0–30 cm)	Land use km^2		Forest area Major species[c]
					SH	4308	*P. pinaster:28%*
Galicia	*HO ws*	A_C (G, Sch, Sl, Q, St)	*Umbrisols (ha, le, an)*	4.5–5.5	FOR	13,712	*P.radiata & P.sylvestris: 9%*
(29,574)	*H ws-cw*	B_A (Gb, Am)	*Andosols (aa)*	4.5–5.5	GR	4497	*Q. robur: 14%*
		U_B (Se)	*Phaeozems (le, ha)*	5.5–6.5	CUL	3794	*E.globulus: 13%*
		Sd	*Cambisols (dy)*	4.5–5.5	VOA	609	*C.sativa* & others: 11%
			Fluvisols (dy)	4.5–5.5	R(es)	634	*pinus, eucaliptus* & others*: 25%*
		A_C (G, Sch, Sl, Q)	*Umbrisols (cm, le, ha),*	4.5–5.5	SH	2427	*C. sativa: 18%*
Asturias	*HO ws*	C_A (St, Cg, Lt, Dt)	*Leptosols (li, rz),*	5.0–7.7	FOR	4419	*F. sylvatica: 15%*
(10,604)	*H ws-cw*	Sd	*Cambisols (dy, eu, ca)*	5.0–7.7	GR	2820	*Quercus* spp.*: 16%*
			Phaeozems (rz)	6.5–8.3	CUL	175	*Eucalyptus* spp.*:13%;*
					VOA	212	*P.radiata, P.pinaster:13%*
					R(es)	227	Deciduous spp. mixtures*: 25%*
		C_A (St, Cg, Lt, Dt)	*Umbrisols (cm, le),*	4.5–5.5	SH	1484	*Quercus* spp: 33%
Cantabria	*HO ws*	A_C (Sl, Q)	*Phaeozems (rz),*	6.5–8.3	FOR	2137	*F. sylvatica*: 15%
(5253)	*H ws-cw*	Sd	*Cambisols (dy, eu, ca),*	5.0–7.7	GR	1140	*Eucalyptus* spp.: 19%
			Leptosols (li, rz)	5.0–8.3	CUL	65	*Deciduous* spp. *mixtures:18%*
					VOA	1	*P. sylvestris, P. radiata*: 9%
					R(es)	175	others: 6%
		C_A (St, Cg, Lt, Dt)	*Umbrisols (cm, le),*	4.5–5.5	SH	987	*P. radiata*: 35%
País Vasco	*HO ws*	A_C (Sl, Q)	*Phaeozems (rz),*	6.5–8.3	FOR	3973	*P.sylvestris, P.pinaster & others:16%*
(7228)	*H ws-cw*	Sd	*Cambisols (dy, eu, ca),*	5.0–7.7	GR	1364	*Quercus* spp.: 22%
			Leptosols (li, rz)		CUL	505	*F. sylvatica*: 15%
					VOA	153	*Eucalyptus* spp.: 3%
					R(es)	152	Mixtures of deciduous spp.: 9%
		C_A (St, Cg, Lt, Ma)	*Regosols (ca, eu, dy),*		SH	1385	
Navarra	*H ws-cw*	A_C (Sl, Sch, Q)	*Cambisols (ca, eu, dy),*	6.0–8.3	FOR	4474	*F. sylvatica*: 26%
(10,391)	*SA/A hs*	Sd	*Calcisols (cm, ha)*		GR	899	*Q. robur* & *F. sylvatica*: 16%
					CUL	2701	*Q. robur* & *Q. petraea*: 25%
					VOA	453	*P. sylvestris* & *P. nigra*: 33%
					R(es)	139	
		C_A (St, Cg, Lt)	*Regosols (dy, eu, ca),*		SH	1270	*F. sylvatica: 15%*
La Rioja	*SH Mws-cw*	A_C (Sl, Q)	*Cambisols (dy, eu, ca),*	6.0–8.3	FOR	1658	*Q. pyrenaica: 18%*
(5045)	*SA/A Mhs*	Sd	*Calcisols (cm, ha)*		GR	80	*Q. ilex & others: 23%*
					CUL	651	*P. sylvestris: 14%*
					VOA	580	*P. nigra: 4%*
					R(es)	107	*Pinus* mixture: 26%
			Regosols (ca, eu),		SH	10,719	
Aragón	*A/SA Mhs*	C_A (St, Cg, Lt, Ma)	*Cambisols (ca, eu),*	7.0–8.3	FOR	15,435	*P. sylvestris*: 35%
(47,720)	*H/SH Mws-cw*	Sd	*Calcisols (cm, ha)*		GR	2403	*P. halepensis*: 11%
			Phaeozem (rz)		CUL	11,115	*P. nigra, P. pinaster* (& others)*: 21%*
					VOA	2669	*Q.ilex, F.sylvatica, Q.faginea*: 33%
					R(es)	1751	
	H ws-cw	C_A (St, Cg, Lt, Ma)	*Regosols (dy, eu, ca),*		SH	3307	*P. halepensis*: 28%
Cataluña	*SH Mws*	A_C (G, Sl, Q)	*Cambisols (dy, eu, ca),*	5,5–8.3	FOR	16,071	*P. sylvestris*: 18%
(32,091)	*DSH Mhs*	B_A (Gb)	*Leptosols (li, rz),*		GR	2024	*P. nigra*: 11%
	A/SA Mhs	Sd	*Calcisols (ha)*		CUL	5080	*Pinus* mixture: 15%
			Umbrisols (le, ha)		VOA	2587	*Q. ilex*: 17%
					R(es)	689	*Fagus,Castanea* & *Quercus*:11%
	SA Mws-cw	A_C (G, Sl, Sch, Q, Cg)	*Regosols (dy, eu, ca),*		SH	18,705	*P. sylvestris*: 18%
Castilla-León	*DSH Mws-cw*	C_A (St, Cg, Lt)	*Cambisols (dy, eu, ca),*	5.0–7.5	FOR	29,447	*P. nigra*: 1%
(94,224)	*SH/H Mws-cw*	Sd	*Umbrisols (le, ha),*		GR	4167	*Pinus* mixture: 34%
			Regosols (ha, cm, vr,)		CUL	30,389	*Q. ilex*: 38%
			Arenosols (eu, ca)		VOA	2156	*F. sylvatica & o. deciduous: 8%*
					R(es)	1347	
Madrid	*SA Mhs*	A_C (G, Sch, Sl)	*Regosols (dy, eu, ca),*		SH	1028	*Q. ilex*: 36%
(8025)	*SH Mws-cw*	St, Cg)	*Cambisols (dy, eu, ca),*	5.5–8.3	FOR	2581	*Q. pyrenaica* & others: 24%
		Sd	*Umbrisols (le, ha)*		GR	1209	*P. nigra*: 12%
					CUL	1313	*P.pinea, P.pinaster* (& others)*: 15%*
					VOA	318	*Pinus* mixture: 12%
					R(es)	47	

(*continued on next page*)

Table 1 (*continued*)

ACs (km^2)	Climate	Lithology (predominant)	Soil type RSG[a] (qualifiers)	pH_{H2O}[b] (0–30 cm)	Land use km²		Forest area Major species[c]
Extremadura	*DSH Mhs*	A_C (Sl, Q, G)	*Regosols (dy, eu, ca),*		SH	7793	*Q. ilex*: 69%
(41,634)	*SA Mhs*	St, Cg)	*Cambisols (dy, eu, ca),*	5.0–8.3	FOR	18,917	*Q. suber*: 10%
	SH Mws-cw	Sd	*Calcisols (cm)*		GR	2466	*P. pinaster* & *P. pinea*: 10%
					CUL	4133	Mixtures: 11%
					VOA	3808	
					R(es)	269	
Castilla-	*SA Mhs*	C_A (St, Cg, Lt)	*Regosols (ca, eu, dy),*		SH	8895	*Q. ilex*: 31%
La Mancha	*A hs*	A_C (Sl, Q, G)	*Cambisols (ca, eu, dy),*	7.0–8.3	FOR	27,081	other deciduous: 20%
(79,470)	*SH Mhs*	Sd	*Calcisols (cm, pt, ha)*		GR	2351	*P. pinaster*: 17%
					CUL	21,399	*P. nigra*: 14%
					VOA	8184	*P. halepensis*: 11%
					R(es)	992	*P.sylvestris, P.pinea* (& others)*:7%*
Comunidad	*SH Mhs*	C_A (Lt, Dt, St, Cg)	*Regosols (ca, eu),*		SH	5192	
Valenciana		Sd	*Cambisols (ca, eu),*	7.0–8.3	FOR	7478	*P. halepensis*: 69%
(23,255)			*Calcisols (cm, pt, ha),*		GR	842	*P. nigra*: 16%
			Kastanozems (ha)		CUL	1000	*P. pinaster*: 3%
					VOA	4225	*Q. ilex, Q. suber*: 12%
					R(es)	169	
Andalucía	*DSH/SA Mhs*	C_A (Lt, St, Cg, Mb)	*Regosols (ca, eu, dy),*		SH	15,445	*Q. ilex*: 47%
(87,598)	*A/D hs*	A_C (Sch, Sl, Q, G)	*Cambisols (ca, eu, dy),*	5.5–9.5	FOR	13,552	*Eucalyptus* spp.: 9%
	SH Mhs	Sd	*Calcisols (cn, pt, ha)*		GR	6174	*Q. suber*: 9%
			Solonchaks (ha)		CUL	11,353	other deciduous: 6%
					VOA	17,783	*P. halepensis* & *P.pinea*: 17%
					R(es)	5649	other *Pinus*: 12%
Murcia	*Ahs*	C_A (Lt, St, Cg, Ma,	*Regosols (ca, dy, le),*		SH	2272	*P. halepensis*: 84%
(11,314)	*Dhs*	Mb)	*Leptosols (li, ca),*	≥ 8.0	FOR	3017	*P.nigra, P.pinaster* (& others)*:* 9%
		A_C (Sch, Sl)	*Calcisols (pt, ha)*		GR	41	*Q. ilex*: 2%
		Sd	*Solonchaks (ha)*		CUL	1608	Mixed: 5%
					VOA	1962	
					R(es)	1019	
pSpain					SH	85,217	(19.6%)
(no urban)					FOR	163,951	(37.6%)
(482,398)					GR	32,477	(7.5%)
					CUL	95,280	(21.9%)
					VOA	45,509	(10.4%)
					R(es)	13,367	(3.1%)
					Total	435,801	

[a] Reference Soil Groups (IUSS Working Group WRB, 2015) (prototype in non-agricultural areas) (there is no direct correspondence between the cells in this column and those in column 3).

[b] Soil pH range in non-agricultural areas.

[c] Source: MAGRAMA 2010, 2012 (updated with information from the ACs).

conditions); carbonate-C (by Bernard calcimeter method), in soils with pH > 7.5; pH in H_2O (1:2.5 ratio); pH in NaF; and BD (by using cylindrical cores of 5 × 5 cm). Treatment and analysis of the databases were implemented with the GIS software tools (ArcGIS 9.3).

2.2. Calculation of SOC stocks

SOC stocks were calculated from:

$$SOC\ (t\ C\ ha^{-1}) = SOC\ (\%)\ x\ BD\ (g\ cm^{-3})\ x\ [1 - (VG/100)]\ x\ LT\ (cm)$$

where *VG* is the gravel volume (%) and *LT* is the soil layer thickness. In profiles that have data for two layers (0–30 and 30–50 cm), the SOC stock was calculated separately for each and finally computed for the upper 30 or 50 cm. When BD data were not available, estimates were made from the correlations between BD and % SOC obtained for soils derived from same parent material and under same land use. Soil thickness was corrected for zones with thin soils. Based on the land cover map, the areas with a predominance of rocky outcrops occupy near 0.45% of the territory, and the sparsely vegetated eroded zones, 2.70% (Section 2.3.3.). In these areas the SOC stocks were computed for conventional thicknesses of 5 and 15 cm, respectively.

2.3. Environmental predictors

Ten environmental variables were used to estimate SOC stocks and their spatial distribution: elevation, mean annual precipitation (MAP), mean annual temperature (MAT), mean summer precipitation (MSP), mean temperature of the coldest month (Tcm), mean temperature of the hottest month (Thm), land use, lithology, soil type and soil pH.

2.3.1. Elevation

The information included in the database was taken directly from the original source or estimated using the digital elevation model MDT05 (IGN, Ministerio de Fomento). The main relief units in Spain are the following (Fig. 2): (1) Hesperian Massif (Galician Massif, Leon Mts, Central System and Sierra Morena) (800–2500 m altitude), uplifted during the Hercynian orogeny; (2) Central Plateau (around 650 m altitude), divided into two sub plateaus by the mountains of the Central System, uplifted during the Alpine orogeny; (3) chain of mountains surrounding the Central Plateau (Cantabrian Range and Iberian System) (>2000 m altitude), of Alpine origin; (4) mountain systems without connection to the Central Plateau, also of Alpine origin (Coastal Catalan System, Pyrenees and Betic System), these latter exceed 3000 m altitude; and (5) two great ancient tectonic depressions (Ebro and Guadalquivir) at 100–200 m altitude.

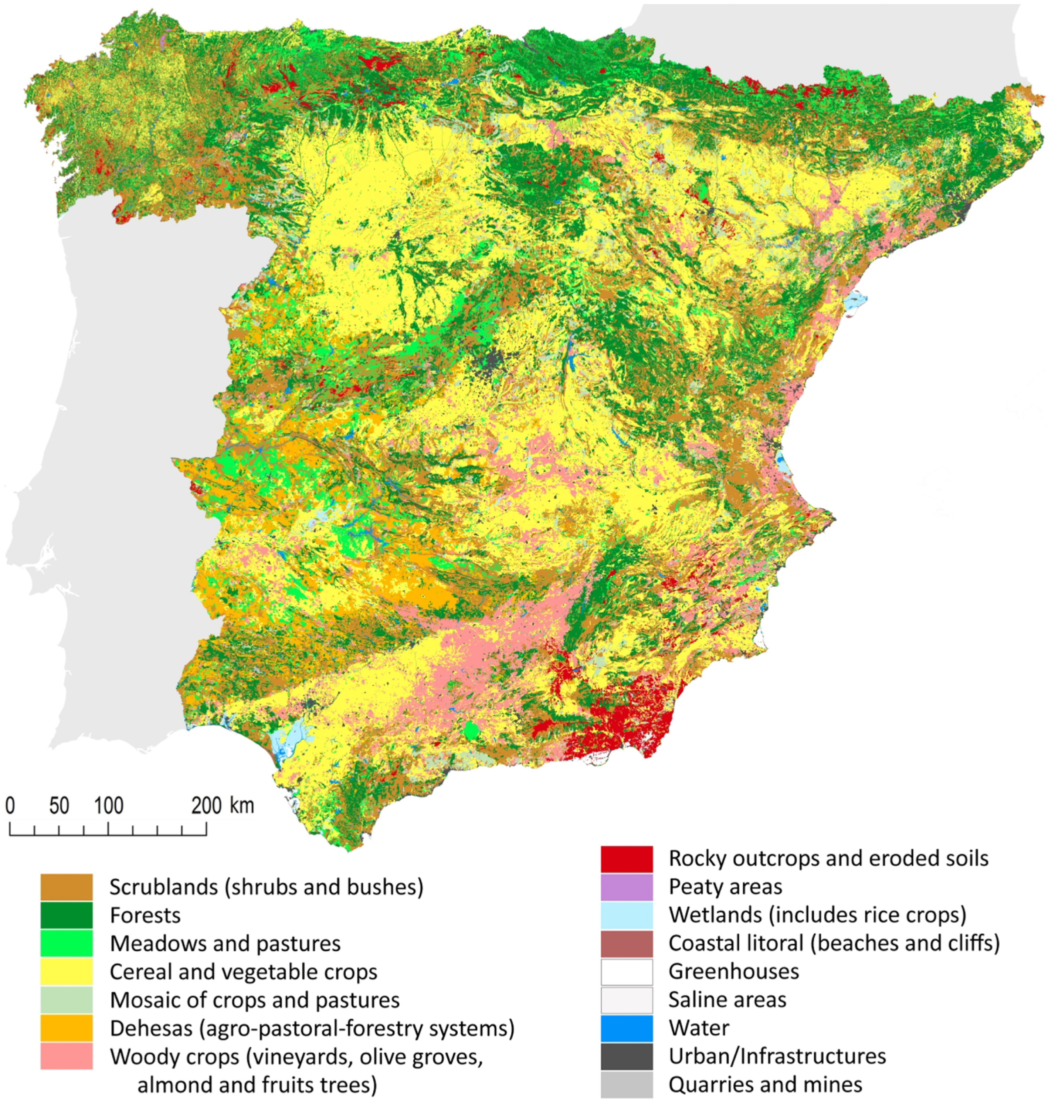

Fig. 4. Land use in pSpain. Synthesis elaborated from different sources: Corine Land Cover (IGN, 2012), NFI3 (MAGRAMA, 2010), and SIOSE (IGN, 2011).

2.3.2. *Climate*

MAP varies widely, from <200 mm (to the southeastern) to >2200 mm (to the north or in elevated areas of some central and southern mountain systems). MAT varies between <2.5 °C in areas of high altitude and >17.5 °C, especially to the south and southeast of the peninsula. Based on the normalized information for the period 1971–2000 (provided by AEMET) a combined map has been prepared for this study, with a sectorization of units according to twelve different climatic types (Fig. 3): desert-hot summer (6^{a}); arid-hot summer (5_s^{a}); semi-arid Mediterranean-hot summer (4_s^{a}); semi-arid Mediterranean-warm summer & cold winter (4_s^{bd}); dry subhumid Mediterranean-hot summer (3_s^{a}); dry subhumid Mediterranean-warm summer & cold winter (3_s^{bd}); subhumid Mediterranean-hot summer (2_s^{a}); subhumid Mediterranean-warm summer (2_s^{b}); subhumid Mediterranean-warm summer & cold winter (2_s^{bd}); humid warm summer (1^{b}, 1_s^{b}); humid-oceanic warm summer (or temperate humid-oceanic) (1_f^{b}); and humid warm summer & cold winter (1_f^{bd}). The predominant climate in a large part of the territory of pSpain (Central Plateau) is semi-arid to dry subhumid Mediterranean with hot summer $(4–3)_s^{a}$ or with warm summer & cold winter $(4–3)_s^{bd}$; it is arid to desert with hot summer $(5–6)_s^{a}$ in the southeast and Ebro depression; humid to subhumid with warm summer and cold winter $(1–2)^{bd}$ in mountains systems without oceanic influence; and temperate humid-oceanic (1_f^{b}) in the north and northwest (Fig. 3 and Table 1).

2.3.3. *Land use*

A digital land use map was prepared using different bases, CORINE Land Cover (IGN, 2012), Third National Forest Inventory (NFI3) (MAGRAMA, 2010) and SIOSE (IGN, 2011) (the latter in regions with a high mix of uses, particularly to the NW of the peninsula). The databases were harmonized and combined according to 16 land use units (Fig. 4). Agricultural and forestry area (AF) occupies around 43.6 Mha (90% of the total area) and includes the following units: areas with predominance of natural forests and plantations (FOR), mainly in mountain systems of subhumid and humid regions (about 38% of the territory); scrublands (SH) (20% of the area); arable crops (CUL), particularly rainfed cereal (22%); woody crops (vineyards, olive groves, almond and fruit trees) (VOA) (10%); meadows and pastures (GR) (8%); rocky outcrops and eroded soils, usually with poor herbaceous coverage, account for about 3% of the AF area (especially in arid zones and in the mountain systems of the Pyrenees, Cantabrian Range and Leon Mts) (Fig. 4 and Table 1).

2.3.4. *Lithology*

Based on the 1:1M digital Geological Map of Spain (IGME, 2015), we have generated a lithological map, in which the 24 units of the

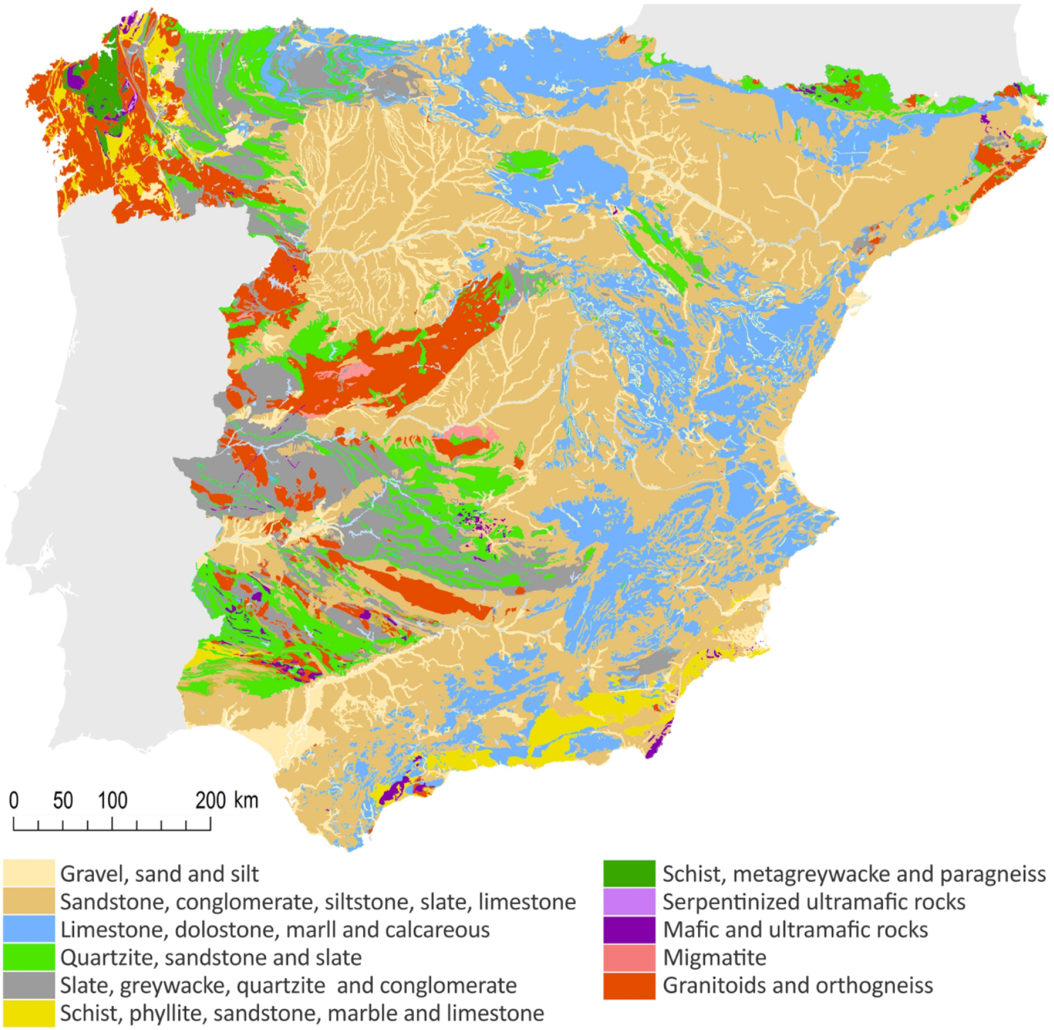

Fig. 5. Lithology of pSpain (synthesis elaborated based on the Geological Map of Spain 1:1M) (IGME, 2015).

geological map were summarized into eleven lithological units (Fig. 5). Siliceous metamorphic rocks (mainly slates, schists and quartzites) and intrusive rocks (mostly granites) are the predominant materials in the Hesperian Massif. Central Plateau presents two large groups of materials: sedimentary rocks (Mesozoic), a large part with calcareous composition or calcareous cement (limestones, marls, sandstones and conglomerates), and clayey to sandy sediments (Tertiary and Quaternary). In the chain of mountains surrounding Central Plateau, in the Pyrenees and in Betic System there is a predominance of calcareous materials. Ebro and Guadalquivir depressions are filled with tertiary and quaternary sediments.

2.3.5. Soil type

The soils of sub-databases 2 and 3 were classified according to World Reference Base (IUSS Working Group WRB, 2015). Since the information taken from the literature (sub-database 1) does not always correspond to this system, we have made a review and adaptation in order to homogenize, as far as possible, the final database. Table 1 shows the predominant Reference Soil Groups (RSGs) in the different autonomous communities (ACs). In arid and desert regions, the most frequent RSGs developed from calcareous materials are *Calcaric Regosols* and *Haplic Calcisols (Calcaric Leptosols* in eroded areas); and from siliceous rocks, *Dystric/Eutric Regosols* and *Leptosols. Solonchaks* are present in certain areas. In semi-arid regions, *Cambic Calcisols* and *Calcaric Cambisols* and *Regosols* predominate on calcareous materials, and *Eutric/Dystric Regosols and Cambisols* on siliceous materials. *Luvisols* and other RSGs with *argic* horizon are frequent on non-calcareous and well-drained sediments. In the dry subhumid to humid transect, the type of soil developed from calcareous materials varies according to the sequence: *Cambic Calcisol - Calcaric Cambisol - Eutric Cambisol - Rendzic Phaeozem/Dystric Cambisol.* In the same transect, the prototype sequence on siliceous materials varies from *Eutric/Dystric Cambisol and Regosols* to *Haplic Umbrisol* (or *Andic Umbrisol/Aluandic Andosol,* in zones with mafic rocks and with no summer moisture deficit). Other RSGs may appear in different parts of the territory depending on the local site characteristics, composition or origin (*Gleysols, Histosols, Arenosols* and *Fluvisols,* among others).

2.3.6. Soil pH

The pH range in the soils of pSpain is very wide, from around 4.5 to 9.5 (Table 1). From the information included in the database, the relationships between soil pH, climate, elevation, parent material and soil type were assessed, and a map of predominant pH ranges was created to be used in this study (Fig. 6). The most acidic soils (pH 4.5–5.5) are preferentially located in humid regions in the N and NW, from siliceous materials (*Umbrisols*) or from intensely decarbonated calcareous materials (*Dystric Cambisols*) (where the pH is often lower than 4.5). The highest pH values are found in arid zones with calcareous materials (*Calcisols* and *Calcaric* RSGs) (pH around 8.3) or in saline environments (*Solonchaks)* (pH $>$ 9.0) (not included as a map unit in Fig. 6). In the rest of the territory, soils are weakly acid to neutral (pH 5.5–7.0), preferentially in subhumid regions with siliceous material (*Dystric/*

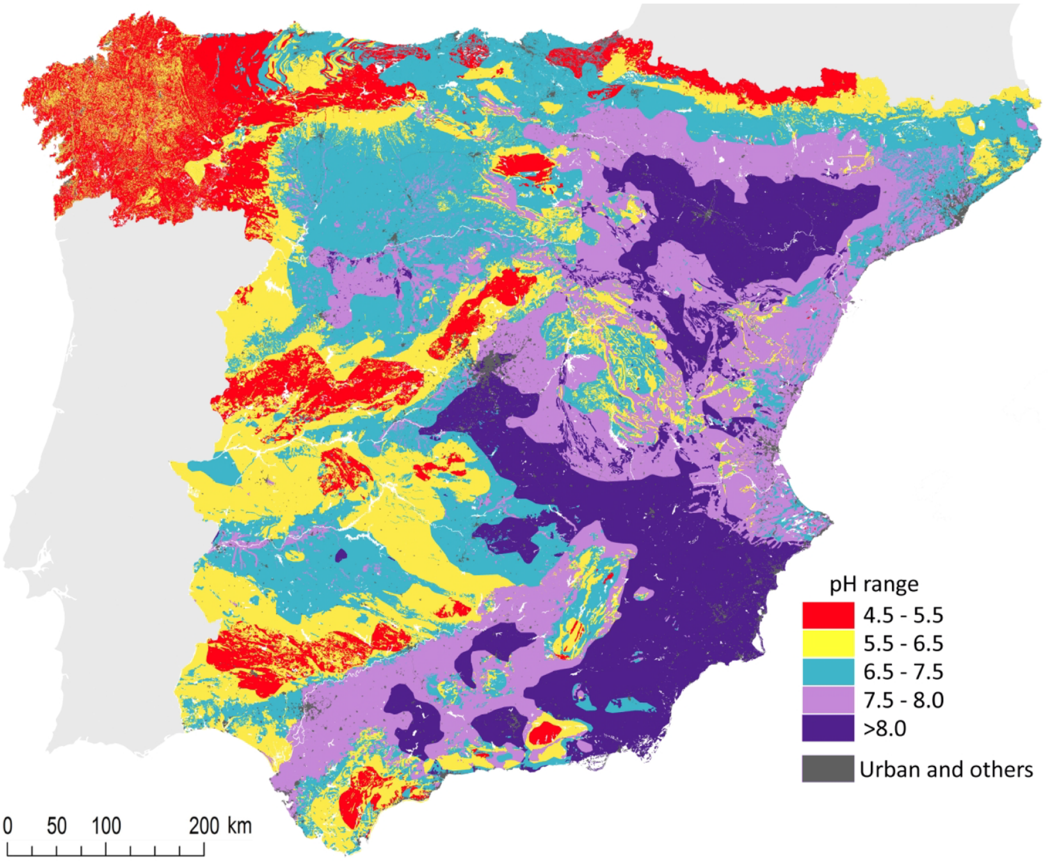

Fig. 6. Predominant soil pH ranges in pSpain. DSM created for this study using RF model (not previously published).

Table 2
SOC content (0–30 cm) in different types of soil and soil/land-use combinations in pSpain. Mean values and coefficients of variation (%) (in brackets) (HS: Histosols; AN: Andosols; CH: Chernozems; UM: Umbrisols; PH: Phaeozems; PZ: Podzols; KS: Kastanozems; GL: Gleysols; AC: Acrisols; LP: Leptosols; AL: Alisols; CM: Cambisols; LV: Luvisols; FL: Fluvisols; CL: Calcisols; RG: Regosols; VR: Vertisols; GY: Gypsisols; SC: Solonchaks; AR: Arenosols; SH: Shrubs; FOR: forests; GR: grasslands; CUL: arable crops; VOA: woody crops; nd: no data available).

% SOC						
RGSs	All data	SH	FOR	GR	CUL	VOA
HS	29.6 (31)	29.1 (36)	31.0 (32)	30.1 (28)	nd	nd
AN	9.2 (43)	9.5 (61)	9.0 (35)	9.3 (48)	4.2 (70)	3.4 (80)
CH	5.9 (63)	4.8 (76)	6.1 (48)	6.0 (60)	nd	nd
UM	5.4 (58)	6.2 (57)	5.5 (59)	4.9 (65)	3.5 (66)	2.7 (64)
PH	4.2 (65)	4.3 (61)	4.0 (66)	4.4 (63)	2.5 (65)	1.5 (60)
PZ	4.2 (68)	3.5 (63)	4.7 (56)	3.6 (68)	nd	nd
KS	3.8 (65)	4.6 (64)	3.6 (69)	3.5 (62)	2.5 (67)	nd
GL	3.4 (98)	4.8 (83)	5.0 (82)	2.8 (70)	1.8 (68)	0.7 (67)
AC	2.4 (65)	1.4 (65)	2.4 (63)	2.9 (67)	1.1 (58)	0.4 (64)
LP	2.4 (81)	2.1 (97)	2.4 (74)	3.4 (62)	nd	0.9 (38)
AL	2.1 (76)	2.0 (75)	2.1 (73)	1.9 (70)	1.0 (68)	nd
CM	2.0 (70)	1.4 (66)	2.8 (64)	1.4 (77)	1.2 (48)	1.0 (59)
LV	1.9 (61)	1.5 (72)	2.3 (60)	2.5 (65)	1.0 (57)	0.7 (53)
FL	1.4 (85)	0.8 (78)	1.6 (75)	1.7 (81)	0.6 (76)	nd
CL	1.2 (77)	1.6 (59)	2.1 (63)	2.3 (75)	1.1 (67)	0.9 (75)
RG	1.2 (77)	1.5 (73)	2.2 (69)	1.2 (81)	1.0 (61)	1.2 (67)
VR	1.0 (53)	1.0 (52)	0.9 (63)	1.2 (72)	1.0 (61)	0.9 (60)
GY	0.9 (74)	1.0 (70)	1.1 (65)	0.5 (66)	0.4 (68)	0.4 (67)
SC	0.7 (83)	0.8 (97)	nd	1.2 (77)	0.9 (64)	0.6 (63)
AR	0.5 (63)	0.6 (60)	0.6 (90)	0.7 (56)	0.1 (62)	0.1 (60)

Eutric RSGs), or neutral to basic (pH 7.0–8.0) on partially decarbonated calcareous materials (*Eutric/Calcaric* RSGs).

2.4. Modeling and spatial prediction

2.4.1. General statistics and distance correlation between predictor variables

In a first step, the relationships between SOC content and the different covariates -taken individually or in combinations- were analyzed, and statistics were calculated for the two layers (0–30 and 30–50 cm) in the different pSpain regions. Later, an analysis of Distance Correlation (DC) (Székely et al., 2007) between all covariates was carried out. DC is a measure of dependence between random vectors and has the advantage over other classical methods in that it can detect more complex relationships than linear ones and characterize the independence of the variables of any dimension (Pearson's correlation coefficient only characterizes independence in univariate Gaussian distributions). DC coefficient takes only values in [0,1] (independence or complete dependence, respectively) as the usual determination coefficient (R^2) although its interpretation is not in terms of explained variability.

2.4.2. Random forest model

Random Forest (RF) regression (Breiman, 2001) was used as modeling tool to predict SOC concentrations and stocks, and its spatial distribution in each of the layers, 0–30 and 30–50 cm depth. RF consists of an ensemble of randomized classification and regression trees (CART) in which many decision trees are built using a random subsample of the available covariates/cases. The final result is one single prediction constructed as a weighted average over all these suboptimal trees. RF is considered to have high predictive performance, low correlation of individual trees, small bias and variance, reliable error estimates and provision of information on the relative importance of

Table 3

SOC content (0–30 and 30–50 cm) under different types of land use in the ACs of pSpain. Mean values and coefficients of variation (%) (in brackets) (SH: Shrubs; FOR: forests; GR: grasslands; CUL: arable crops; VOA: woody crops; in italics: number of samples).

% SOC

AC_s	0–30 cm						30–50 cm					
	SH	FOR	GR	CUL	VOA	W. Mean[a]	SH	FOR	GR	CUL	VOA	W. Mean[a]
Galicia (1)	6.5 (51) *658*	6.9 (43) *901*	5.1 (57) *630*	4.2 (57)[b] *428*	3.1 (45)[b] *107*	6.1 (49) *2724*	3.6 (50) *59*	4.9 (43) *166*	2.8 (64) *92*	1.8 (39) *159*	2.6 (46) *157*	3.8 (45) *633*
Asturias (2)	4.3 (44) *65*	4.1 (46) *324*	3.9 (46) *167*	2.1 (52) *124*	1.5 (60) *94*	3.8 (46) *774*	2.5 (60) *15*	3.2 (51) *93*	1.5 (70) *30*	1.2 (46) *32*	0.9 (42) *12*	2.4 (56) *182*
Cantabria (3)	3.5 (51) *90*	3.3 (52) *455*	3.2 (56) *234*	2.1 (52) *173*	1.5 (60) *131*	3.3 (53) *1083*	2.0 (80) *13*	1.5 (62) *60*	1.3 (73) *49*	0.8 (69) *46*	0.4 (46) *15*	1.6 (70) *183*
P.Vasco (4)	3.6 (47) *106*	3.4 (47) *517*	3.2 (55) *368*	2.2 (50) *198*	1.6 (50) *49*	3.2 (49) *1238*	2.2 (90) *16*	2.0 (60) *76*	1.2 (83) *43*	0.7 (71) *28*	0.5 (40) *21*	1.5 (70) *184*
Navarra (5)	2.6 (73) *35*	2.9 (64) *104*	3.3 (61) *40*	1.1 (46) *238*	1.0 (71) *131*	2.3 (59) *548*	0.8 (65) *21*	1.0 (95) *44*	1.3 (63) *29*	0.9 (55) *63*	0.9 (26) *57*	1.0 (74) *214*
La Rioja (6)	2.5 (68) *18*	2.8 (61) *35*	3.4 (59) *29*	1.0 (50) *17*	1.1 (75) *26*	2.0 (62) *125*	0.7 (57) *9*	0.8 (86) *13*	1.3 (65) *18*	0.7 (47) *13*	0.8 (37) *16*	0.8 (61) *69*
Aragón (7)	2.1 (71) *17*	2.8 (68) *67*	3.7 (67) *43*	1.2 (56) *178*	1.2 (67) *38*	2.1 (64) *343*	1.3 (67) *7*	0.9 (78) *29*	2.3 (65) *15*	1.1 (97) *63*	1.8 (76) *27*	1.1 (82) *141*
Cataluña (8)	4.0 (25) *40*	3.3 (70) *140*	3.4 (82) *37*	1.6 (69) *140*	1.3 (62) *70*	2.8 (65) *427*	2.9 (73) *27*	0.9 (89) *21*	2.9 (89) *15*	0.9 (68) *33*	0.8 (81) *32*	1.2 (82) *128*
C-León (9)	3.5 (57) *47*	2.8 (71) *229*	2.1 (76) *101*	1.0 (60) *356*	0.7 (67) *16*	2.1 (61) *749*	1.5 (78) *73*	1.9 (96) *91*	1.1 (87) *39*	0.4 (66) *84*	0.3 (42) *13*	1.1 (78) *300*
Madrid (10)	0.9 (56) *45*	2.5 (80) *95*	1.1 (82) *31*	0.9 (83) *42*	0.7 (60) *10*	1.4 (77) *223*	0.5 (53) *14*	1.5 (90) *19*	0.5 (95) *16*	0.4 (35) *8*	0.4 (26) *6*	0.9 (70) *63*
Extremadura (11)	1.1 (55) *53*	2.3 (65) *46*	1.1 (64) *63*	0.9 (56) *101*	0.9 (56) *49*	1.6 (60) *312*	0.8 (56) *9*	1.7 (77) *34*	0.5 (72) *23*	0.7 (97) *46*	0.2 (71) *10*	1.2 (84) *122*
C-La Mancha (12)	1.6 (69) *63*	2.6 (58) *72*	1.6 (94) *70*	1.1 (55) *540*	1.0 (60) *204*	1.7 (59) *949*	1.2 (87) *18*	0.4 (30) *30*	1.5 (87) *18*	0.8 (38) *28*	0.7 (36) *18*	0.7 (42) *112*
C. Valenciana (13)	2.9 (76) *76*	2.6 (77) *116*	2.2 (68) *11*	1.0 (50) *66*	1.0 (60) *120*	2.0 (68) *389*	1.7 (90) *35*	1.9 (85) *51*	1.0 (72) *8*	1.0 (89) *30*	0.8 (59) *23*	1.4 (83) *147*
Andalucia (14)	1.5 (87) *129*	2.1 (71) *393*	1.7 (76) *91*	1.0 (60) *429*	0.9 (56) *374*	1.3 (67) *1416*	1.1 (90) *145*	1.5 (87) *328*	1.0 (95) *15*	0.8 (67) *377*	0.7 (66) *109*	1.0 (75) *974*
Murcia (15)	1.1 (82) *276*	1.6 (69) *261*	0.9 (67) *161*	0.9 (67) *456*	0.7 (86) *270*	1.2 (76) *1424*	0.4 (57) *27*	0.4 (63) *27*	0.5 (73) *29*	0.4 (64) *49*	0.2 (75) *23*	0.4 (65) *155*
pSpain[a]	2.7 (65) *1718*	3.1 (66) *3755*	2.9 (70) *2076*	1.2 (58) *3486*	1.0 (61) *1689*	2.2 (62) *12,724*	1.4 (70) *479*	1.7 (71) *1091*	1.6 (77) *439*	0.8 (63) *1065*	0.7 (52) *533*	1.2 (70) *3607*

[a] Weighted mean based on the area occupied by each type of land use in the different ACs.

[b] In AC1, traditional cropland management includes organic amendments (manure and others).

predictors (Breiman, 2001; Liaw and Wiener, 2002; Svetnik et al., 2003; Arun and Langmead, 2005). The model has been successfully used in spatial prediction of SOC stocks (e.g. Grimm et al., 2008; Wiesmeier et al., 2011; Vågen et al., 2013; Sreenivas et al., 2016; Apka et al., 2016). For RF computations, we used the "RandomForest" package (Liaw and Wiener, 2018). The default options of the RF function in the R package were applied. We have only defined the following parameters required by the model: number of trees to be built in the forest (ntree = 2000); number of predictors to be used in each tree-building process (mtry = 2); and minimum number of data points in each terminal node (nodesize = 10).

2.4.3. Model performance

RF model carry out an internal performance evaluation by using a large independent test dataset that were not used in the training procedure, Out-Of-Bag (OOB) samples (one third of the data). These OOB samples are predicted by its corresponding bootstrap training tree and by aggregating the OOB predictions from all trees the mean error (ME), root mean square error (RMSE) and coefficient of determination (R^2) are calculated (Eqs. (2), (3) and (4)).

$$ME = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \tag{2}$$

$$RMSE = \sqrt{n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4}$$

where n is the number of observations, y_i and $\hat{y}_i$ are observed and predicted SOC content values at the ith point and $\bar{y}$ is the average of the y_i values.

Additionally, we have applied a cross-validation test. The entire dataset was randomly divided into training and validation subsets (90% and 10% of the data, respectively). In the training subset RF was applied (ntree = 500) and the OOB statistics for internal validation were calculated. These same parameters were computed for the validation subset. The procedure was repeated 100 times and at each run the selection of both subsets was random.

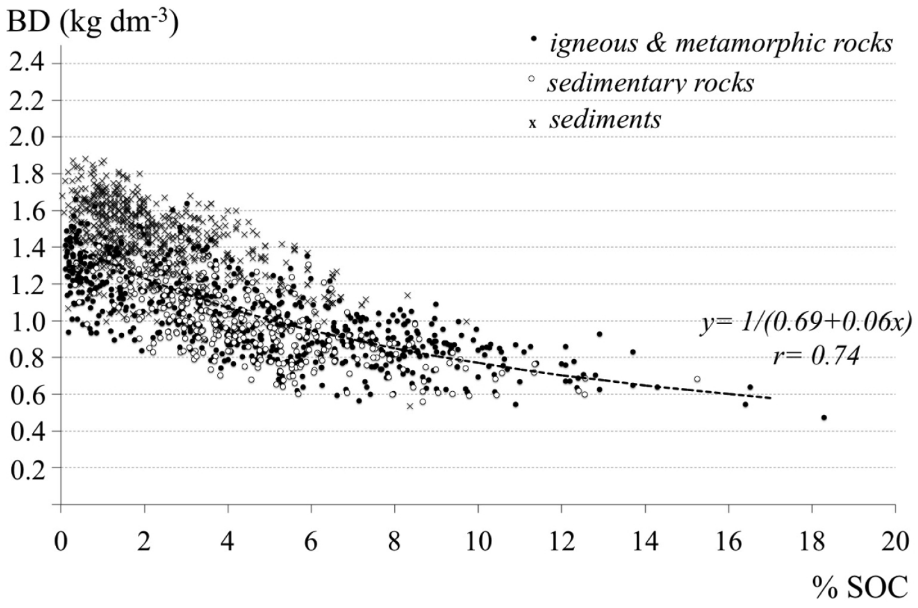

Fig. 7. Bulk density *versus* SOC content in soils developed from different materials (n = 1696) (data for tilled lands are not included here; the equation corresponds to the set of points).

3. Results and discussion

3.1. Observed SOC content and stocks

3.1.1. Global and regional statistics for different soil type and land use

The average concentration of SOC for the main soil types existing in pSpain ranges between 0.5% in *Arenosols* and 30%, in *Histosols* (Table 2). In all cases, the highest values are located under forests and scrubs, and the lowest under crops (especially woody crops). Anyway, the variability is very high and these mean values cannot be used to make estimates at the national level (the coefficient of variation -CV- in the units with the greatest representation in the territory is >70%).

A best approximation is observed from data obtained at a regional scale (Table 3). The lowest mean values (0.7–1.6%) (0–30 cm) are registered in AC15, and the highest (3.1–6.9%) in AC1 (autonomous communities located at the SE and NW ends of the peninsula, respectively). In the subsurface (30–50 cm) layer the SOC content decreases between 1.4 and 2 times, and minimum and maximum remain to the SE and NW, respectively.

The SOC stocks were calculated by applying eq. 1. In some samples, BD was estimated from the relationships between BD and %SOC observed in soils with the same lithology and land use. In all cases, the best fit was obtained from the equation $BD = 1/(a + b\,\%SOC)$, where a and b are constants (Fig. 7). The largest SOC stocks are in forest soils. In these soils, the mean SOC density ranges between 58 and 180 t ha^{-1} (0–30 cm) and between 70 and 290 t ha^{-1} (0–50 cm), from the SE to the NW respectively (Table 4). In general, carbon pools are higher in regions with the largest forest area. The only notable exception is observed in AC1, a small region that ranks 6th in forest area and 2nd in total SOC reserves. The resulting mean weighted SOC pool for pSpain is 3254.6 ± 1649 Mt (0–30 cm) and 4276.9 ± 2095 Mt (0–50 cm). They are rough estimates (CV is around 50%).

3.1.2. Climate

The strong climatic contrasts in pSpain exert an important influence on the variability of the SOC reserves. Fig. 8 shows the median values and interquartile range of SOC content -under forests and scrublands-plotted against precipitation ranges, in regions with different moisture and temperature regime. The SOC content is minimal in arid and desert regions and increases as the MAP increases. However, the behavior pattern is not homogeneous for all soils in humid zones. As prototype examples of this variability, Fig. 8 shows the data from the ACs 1 to 4, neighboring regions and with the same temperate humid-oceanic climate. The variation range is similar under forest and scrubland. SOC content varies gradually between about 1 and 4% (0–30 cm) from 200 to 1700 mm of MAP, if the ACs 2–4 data are considered. However, the AC1 data are not consistent with this sequence, showing a high increase rate from 1000 mm of precipitation, up to double the mean values in the other humid regions, in the MAP range of 1500–1700 mm. According to this, the climate factor is far from explaining properly the variations in SOC concentration for a same land use.

3.1.3. Parent material and soil properties

The variations of SOC content for different lithology and soil type, in forest soils of Mediterranean and Humid-Oceanic regions, are shown in Fig. 9. The influence of materials is minimal in arid and desert regions, where the moisture deficit is the decisive factor. In semi-arid to subhumid zones (MAP 400–800 mm), the average SOC content is around 1% higher in soils developed on calcareous materials (predominantly *Calcisols* and *ca Regosols*), pH ≈ 8.3, than in neutral or weakly acid soils derived from siliceous materials (*eu/dy Regosols* and *Cambisols*). The situation is reversed when the precipitation increases and most of the calcareous materials have undergone decarbonation, at least in topsoil. For a range of MAP between 900 and 1000 mm, the average SOC content of these soils (mainly *ca/eu Cambisols*) (pH 8.0–6.5) is similar to or lower than in siliceous materials (*dy/eu Cambisols* and *Regosols*).

The highest variability is definitely observed in humid zones (Fig. 9, right). In these climatic conditions, most of the calcareous materials are completely decarbonated in the upper layer and the pH ranges from 7.5 to <4.0 (CEC saturated by H^+*)* (*rz Phaeozems* to *dy Cambisols)*. The SOC concentration is lower than in other materials and remains practically constant (at around 3.0%) from 1000 to 1700 mm of precipitation. Similar values are observed in neutral to acid soils developed from serpentinized ultramafic rocks (*ha/le Phaeozems* and *dy/eu Cambisols)*. Soils from granite and other quartz-rich rocks are acidic (pH 5.5–4.5) (Al buffer) and have an *umbric* epipedon with a mean SOC content two to three times higher than those from limestone and serpentinites. In addition, SOC increases at a high rate as the MAP increases (up to about 8% for 1500–1700 mm). Soils on mafic rocks are also acidic (pH

Table 4
SOC stocks (0–30 and 0–50 cm) under different types of land use in the ACs of pSpain. Upper row: Mean and standard deviation (tons C ha^{-1}); lower row: Mean (million tons C) and coefficient of variation (%) (in brackets) (SH: Shrubs; FOR: forests; GR: grasslands; CUL: arable crops; VOA: woody crops).

SOC stock												
	0–30 cm						0–50 cm					
ACs	SH	FOR	GR	CUL	VOA	W. Mean[a]	SH	FOR	GR	CUL	VOA	W. Mean[a]
Galicia (1)	175 ± 104	182 ± 119	146 ± 93	126 ± 79[b]	98 ± 49[b]	163 ± 103	245 ± 139	291 ± 185	205 ± 125	164 ± 98[b]	128 ± 64[b]	237 ± 153
	75.3 (37)	249.7 (31)	65.8 (44)	47.8 (46)	6.0 (38)	450.1 (35)	105.4 (39)	399.6 (35)	92.1 (34)	62.1 (23)	7.8 (20)	672.5 (25)
Asturias (2)	128 ± 67	128 ± 67	128 ± 73	73 ± 42	55 ± 35	120 ± 64	173 ± 88	193 ± 99	154 ± 87	92 ± 52	71 ± 45	166 ± 62
	31.2 (35)	56.7 (35)	36.2 (39)	1.3 (45)	0.1 (52)	127.0 (36)	42.1 (26)	85.1 (42)	43.5 (36)	1.6 (37)	0.2 (53)	173.8 (45)
Cantabria (3)	110 ± 61	106 ± 64	103 ± 65	73 ± 42	55 ± 35	104 ± 62	143 ± 85	148 ± 84	113 ± 72	88 ± 50	71 ± 44	134 ± 78
	16.4 (42)	22.5 (42)	11.7 (47)	0.5 (45)	0.01 (52)	52.1 (42)	21.3 (38)	31.6 (46)	12.9 (40)	0.6 (43)	0.01 (23)	67.3 (43)
P.Vasco (4)	113 ± 58	108 ± 61	103 ± 61	76 ± 42	58 ± 31	103 ± 56	147 ± 80	151 ± 79	113 ± 65	92 ± 47	76 ± 39	135 ± 73
	11.1 (38)	42.9 (38)	14.1 (44)	3.9 (42)	0.9 (43)	73.7 (39)	14.5 (34)	60.0 (35)	15.5 (48)	4.6 (45)	1.2 (43)	96.6 (40)
Navarra (5)	87 ± 67	95 ± 67	106 ± 70	42 ± 20	42 ± 31	77 ± 50	114 ± 89	124 ± 87	137 ± 90	58 ± 28	58 ± 45	101 ± 66
	12.1 (61)	42.7 (54)	9.5 (50)	11.2 (40)	1.9 (64)	78.0 (53)	15.7 (57)	55.5 (51)	12.3 (45)	15.8 (37)	2.6 (55)	102.6 (50)
La Rioja (6)	87 ± 67	95 ± 65	106 ± 70	42 ± 20	42 ± 31	77 ± 50	114 ± 83	124 ± 85	137 ± 91	58 ± 27	58 ± 40	101 ± 65
	11.1 (40)	15.8 (60)	0.8 (49)	2.7 (38)	2.4 (65)	33.4 (55)	14.4 (47)	20.6 (61)	1.1 (55)	4.1 (38)	3.1 (34)	43.8 (41)
Aragón (7)	73 ± 55	93 ± 67	115 ± 85	45 ± 27	45 ± 31	71 ± 50	95 ± 71	121 ± 88	150 ± 90	63 ± 38	63 ± 43	90 ± 64
	78.6 (61)	143.2 (57)	27.6 (55)	50.1 (51)	12.0 (58)	318.1 (56)	102.2 (55)	186.2 (53)	35.9 (44)	70.1 (51)	16.8 (58)	417.8 (58)
Cataluña (8)	122 ± 38	106 ± 79	108 ± 93	58 ± 42	48 ± 31	93 ± 62	171 ± 53	148 ± 100	151 ± 120	76 ± 45	68 ± 36	93 ± 86
	40.3 (20)	169.6 (57)	21.8 (68)	29.5 (59)	12.5 (54)	278.0 (52)	56.4 (51)	237.4 (55)	30.6 (55)	38.4 (47)	17.5 (34)	384.5 (49)
C-León (9)	110 ± 70	93 ± 70	73 ± 58	38 ± 24	24 ± 16	74 ± 49	143 ± 84	111 ± 90	95 ± 75	54 ± 31	33 ± 21	94 ± 62
	206.4 (47)	273.3 (60)	30.6 (65)	116.0 (53)	5.1 (60)	638.8 (54)	268.3 (44)	327.9 (55)	39.8 (48)	162.5 (51)	7.2 (58)	813.0 (51)
Madrid (10)	35 ± 20	85 ± 70	42 ± 35	24 ± 20	24 ± 19	53 ± 41	45 ± 25	102 ± 85	58 ± 47	31 ± 25	30 ± 24	66 ± 51
	3.6 (49)	21.9 (68)	5.0 (72)	3.1 (74)	0.8 (74)	34.4 (67)	4.6 (47)	26.2 (63)	7.0 (70)	4.1 (63)	1.0 (64)	43.0 (55)
Extremadura (11)	42 ± 24	79 ± 55	42 ± 27	35 ± 20	35 ± 19	59 ± 37	54 ± 28	103 ± 70	54 ± 35	49 ± 24	48 ± 23	78 ± 49
	32.5 (48)	149.7 (55)	10.3 (56)	14.3 (49)	13.2 (49)	220.5 (53)	42.2 (46)	194.6 (49)	13.4 (45)	20.1 (37)	18.5 (38)	289.2 (46)
C-La Mancha (12)	58 ± 55	87 ± 42	58 ± 25	38 ± 24	38 ± 24	61 ± 37	70 ± 49	101 ± 62	70 ± 63	50 ± 26	50 ± 29	72 ± 44
	51.7 (59)	236.9 (48)	13.7 (81)	89.1 (48)	31.3 (53)	425.5 (51)	62.1 (41)	272.4 (43)	16.4 (74)	107.0 (36)	40.6 (41)	501.3 (46)
C. Valenciana (13)	95 ± 70	87 ± 76	76 ± 55	38 ± 20	38 ± 24	75 ± 52	124 ± 89	114 ± 99	99 ± 70	50 ± 25	54 ± 30	98 ± 68
	49.5 (63)	65.4 (65)	6.4 (58)	3.8 (44)	16.1 (53)	142.1 (62)	64.4 (59)	85.0 (61)	8.4 (52)	5.0 (41)	22.6 (50)	186.1 (47)
Andalucia (14)	55 ± 55	73 ± 48	61 ± 48	38 ± 24	35 ± 20	48 ± 33	71 ± 60	88 ± 65	80 ± 59	50 ± 29	49 ± 27	61 ± 41
	84.9 (75)	99.4 (61)	37.8 (66)	43.4 (53)	61.6 (49)	342.7 (59)	110.4 (84)	119.3 (65)	49.2 (58)	56.4 (62)	86.3 (48)	437.0 (50)
Murcia (15)	42 ± 35	58 ± 42	35 ± 24	34 ± 24	27 ± 21	41 ± 31	54 ± 43	70 ± 49	45 ± 29	45 ± 28	27 ± 24	49 ± 35
	9.5 (72)	17.5 (59)	0.1 (59)	5.6 (59)	5.4 (76)	40.2 (61)	12.3 (63)	21.1 (52)	0.2 (47)	7.2 (48)	5.4 (55)	48.3 (57)
pSpain[a]	84 ± 57	98 ± 68	100 ± 66	44 ± 26	38 ± 23	75 ± 48	110 ± 75	129 ± 89	130 ± 83	57 ± 35	51 ± 30	97 ± 43
	714.1 (51)	1607.3 (51)	291.5 (54)	422.4 (51)	169.3 (52)	**3254.6 (51)**	936.3 (46)	2122.5 (49)	378.1 (50)	559.1 (49)	231.0 (46)	**4276.9** (49)

[a] Weighted mean (data of eroded lands, usually with poor herbaceous coverage, are included).

[b] Traditional cropland management includes organic amendments (manure and others).

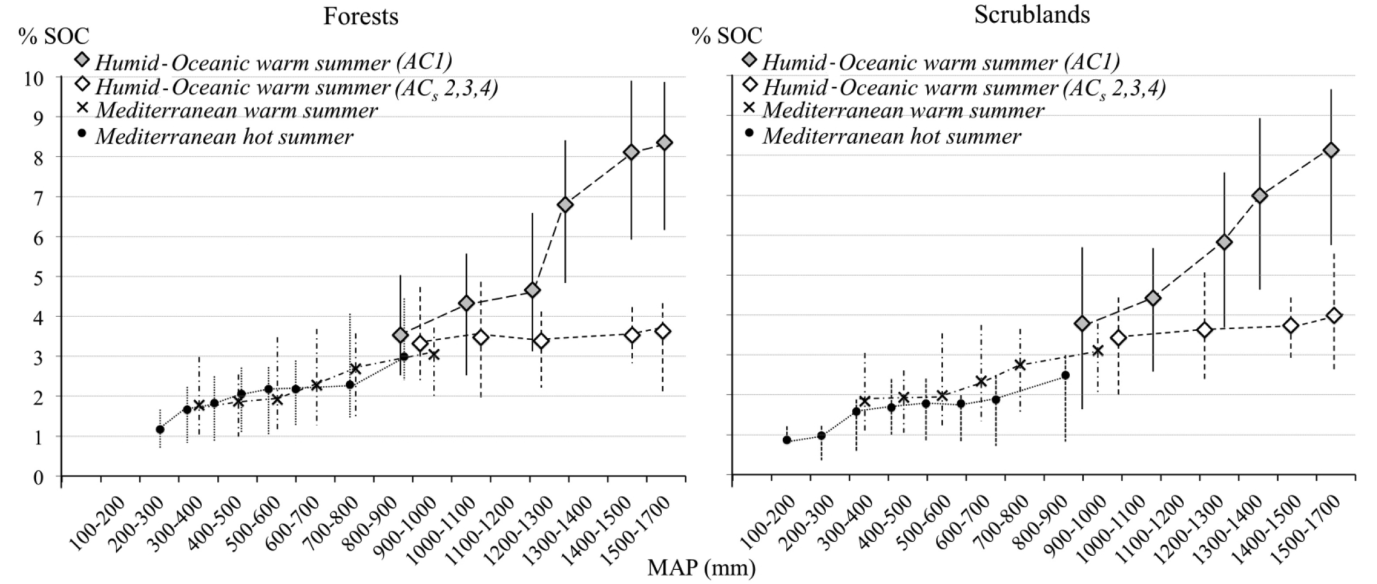

Fig. 8. SOC content plotted against ranges of MAP. Median and interquartile range (0–30 cm) in forest and scrubland soils in different climate zones of pSpain (for humid zones, data from the ACs 1 to 4 are shown) (see Fig. 2).

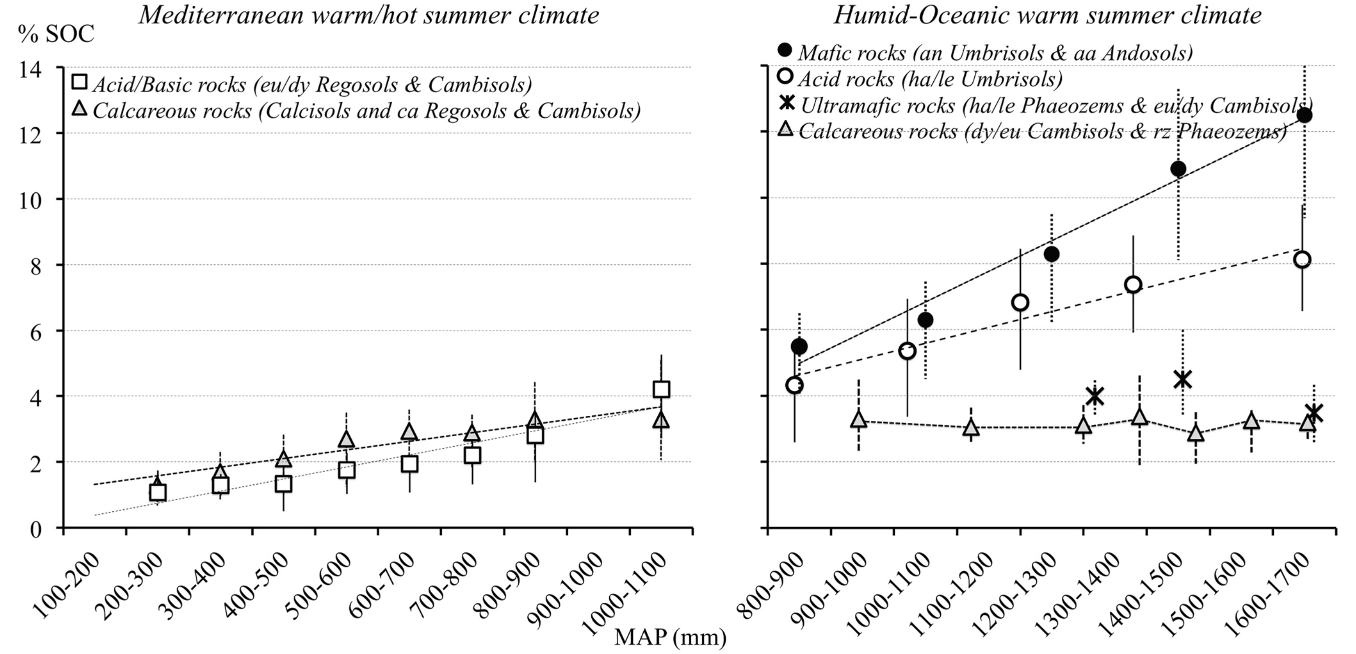

Fig. 9. Variation in SOC content in relation to climate and lithology-soil type: % SOC *vs* MAP in forest soils of Mediterranean and humid-oceanic regions (median and interquartile range, 0–30 cm) (ultramafic rocks are serpentinites and their vegetal cover, scrub).

5.0–5.5) and develop an *umbric* epipedon. In these soils the mean SOC content is 1.6 times higher than in granites (around 13% for 1500–1700 mm of precipitation). A synthesis of SOC content variations against different pH intervals is shown in Fig. 10.

Based on these results, the distribution of SOC content for a given land use seems to be regulated by the presence or absence of active lime, in Mediterranean semi-arid to subhumid regions; and by the presence or absence of active Al, in Humid-Oceanic regions. Selective extractions of C, Al and Fe in these last zones help to support this hypothesis.

3.1.3.1. Selective extractions and estimation of SOC fractionation in humid regions. The results obtained by selective extractions of C, Al and Fe in forest soils of humid regions are shown in Fig. 11 and Table 5. From these data we have estimated the relative importance of different organic carbon species, such as metal-C complexes (by pyrophosphate extraction); active, short-range-order or amorphous Al and Fe (by oxalate extractions); and free Al and Fe (by dithionite extractions). Simple regression analysis reveals the key role of Al_p in explaining the variation in SOC concentration. For most materials, moderate to high positive correlation between Al_p and C_p, and between C_p and total SOC is observed (Fig. 11). There is no correlation between C_p and Fe_p, nor between C_p and $Fe_p + Al_p$ (data not shown).

3.1.3.1.1. Calcareous materials. Soils developed on calcareous rocks have very low concentrations of free Al and Al_p, so the % C_p is also very low and represent a small proportion of the total organic C (C_t) (the

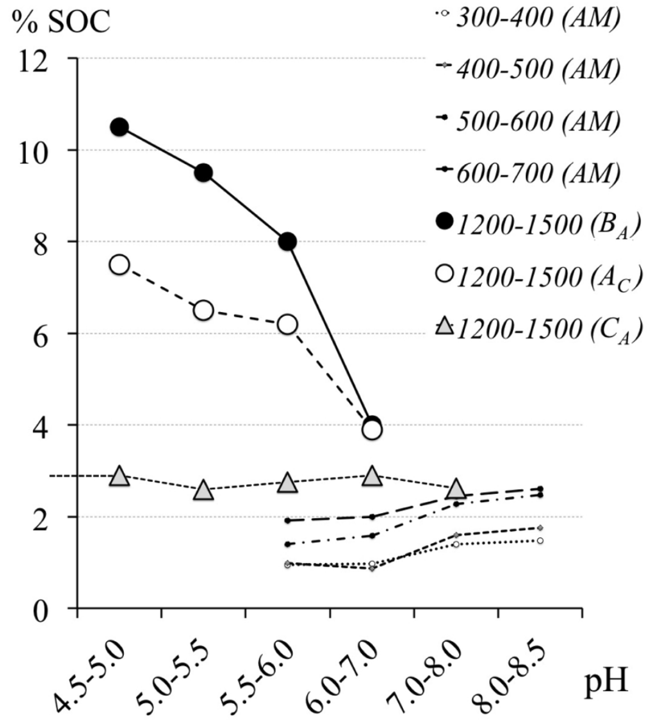

Fig. 10. Average SOC content *versus* soil pH range in zones with different mean annual precipitation (mm) (in italics) (AM- all materials; B_A: basic rocks; A_C: acid rocks; C_A: calcareous materials).

average C_p/C_t ratio is 0.19) (Table 5). In addition, the mean Al_p/C_p ratio is <0.1, indicating a low saturation of the organic component of the complex (Higashi et al., 1981; Oades, 1989). Non-oxidizable organic C (recalcitrant-C) represents $<3\%$ of the total organic C. Thus, most of the SOC in these soils (almost 80%) appears to be labile-C and around 19% is involved in C-metal complexes of low stability, which allows us to suppose a high rate of organic matter decomposition in these temperate humid environments. The scarce variability in the SOC concentration in a wide range of MAP (Fig. 9, right), despite the likely increase in C input, appears to support this hypothesis, and even that these soils may have reached a steady state for the organic C (around 3%). Under this assumption, and supposing a similar *climate balance* in all soils in the same region, excess SOC above this threshold value would enable estimation of the importance of mechanisms stabilization and accumulation of carbon in other materials.

3.1.3.1.2. Granitic rocks. The previously indicated threshold is greatly exceeded in the soils *(ha Umbrisols)* derived from granites and other acid rocks. In these environments, the hypothesis of *climate balance* for organic carbon seems very likely given the high decomposition rate they present. In a study carried out in a climate transect in Europe, from subartic to subtropical and Mediterranean regions, the forest soils from granite on NW Spain showed the highest litter mass-losses rates (40–50% in the first year after falling) (Berg et al., 1993). Moreover, the importance of the processes of stabilization by formation of C-Al complexes can be deduced from the concentrations of Al_{DC}, Al_p and C_p, which are several times higher than those of calcareous materials. C-Al fraction represents around 40% of the total C (Table 5 and Fig. 11) and, given the higher availability of active Al, the stability of these species is also higher ($Al_p/C_p > 0.1$).

A study on the formation of organo-metal complexes in forest soils in these regions has enabled recognition of the presence of mobile precursors in the liquid fraction of the upper layers (O/A horizons), either in the form of water-soluble C-Al monomers or as acid-soluble polymers (Calvo de Anta and Álvarez, 1992). An increase in the complexity of these polymers would lead its precipitation in the few cm of soil, or at greater depth if a vertical or lateral mobilization of these species occurs (Masiello et al., 2004; Schrumpf et al., 2013). The latter could explain the high SOC content in subsurface layer (30–50 cm) of these soils (data for AC1 in Table 3). Finally, the existence of stabilization mechanisms via the formation of organo-mineral complexes must also be considered, as these media have a predominance of variable-charge clays (kaolinite/halloysite and gibbsite) (Calvo de Anta and Macías, 1993) and show a high anion fixation capacity (mean phosphate retention is $>85\%$) (Table 5). The importance of C stabilization processes in these climate environments has been revealed by radiocarbon dating analysis in deep horizons (60–170 cm) of buried soils. The values obtained range between 1000 and 6150 ± 30 years before present, i.e. 1950) (Calvo de Anta et al., 2014).

Based on the data in Table 5 we have estimated a possible organic C fractionation as follows: for a total content of 8.0% (average for 1400–1700 mm of precipitation), C_r and C_p would be around 0.5% and 3.2% respectively. Assuming *climate balance*, the labile-C fraction would be close to 3.0%. The remaining 1.3% could correspond to clay-C complexes. That is, a distribution of 6, 40, 17 and 37% for recalcitrant-C, metal-C complexes, mineral-C complexes and labile-C respectively.

3.1.3.1.3. Mafic rocks. Soils derived from basic rocks (predominantly gabbro and amphibolite) have the highest values of Al_{DC}, Al_p and C_p (this latter is 1.3 to 6 times higher than in soils developed from other materials), and the Al_p/C_p ratio approaches 0.2. Namely, they have the highest, and most stable, content of metal-C complexes, which account for 42% of the total organic C (Table 5). In addition, in these climate zones these materials develop *andic* properties *(Andic Umbrisols and Aluandic Andosols*) (García-Rodeja et al., 1987; Macías and Calvo de Anta, 1992), i.e. high capacity of anion fixation and formation of organo-mineral complexes (e.g. Wada 1985; Torn et al., 1997). An estimate of the organic C fractionation in these soils is as follows: for an SOC content of 11% (mean for MAP 1400–1700 mm), C_r and C_p would be around respectively 2.1% and 4.6% and (assuming a labile-C content of 3.0%) the remaining 1.3% would be constituted by organo-mineral complexes (19, 42, 12 and 27% for respectively recalcitrant-C, metal-C complexes, mineral-C complexes and labile-C).

3.1.3.1.4. Serpentinized ultramafic rocks. Soil properties in these environments vary greatly depending on the degree of serpentinization of the material (Table 5). In the most intensely serpentinized rocks, the soils are poorly developed, neutral and saturated by Mg (Ca/Mg ratio in CEC < 0.2) (*ha/le Phaeozems* are predominant). These soils have very low contents of Al_{DC}, Al_{ox}, Al_p and C_p, and the total SOC is of the same order as in calcareous materials and also appears to be mainly constituted by labile-C. The C-Al complexes represent 27% of the total organic C and are unstable (Al_p/C_p ratio is 0.07). When the degree of serpentinization is low, soils are not very different from those derived from other siliceous rocks (Table 5).

3.1.4. Soil carbon versus biomass carbon in forest systems

From the third National Forest Inventory (NFI3), the total over bark wood volume in forests of pSpain was estimated to be around 907 Mm^3 (about 412.3 Mt C), 75% aboveground and 25% belowground (MAGRAMA, 2010, 2012). The mean density for the whole territory was 23 t C ha^{-1}, with important differences between arid and semi-arid Mediterranean zones (8–25 t C ha^{-1}) and temperate humid zones (45–65 t C ha^{-1}). The mean data for the different ACs were compared with the mean SOC stocks in forest soils (Fig. 12). In all regions the soil is a much more important sink than the forest biomass (≥ 2 times higher in the upper 30 cm) and the differences are multiplied by 1.3 to 1.6 times when the upper 50 cm layer is considered (C_S/C_B ratio in Fig. 12).

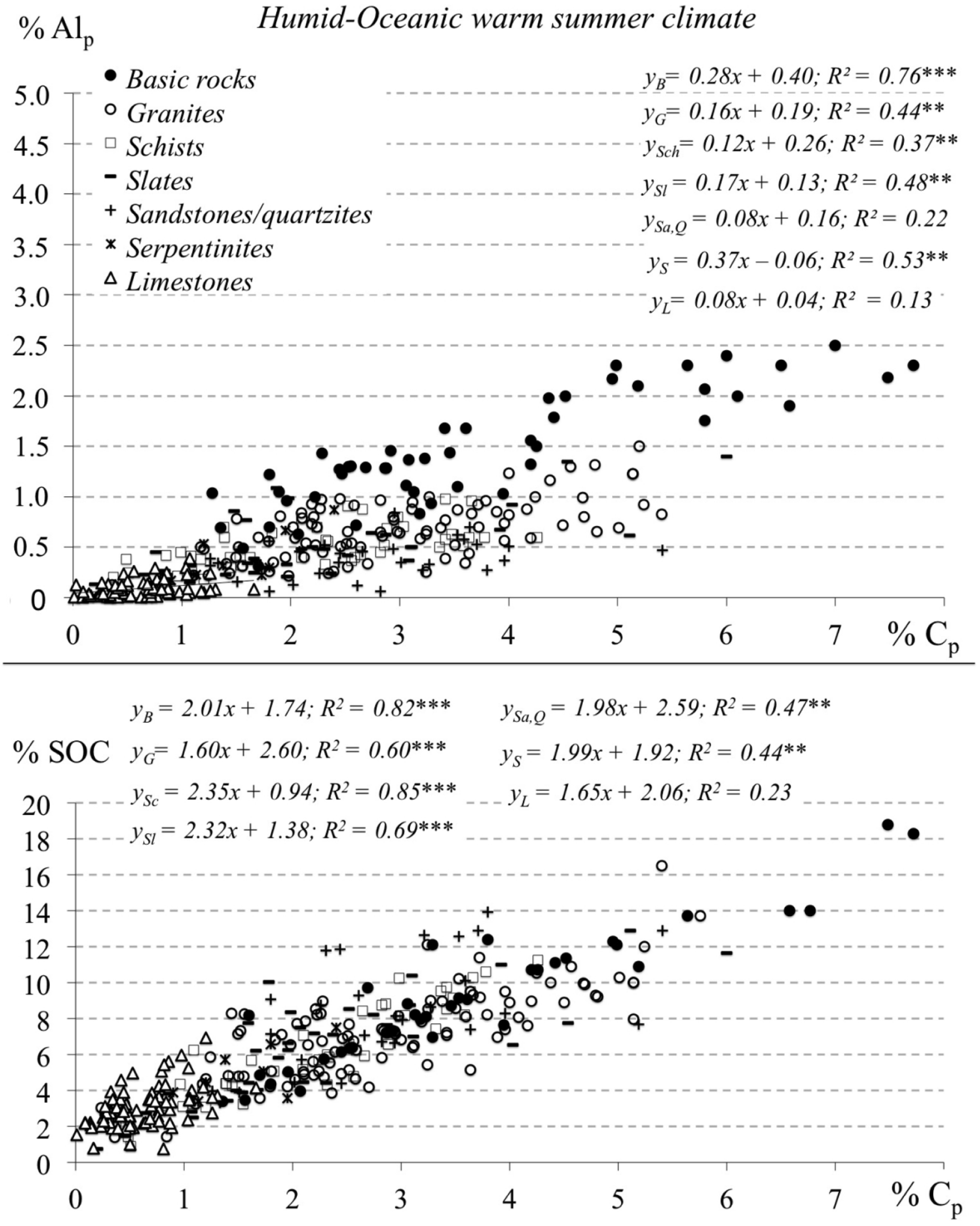

Fig. 11. Relationship between Al and C extracted by Na-pyrophosphate (Al_p, C_p) and between C_p and total SOC content (0–30 cm) in forest soils developed from different parent materials, in temperate humid-oceanic zones of northern pSpain (** and ***: significant at $P < 0.01$ and $P < 0.001$, respectively).

3.2. Modeling and digital mapping of SOC concentration and stock

3.2.1. Distance correlation

Distance correlation analysis between the variables considered in the study is shown in Table 6. The results evidence the logical independence between some of them, such as MAP, MAT, lithology, elevation and land use (even between MAT and elevation), as well as the relationship between soil properties and the environmental variables. Thus, soil pH is significantly related to lithology (DC = 0.44) and MAP (DC = 0.20), and SOC is related to MAP, land use and lithology (DC are 0.14, 0.14 and 0.11, respectively). And both, SOC and pH, are related to each other (DC = 0.29). Likewise, the relationship between soil type and some formation factors (MAP, lithology and land use) is recognizable. Only MAT and elevation do not seem to be related to the rest of the variables, and particularly with the SOC content (DC is 0.02 and 0.01, respectively). This last comment is extensible to the variables Tcm and Thm (not shown in Table 6).

3.2.2. RF model performance and validation of the prediction

RF was applied using all the variables in Table 6, although lithology and soil type were considered in combination (Li/RSG) since better results were observed than when they were utilized independently. The model performance, on the basis of the ME, RMSE and R^2 indices, is shown in Table 7. The results obtained by internal validation (in all or in 90% of the samples) were similar to those assessed by cross-validation. The quality of the fit (R^2) is higher in the 0–30 cm than at 30–50 cm depth, capturing around 61% and 37% of the variance, respectively. The ME is also lower in topsoil and in both layers a certain

Table 5

Selective extractions of C, Al and Fe (and other parameters) in forest soils of temperate humid-oceanic regions. C_t: total organic C; C_r: recalcitrant-C (non-oxidizable organic C); C_p, Al_p and Fe_p: C, Al and Fe extracted by Na-pyrophosphate; Al_{ox}, Fe_{ox}: Al and Fe extracted by acid ammonium oxalate; Al_{DC}, Fe_{DC}: Al and Fe extracted by dithionite-citrate-bicarbonate; BD: bulk density; PO_4^{3-} ret.: phosphate retention (data for 361 soil samples) (source: Calvo de Anta, 2000–2004; Verde, 2009).

														Andic properties			
	Soil prototype		pH	C_t	C_r	C_p	C_p/C_t	Al_{DC}	Al_{ox}	Al_p	Fe_{DC}	Fe_p	Al_p/C_p	Al_{ox} + ½ Fe_{ox}	BD	PO_4^{3-}ret.	pH
Lithology	(RSGs)			%	%	%		%	%	%	%	%	(molar)	%	(kg dm^{-3})	(%)	(NaF)
Mafic	*aa Andosol,*	M	5.2	9.5	1.8	3.7	0.42	2.3	1.7	1.3	3.9	0.7	0.16	2.3	0.8	91	10.8
rocks[a]	*an Umbrisol*	SD	0.4	4.2	0.08	1.7	0.12	0.9	0.8	0.6	1.9	0.4	0.04	0.8	0.1	8	1.0
Granites	*ha/le Umbrisol*	M	4.7	6.1	0.4	2.9	0.40	0.9	0.8	0.7	1.0	0.3	0.12	1.1	0.9	86	9.4
		SD	0.4	2.9	0.06)	1.2	0.13	0.4	0.4	0.4	0.5	0.3	0.07	0.5	0.2	8	1.3
Schists	*cm Umbrisol,*	M	4.9	5.5	0.6	2.3	0.37	1.0	0.9	0.7	2.0	0.5	0.13	1.3	0.9	85	9.6
	ha Umbrisol	SD	0.5	2.5	0.05	1.1	0.08	0.6	0.5	0.4	1.0	0.3	0.10	0.6	0.2)	6	1.2
Slates	*le Umbrisol*	M	4.5	6.1	0.6	2.1	0.34	0.6	0.5	0.5	2.4	0.6	0.12	0.8	0.9	83	8.3
		SD	0.5	3.1	0.05	1.3	0.09	0.4	0.3	0.3	1.1	0.4	0.07	0.4	0.2	5	1.0
Sandstones	*le Umbrisol,*	M	4.3	6.0	0.5	2.7	0.35	0.5	0.4	0.4	2.1	0.5	0.09	0.7	0.9	83	8.3
	um Podzol	SD	0.4	3.0	0.07	1.3	0.11	0.2	0.3	0.2	1.1	0.3	0.04	0.4	0.1	6	0.8
Serpentinites	*le Phaeozem,*	M[b]	6.3	3.2	0.6	1.1	0.27	0.4	0.3	0.2	8.6	0.4	0.07	0.8	1.0	56	8.3
	ha/le Umbrisol & eu/dy Cambisol	M[c]	5.5	5.6	0.5	1.8	0.34	1.1	0.6	0.4	2.4	0.9	0.15	2.0	0.9	80	9.1
Calcareous	*dy/le Cambisol,*	M	6.0	3.2	0.09	0.6	0.19	0.2	0.2	0.1	2.5	0.2	0.05	0.4	1.0	60	8.1
rocks	*rz Phaeozem*	SD	1.2	1.4	0.006	0.3	0.10	0.09	0.1	0.05	1.8	0.1	0.05	0.2	0.2	11	0.3

[a] Gabbros and amphibolites.

[b,c] Materials with high and low degree of serpentinization, respectively.

bias of the prediction -by underestimation- is recognized. However, the model accuracy expressed by the RMSE is poorer in surface than in depth (1.48 and 1.07% C, respectively), which could be related to the higher SOC contents in the first case.

3.2.3. Relative importance of predictor variables

The importance of the variables is assessed by applying two measures, the Mean Decrease Accuracy (% IncMSE), difference between the MSE_{OOB} of each tree computed before and after permuting a variable, and the Mean Decrease Gini (IncNodePurity), that calculates each feature importance as the sum over the number of splits (across all tress) that include the feature, proportionally to the number of samples it splits. Taking into account the first method, the most important predictor in upper layer is MAP, whose removal or permutation generates an increase in the MSE > 200% on an average, followed by far the rest of the variables, and MAT is the worst predictor (Fig. 13). Soil pH is a property dependent on other variables and as a simple parameter has a lower (or complementary) predictive power. The second method allows to establish the importance of the predictors in relation to the increase in node purity (IncNodePurity), thus, it selects the split that has a high inter node variance and a small intra node variance. Again MAP and MAT are the best and worst predictor, respectively. In this case, pH appears as an important variable in the reduction of the impurities of the nodes, when it is used for its division. In the subsurface layer the results are not very different, although the range of variation of both indices is much narrower (Fig. 13).

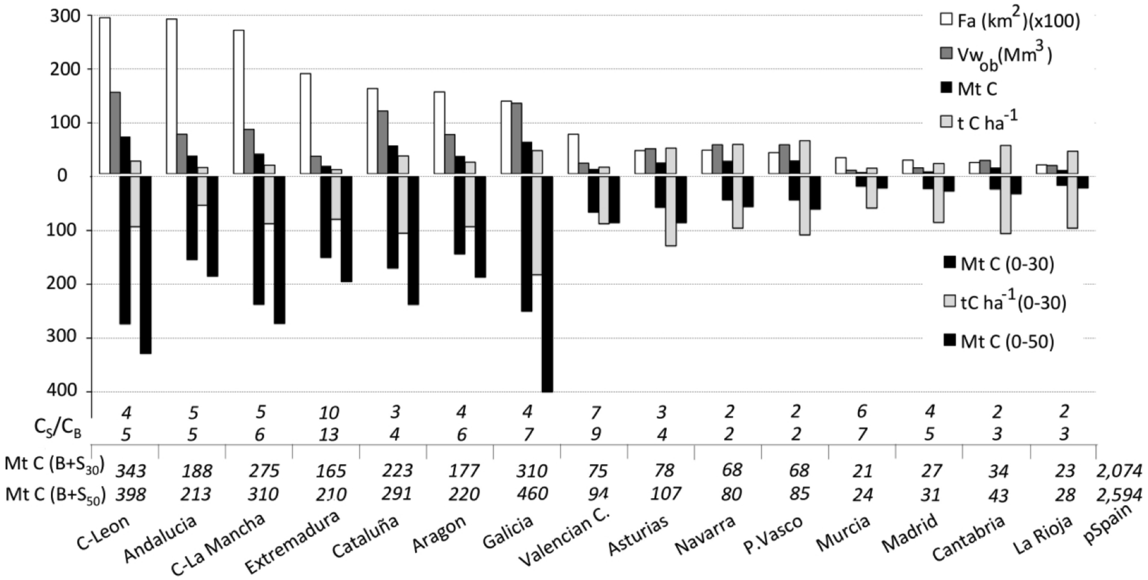

Fig. 12. Soil carbon *versus* biomass carbon in forest systems of the autonomous communities of pSpain. In Y axis: upwards- a single scale is used to represent the forest area (Fa) in each region (km^2 = value × 100), volume of wood over bark (Vw_{ob}) (million m^3) in live biomass, aboveground and belowground, and average C stocks (million tons and tons ha^{-1}) (source: MAGRAMA, 2010, 2012); downwards- average SOC stocks in forest soils (0–30 and 0–50 cm) and mean SOC density in the upper 30 cm. In X axis: soil-C/biomass-C ratio (C_S/C_B) considering the mean SOC stock for 0–30 and 0–50 cm (upper and lower line, respectively), and total organic C in biomass and soil (0–30 and 0–50 cm) (B + S_{30} and B + S_{50}, respectively) (mean values) (* in Andalucia, the data for soil and biomass in olive groves have been included).

Table 6
Distance Correlation matrix between different variables used in this study, SOC content (0–30 cm), land use (LU), elevation (E), mean annual precipitation (MAP), mean annual temperature (MAT), lithology (Li), Reference Soil Groups (RSG), and soil pH.

	SOC	LU	E	MAP	MAT	Li	RSG	pH
SOC	1.00	0.14	0.01	0.14	0.02	0.11	0.27	0.29
LU	0.14	1.00	0.03	0.06	0.03	0.08	0.13	0.13
E	0.01	0.03	1.00	0.06	0.10	0.01	0.04	0.02
MAP	0.14	0.06	0.06	1.00	0.03	0.07	0.16	0.20
MAT	0.02	0.03	0.10	0.03	1.00	0.01	0.02	0.02
Li	0.11	0.08	0.01	0.07	0.01	1.00	0.14	0.44
RSG	0.27	0.13	0.04	0.16	0.02	0.14	1.00	0.24
pH	0.29	0.13	0.02	0.20	0.02	0.44	0.24	1.00

3.2.4. *Uncertainty of the prediction*

Uncertainty of prediction was estimated based on Malone et al. (2011), who use an empirical method where model output uncertainties are expressed as intervals of spatial distribution of the prediction errors, estimated for sample groups that share similar environmental attributes. In our study, model uncertainty was estimated from the error intervals -computed in the internal validation set (OOB samples)- for groups of samples within the same range of observed SOC stocks. The results indicate a positive correlation between the observed values and the estimated errors (Fig. 14). In the surface layer (0–30 cm) a slight underestimation is recognized in soils with low reserves of carbon (up to 4 kg m^{-2}), with a median error of around 2 kg m^{-2} and a range of uncertainty on this error of 5 kg m^{-2}. In classes with 4–12 kg m^{-2} (almost 60% of the samples) the error/observed value ratio is close to 0.0 and the error range is between 5 and 15 kg m^{-2}. This ratio increases as the observed value increases, from 0.2 to 0.6 for 12 and >26 kg m^{-2}, respectively, and in all classes the uncertainty range is same (17 kg m^{-2}). The reliability of the model is low in the last three classes and especially in soils with >26 kg C m^{-2}. Peaty soils have very little representation in the database (0.5%) and in the land cover they map occupy $<0.002\%$ of the total (no urban) area of pSpain. In these zones the predictive model was not used and the average values observed (for the upper 30 or 50 cm) in the organic soils closest to the area were applied. In the subsurface layer (30–50 cm) the uncertainty distribution show a pattern of behavior similar to that of the surface. Most of the samples (87%) are in the three first classes (observed SOC stocks are less than or equal to 6 kg m^{-2}) and the estimated mean prediction error is close to zero, with a range between 2 and 6 kg m^{-2} (Fig. 14). Uncertainty increases in the remaining classes, always with a lower error range than on the surface.

Table 7
Performance and validation of model for the SOC content prediction: Internal and cross-validation. Mean error (ME) (g/100 g), root mean square error (RMSE) (g/100 g) and coefficient of determination (R^2) for the Out-Of-Bag samples and from cross-validation test.

	Validation indices	Depth (cm)	
		0–30	30–50
All data	ME_{OOB}	−0.008	−0.013
	$RMSE_{OOB}$	1.48	1.08
	R^2_{OOB}	0.60	0.37
Training subset (90% of the samples)	ME_{OOB}	−0.008	−0.015
	$RMSE_{OOB}$	1.49	1.08
	R^2_{OOB}	0.60	0.37
Validation subset (10% of the samples)	ME_{test}	−0.005	−0.023
	$RMSE_{test}$	1.48	1.07
	R^2_{test}	0.61	0.37

3.2.5. *Spatial distribution of SOC content and density*

The spatial distribution of predicted SOC percentage and stock is shown in Figs. 15 and 16. In most of the territory ($>50\%$) the mean density in the upper 30 cm is between 40 and 80 t C ha^{-1} (approximately 1.0 and 2.5%, respectively); in 23% of the area it is between 80 and 140 t C ha^{-1} (about 2.5–5.0%); and it is <40 or >140 t C ha^{-1}, in 15% and 9% of the surface, respectively. For a total area of around 43.6 Ma (except urban soils, marshes, wetlands and others) the SOC stock is estimated in an average of 3334.63 Mt, with a CI95 ranging between 2903.05 and 3756.92 Mt. Most of C is stored under forests (48% of the total C in 38% of the total area), followed by scrublands (24% and 20%, respectively), cropland (18% and 32%), meadows & pastures (9% and 8%) and sparsely vegetated eroded areas (1% and 3%). The total predicted for this upper layer is similar to that calculated by weighting the average values observed for different land use at the regional level (Table 4). They are values of the same order as those estimated by Hiederer (2010) for topsoil in Spain (3.50 Pg C), and higher than those assessed by Rodriguez Martín et al. (2016) (2.82 Pg C for the peninsular territory and Balearic Islands). The SOC density in the subsurface layer (30–50 cm) is <40 t ha^{-1} in 88% of the area, and the total reserve is 854.01 Mt (432.76–1275.23 Mt). According to these results, the total stock estimated for pSpain (0–50 cm) is 4188.63 Mt C, ranging between 3335.75 and 5032.22 Mt (about 20% of uncertainty).

The complexity of RF model does not allow an easy interpretation, since it acts as a “black box” and it is not possible to establish the joint structure of all the trees in the forest and recognize the functional model that incorporates all the variables. As a simple approximation to the relationships among SOC stock and the covariates, statistics for sample sets grouped considering the most important variables predicted by the model have been calculated. The mean range for these groups is shown in Fig. 17. In cropland, SOC stock is between 10 and 40 t ha^{-1} (0–30 cm) in regions with <700 mm of precipitation, and between 30 and 60 t ha^{-1} in the rest of the territory (around 150 t ha^{-1} in acid soils that traditionally receive organic amendments). In forest soils, the mean range for different groups is as follow: 40–70 t ha^{-1} in arid to desert zones (MAP $<$ 400 mm); 70–100 t ha^{-1}, in semi-arid and dry subhumid regiones (MAP 400–700 mm) with neutral or weakly acid soils (*eu/dy Regosols* and *Cambisols*); 90–110 t ha^{-1}, in the same climatic region and soils with high levels of active lime (pH $\approx$ 8.3) (*Calcisols* and *ca Regosols*); 100–120 t ha^{-1}, in subhumid and humid zones (900–1700 mm) and decarbonated calcareous materials (*dy/eu Cambisols* and *rz Phaeozems*) (pH $<$ 4.0–7.0) and in neutral soils from serpentinites (*ha/le Phaeozems*); 120–200 t ha^{-1}, in humid zones and soils developed from granite and other quartz-rich rocks (*ha Umbrisols*); and 150–240 t ha^{-1}, in soils derived from mafic rocks (*an Umbrisols and aa Andosols*).

4. Conclusions

In this study the relationships between soil organic carbon and some environmental factors in peninsular Spain are analyzed from a database (12,724 profiles) covering different climate, land use, elevation, parent material, soil type and soil pH. Random Forest regression was used to predict the spatial distribution of SOC content and stock at 0–30 and 30–50 cm. The main conclusions were as follows:

- The large spatial heterogeneity of environmental factors in Spain leads to very large variability in SOC stocks. In the upper layer, the lowest mean values (10–30 t C ha^{-1}) (0–30 cm) are recorded in agricultural soils of arid regions; and the highest (170–240 t C ha^{-1}), in acid forest soils of humid regions. In subsurface layers there are a mean decrease of about 3 times, in the first environments, and of about 2 times, in the second.
- The presence/absence of active lime and active Al are the key factors of SOC content variability in semi-arid and humid regions, respectively. In semi-arid zones, calcaric soils (pH $\approx$ 8.3) have higher

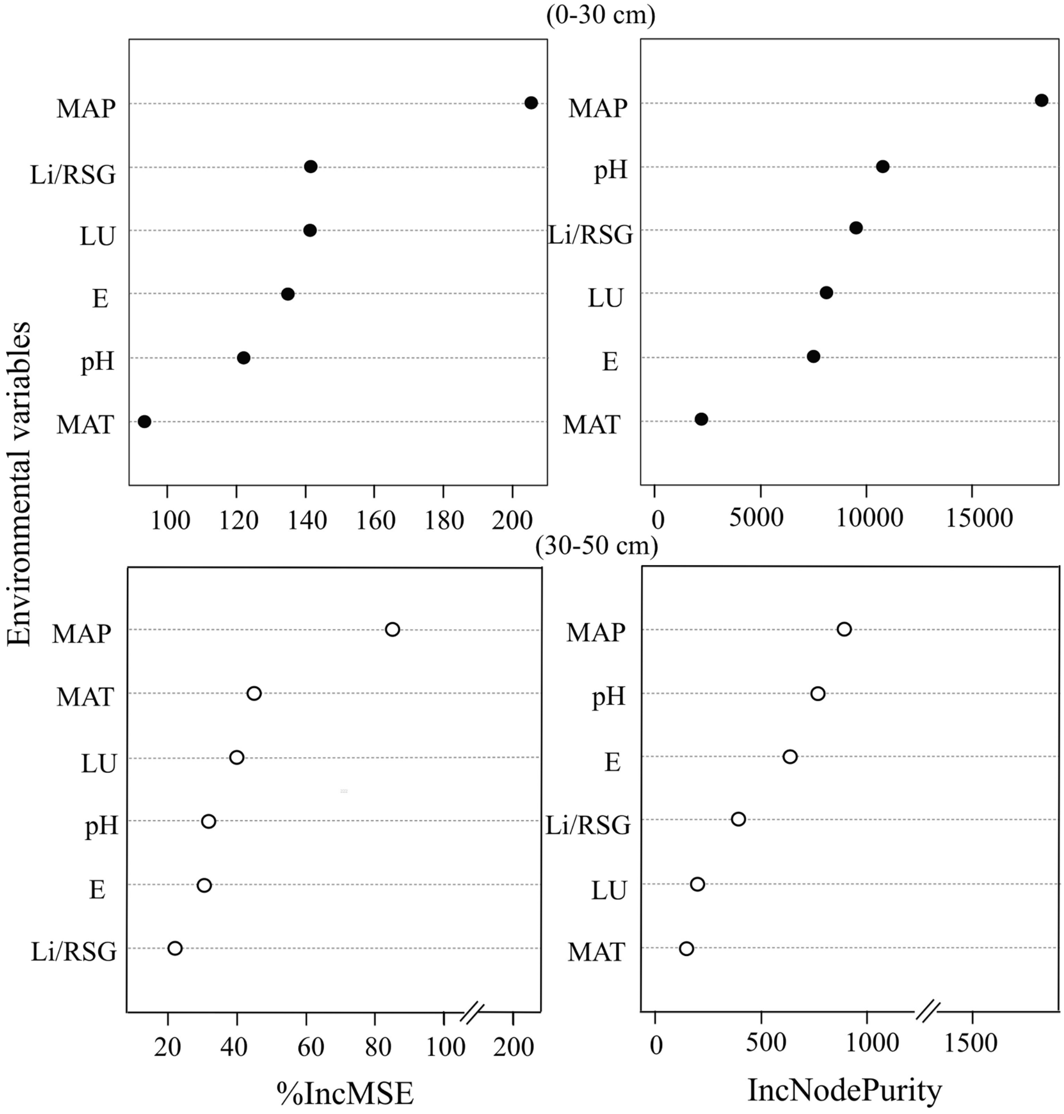

Fig. 13. Importance of the variables used in Random Forest regression modeling. %IncMSE: Mean Decrease Accuracy (percent increase in the mean squared error of each tree after permuting a variable); IncNodePurity: Mean Decrease Gini (total increase in node purities from splitting on the variable); MAP: mean annual precipitation; MAT: mean annual temperature; Li: lithology; RSG: Reference Soil Groups; LU: land use; E: elevation.

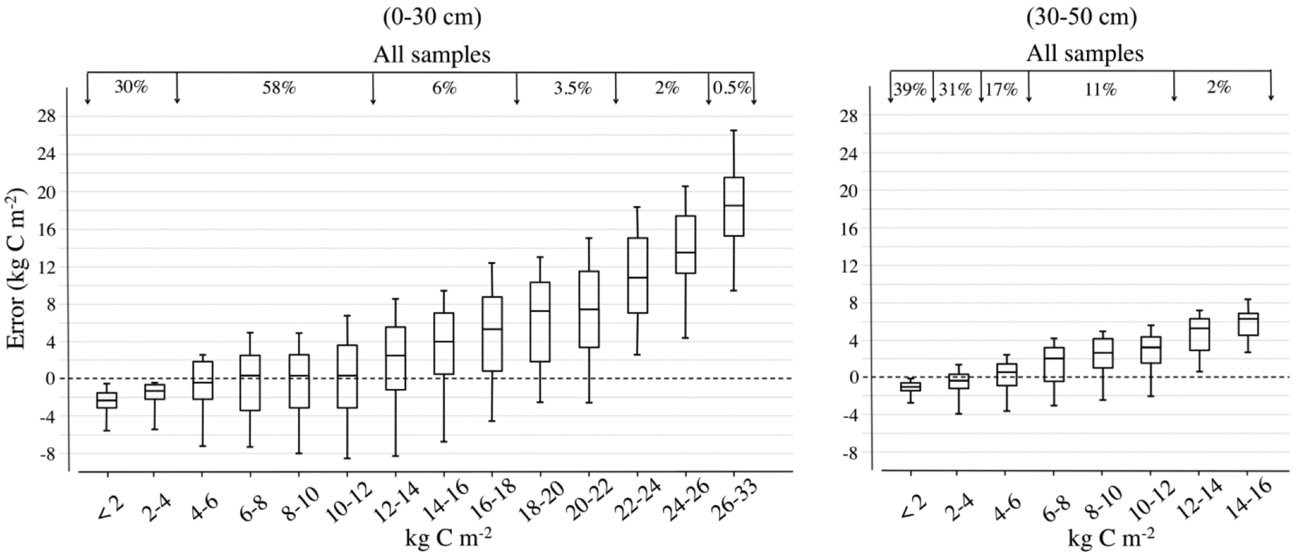

Fig. 14. Uncertainty of the model as a function of the observed SOC stocks (0–30 and 30–50 cm). Boxplots and confidence intervals (CI 95%) of estimated errors for soil groups with different range of observed SOC stock (computed from the internal validation test, in Out-Of-Bag samples).

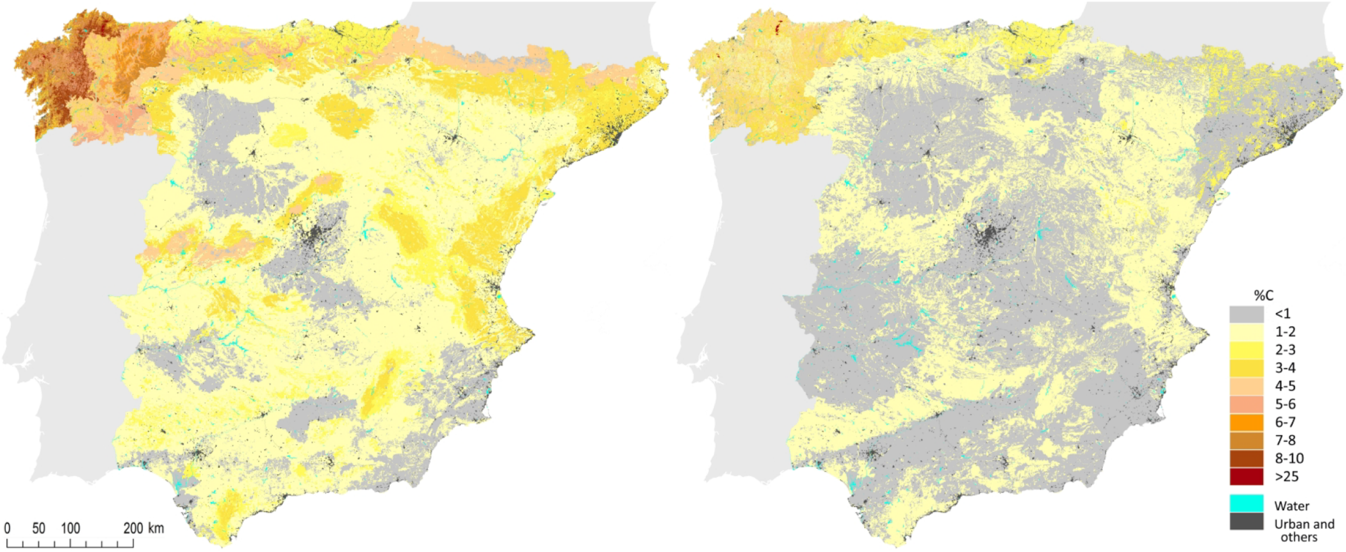

Fig.15. Predicted spatial distribution of SOC content at 0–30 cm (left) and 30–50 cm depth (right).

contents of SOC than neutral to weakly acidic soils from siliceous materials. In humid regions, acidic soils with high levels of active Al (pH 4.5–5.5) have markedly higher reserves (up to >3 times) than neutral and acidic soils from decarbonated calcareous materials (pH $<$ 4.0–7.5) or from serpentinized ultrabasic rocks (pH 5.5–6.5).

- SOC content varies according to the following sequence of soil types (from arid to humid regions): *eu/dy Regosols* (or *Cambisols)* $<$ *Calcisols* and *ca Regosols (or Cambisols)* $<$ *rz/ha Phaeozems* $<$ *ha Umbrisols* $<$ *an Umbrisols* $<$ *aa Andosols* (*Histosols* are not considered).
- In all regions, the forest soil is a much more important carbon sink than live forest biomass (2–4 times higher in the upper 30 cm and 3–6 times greater in the upper 50 cm).
- RF model performance (evaluated by internal validation and cross validation) was considered acceptable. In topsoil, ME, RMSE and R^2 were −0.007% C, 1.48% C and 0.61, respectively; and, in 30–50 cm layer, these indices were −0.020, 1.07 and 0.37, respectively.
- From modeling, mean annual precipitation was the most important predictor variable, followed by lithology/soil type, land use and soil pH.
- The estimated total SOC stock at 0–30 cm soil depth was about 3.33 Pg and that for 0–50 cm was 4.19 Pg, with a 95% confidence interval ranging between 3.33 and 5.03 Pg.
- Most of SOC is stored under forests, followed by scrubland and cropland. Forest soils cover 38% of the total area and contain 48% of the total SOC; scrubland occupy 24% of the surface and store 20% of the SOC; and cropland occupies 32% and stores 18%.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

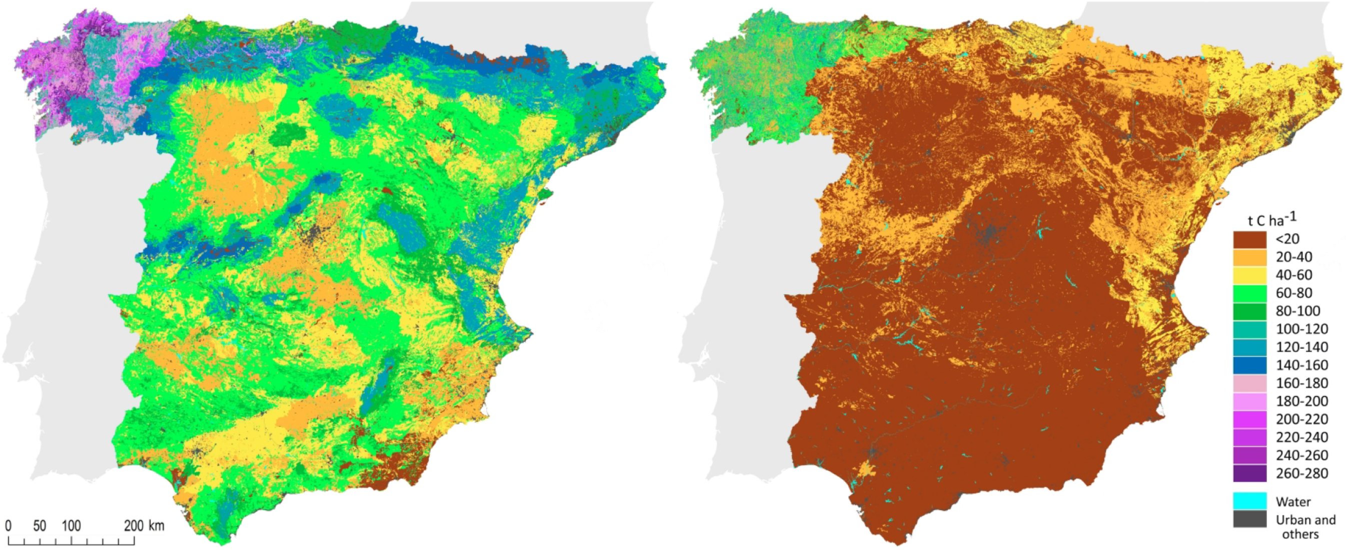

Fig.16. Predicted spatial distribution of SOC stock at 0–30 cm (left) and 30–50 cm depth (right).

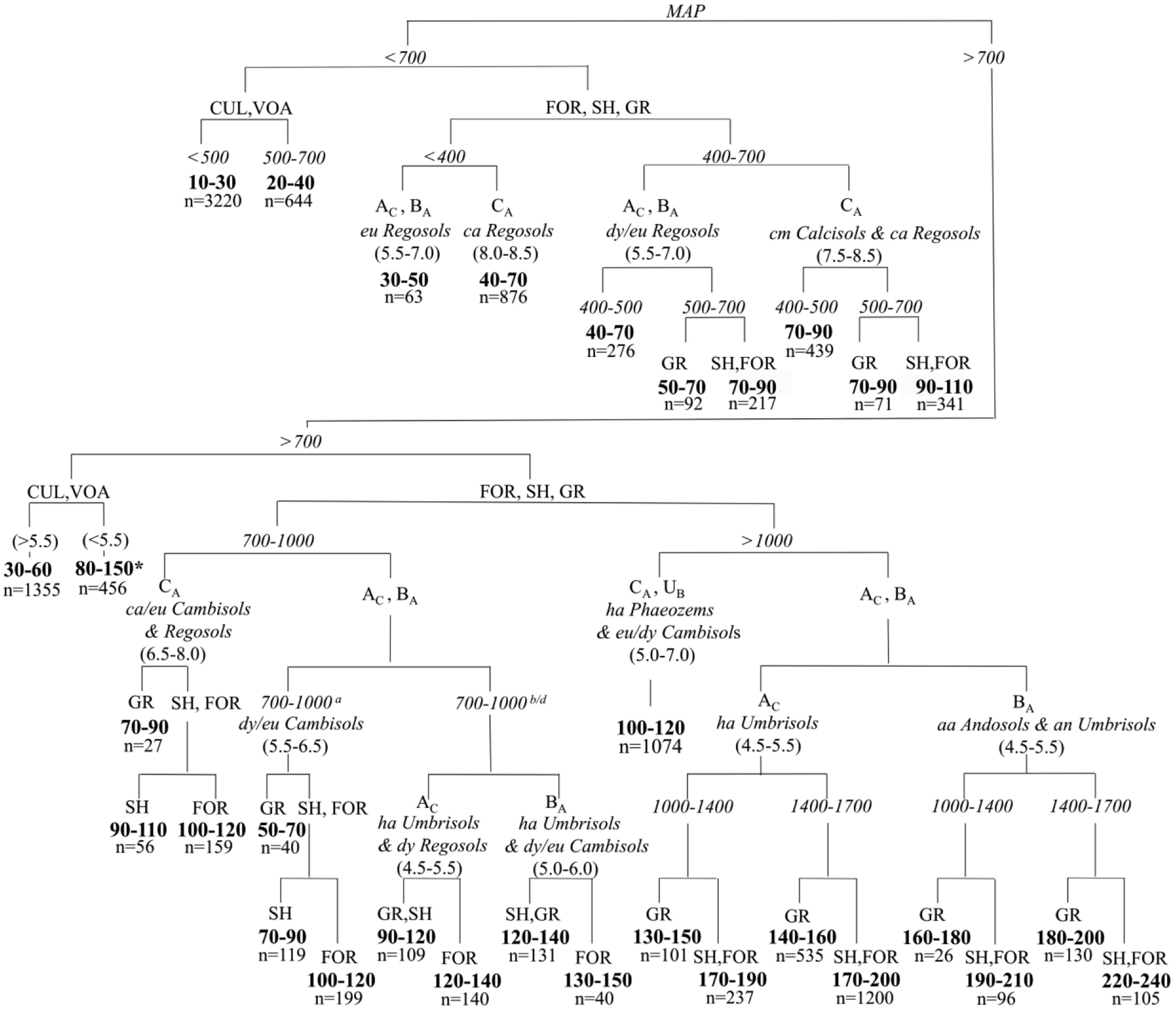

Fig. 17. Average range of observed SOC density (t C ha^{-1}) (in bold) in soils of pSpain grouped according the importance of the variables predicted by Random Forest (0–30 cm) (in italics: mm of mean annual precipitation- MAP; in brackets: soil pH ranges; ^{a}: hot summer; ^{b}: warm summer; ^{d}: cold winter -see Fig. 3-; A_C: acid rocks; B_A: basic rocks; C_A: calcareous materials; SH: Shrubs; FOR: forests; GR: grasslands; CUL: arable crops; VOA: woody crops; n: number of samples) (*: agricultural areas which traditional management includes organic amendments, manure and others).

Acknowledgments

We thank the AEMET for facilitating the digital climatic coverage of Spain and the Galician government for providing us with the basis for Galicia of the Land Use Information System (SIOSE). We acknowledge the organizations that have provided access to information for this study (IGN, IGME and INIA) and to the researchers who have collaborated in the search for data, in particular F. Arricibita, P. Bescansa and J.Virto.

Funding

This work was supported by the Ministry of Science and Innovation, Spain (CGL2009, 13857: 2010-2014); and the Consellería de Educación-Xunta de Galicia, Spain (2014-GRC GI-124, 2014-2017).

References

Adhikari, K., Hartemink, A.E., Minasny, B., Bou Kheir, R., Greve, M.B., Greve, M.H., 2014. Digital mapping of soil organic carbon contents and stocks in Denmark. PLoS ONE 9.

AEMET, 2011. Iberian Climate Atlas. Air temperature and Precipitation (1971-2000). Agencia Estatal de Meteorología. Ministerio de Medido Ambiente y Medio Rural y Marino, Madrid.

Apka, S., Odeh, I., Bishop, T., Hartemink, A., Amapu, I., 2016. Total soil organic carbon and carbon sequestration potential in Nigeria. Geoderma 271, 202–215.

Arun, K., Langmead, C.J., 2005. Structure based chemical shift prediction using random forest non-linear regression. CarnegieMellon University, Pittsburgh. CMU-CS-05-163. http://reports-archive.adm.cs.cmu.edu/anon/2005/abstracts/05-163.

Arrouays, D., Deslais, W., Badeau, V., 2001. The carbon content of topsoil and its geographical distribution in France. Soil Use and Manage. 17, 7–11.

Bascomb, C.L., 1968. Distribution of pyrophosphate-extractable iron and organic carbon in soils of various groups. Eur. J. Soil Sci. 19 (2), 251–268.

Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci. 47 (2), 151–163.

Batjes, N.H., 2016. Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. Geoderma 269, 61–68.

Berg, B., Berg, M.P., Bottner, P., Box, E., Breymeyer, A., Calvo de Anta, R., Couteaux, M., Escudero, A., Gallardo, A., Kratz, W., Madeira, M., Mälkönen, E., McClaugherty, C., Meentemeyer, V., Munoz, F., Piussi, P., Remacle, J., de Santo, A.V., 1993. Litter mass loss rates in pine forests of Europe and Eastern United States: some relationships with climate and litter quality. Biogeochemistry 20 (3), 127–159.

Blakemore, L.C., Searle, P.L., Daly, B.K., 1987. Soil Bureau analytical methods. A method for chemical analysis of soils. NZ Soil Bureau. Sci. Report 80, DSIRO.

Bolliger, J., Hagedorn, F., Leifeld, J., 2008. Effects of land-use change on carbon stocks in Switzerland. Ecosystems 11 (6), 895–907.

Bradley, R.I., Milne, R., Bell, J., Lilly, A., Jordan, C., Higgins, A., 2005. A soil carbon and land use database for the United Kingdom. Soil Use Manage. 21, 363–369.

Breiman, L., 2001. Random forests. Machine Learning 45, 5–32.

de Brogniez, D., 2015 Quantification of Topsoil Organic Carbon in Europe: A Statistical Approach Combining Harmonized Soil and Land Cover Point Data with Spatial Layers (PhD thesis). Université Catholique de Louvain. Belgium http://hdl.handle.net/2078.1/158816.

de Brogniez, D., Ballabio, C., Stevens, A., Jones, R.J.A., Montanarella, L., van Wesemael, B., 2015. A map of the topsoil organic carbon content of Europe generated by a generalized additive model. Eur. J. Soil Sci. 66 (1), 121–134.

Calvo de Anta, R., Álvarez, E., 1992. Aluminiun activity in soil solution and mineral stability in soils from Galicia (N.W. Spain). Clay Miner. 27, 325–330.

Calvo de Anta, R., Macías, F., 1993. Rôle des processus pédogenétiques dans le maintien de la qualité des eaux superficielles en zones temperées humides. Exemple des sols acides sur roches granitiques et schisteuses de la Galice (NW Espagne). Comptes Rendus de l'Académie des Sciences 316 (II), 799–805.

Calvo de Anta, R. (coord.), 2000-2004. Mapa de Suelos de Galicia 1:50.000. SIAM. Xunta de Galicia. https://siam.xunta.gal/mapa-de-solos.

Calvo de Anta, R., Luís, E., Matilla, N., Casás, F., Macías, F., Camps, M., Vázquez, N., Galiñanes, J.M., 2014. Mapa digital de carbono en suelos del norte de España (Galicia y Cornisa Cantábrica). In: Macías, F., Díaz-Raviña, M., Barral, M. (Eds.), Retos y oportunidades en la ciencia del suelo. VI Congreso Ibérico de la Ciencia del Suelo, Santiago de Compostela, pp. 65–68.

Calvo de Anta, R., Luis, E., Casás, F., Galiñames, M., Matilla, N., Macías, F., Camps, A., Vazquez, N., 2015. Soil organic carbon in northern Spain (Galicia, Asturias, Cantabria and País Vasco). Span. J. Soil Sci. 5 (1), 41–53.

Eswaran, H., Van Den Berg, E., Reich, P., 1993. Organic carbon in soils of the world. Soil Sci. Soc. Am. J. 57, 192–194.

Fieldes, M., Perrott, K.W., 1966. The nature of allophane soils: 3. Rapid field and laboratory test for allophane. N. Z. J. Sci. 9, 623–629.

García-Rodeja, E., Silva, B., Macías, F., 1987. Andosols developed from non-volcanic materials in Galicia, NW Spain. Eur. J. Soil Sci. 38, 573–591.

Gingrich, S., Erb, K.H., Krausmann, F., Gaube, V., Haberl, H., 2007. Long-term dynamics of terrestrial carbon stocks in Austria: a comprehensive assessment of the time period from 1830 to 2000. Reg. Environ. Change 7 (1), 37–47.

Grimm, R., Behrens, T., Märker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island-Digital soil mapping using Random Forest analysis. Geoderma 146, 102–113.

Hiederer, R., 2010. Organic carbon in topsoils (0-30) per country, http://eusoils.jrc.ec.europa.eu/ESDB Archive/octop/Resources/OC Per Country.

Hiederer, R., Köchy, M., 2011. Global Soil Organic Carbon Estimates and the Harmonized World Soil Database. EUR 25225 EN. Publications Office of the EU, Luxembourg.

Higashi, T., De Coninck, F., Gelaude, F., 1981. Characterization of some spodic horizons of the campine (Belgium) with dithionite-citrate, pyrophosphate and sodium hydroxide-tetraborate. Geoderma 25, 131–142.

ICONA, 1986-1992. LUCDEME project. Mapas de Suelos de España (1:100,000). Ministerio de Agricultura, Pesca y Alimentación, Madrid.

IGME, 2015. Mapa Geológico de España 1:1M. http://mapas.igme.es/gis/rest/services/Cartografia_Geologica/IGME_EP_Geologico_1M_2015/MapServer.

IGN, 2011. Information System of Land Occupation of Spain. SIOSE. Centro Nacional de Información Geográfica, España.

IGN, 2012. CORINE Land Cover 2012. Centro Nacional de Información Geográfica, España.

Iñiguez, J. (coord.), 1982-1992. Mapa de Suelos de Navarra 1:50,000. Publ. Principe de Viana. Suplemento de Ciencias. Pamplona.

IPCC, 1990. Climate Change. Cambridge University Press, Cambridge.

IUSS Working Group WRB, 2015. World Reference Base for Soil Resources 2014 (updated in 2015). International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome.

Jobbágy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. 10, 423–436.

Jones, R.J.A., Hiederer, R., Rusco, E., Montanarella, L., 2005. Estimating organic carbon in the soils of Europe for policy support. Eur. J. Soil Sci. 56, 655–671.

Krogh, L., Noergaard, A., Hermansen, M., Greve, M.H., Balstroem, T., Breuning-Madsen, H., 2003. Preliminary estimates of contemporary soil organic carbon stocks in Denmark using multiple datasets and four scaling-up methods. Agric. Ecosyst. Environ. 96, 19–28.

Kuikman, P., de Groot, W., Hendriks, R., Verhagen, J., de Vries, F., 2003. Stocks of C in soils and emissions of CO2 from agricultural soils in the Netherlands. Alterra-rapport 561. Alterra, Green World Research, Wageningen.

Liaw, A., Wiener, M., 2002. Classification and regression by Random Forest. R News 2 (3), 18–22.

Liaw, A., Wiener, M., 2018. Classification and regression by Random Forest R package Ver.: 4.6-14 https://cran.r-project.org/package = randomForest.

Macías, F., Calvo de Anta, R., 1992. Caractérisation pédogéochimique des sols de la Galice (NW Espagne) en relation avec la diversification lithologique. Mise en évidence d'un milieu de transition entre les domaines tempérés et subtropicaux humides. Comptes Rendus de l'Acad. Sci. 315 (II), 1803–1810.

MAGRAMA, 2010. Third National Forest Inventory (NFI3) 2007. Ministerio de Agricultura, Alimentación y Medio Ambiente. Madrid.

MAGRAMA, 2012. Criteria and indicators for sustainable forest Management in Spanish forest. Ministerio de Agricultura, Alimentación y Medio Ambiente, Madrid.

Malone, B.P., McBratney, A.B., Minasny, B., 2011. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. Geoderma 160, 614–626.

Martin, M.P., Wattenbach, M., Smith, P., Meersmans, J., Jolivetl, C., Boulonne, L., Arrouays, D., 2011. Spatial distribution of soil organic carbon stocks in France. Biogeosciences 8, 1053–1065.

Masiello, C.A., Chadwick, O.A., Southon, J., Torn, M.S., Harden, J.W., 2004. Weathering controls on mechanisms of carbon storage in grassland soils. Global Biogeochem. Cycles 18, GB4023. https://doi.org/10.1029/2004GB002219.

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52.

Mehra, O.P., Jackson, M.L., 1960. Iron oxide removal from soils and clay by a dithionite-citrate system buffered with sodium bicarbonate. Clays Clay Miner. 7, 317–327.

Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I., 2013. Digital mapping of soil carbon. Adv. Agron. 118, 1–47.

Montoya, R., López-Arias, M., 1997. La Red Europea de Seguimiento de Daños en los Bosques (Nivel I). España 1987-1996. Publ. del O.A. Parques Nacionales. Ministerio de Medio Ambiente. Madrid.

Morvan, X., Saby, N.P.A., Arrouays, D., Le Bas, C., Jones, R.J.A., Verheijen, F.G.A., Bellamy, P.H., Stephens, M., Kibblewhite, M.G., 2008. Soil monitoring in Europe: a review of existing systems and requirements for harmonisation. Sci. Total Environ. 391 (1), 1–12.

Mulder, V.L., Lacoste, M., Richer-de-Forges, A.C., Martin, M.P., Arrouays, D., 2016. National versus global modelling the 3D distribution of soils organic carbon in mainland France. Geoderma 263, 16–34.

Oades, J.M., 1989. An introduction to organic matter in mineral soils. In: Dixon, J.B., Weed, S.B. (Eds.), Minerals in Soil Environments, 2nd Edition. Soil Sci. Soc. of Am, Madison, WI, pp. 89–159.

Poggio, L., Gimona, A., 2014. National scale 3D modelling of soil organic carbon stocks with uncertainty propagation – an example from Scotland. Geoderma 232–234, 284–299.

Rodriguez Martín, J.A., Álvaro-Fuentes, J., Gonzalo, J., Gil, C., Ramos-Miras, J.J., Grau Corbí, J.M., Boluda, R., 2016. Assessment of the soil organic carbon stock in Spain. Geoderma 264, 117–125.

Scharlemann, J., Hiederer, R., Kapos, V., 2009. Global Map of Terrestrial Soil Organic Carbon Stocks. UNEP-WCMC & EU-JRC, Cambridge, UK.

Scharlemann, J., Tanner, E., Hiederer, R., Kapos, V., 2014. Global soil carbon: understanding and managing the largest terrestrial carbon pool. Carbon Manag. 5 (1), 81–91. https://doi.org/10.4155/cmt.13.77.

Schils, R.L.M., Kuikman, P., Liski, J., van Oijen, M., Smith, P., Webb, J., Alm, J., Somogyi, Z., van den Akker, J., Billett, M., Emmett, B., Evans, C., Lindner, M., Palosuo, T., Bellamy, P., Jandl, R., Hiederer, R., 2008. Service Contract: Review of Existing Information on the Interrelations Between Soil and Climate Change (CLIMSOIL). Final Report, 208 pp. European Commission, Brussels, Belgium. Web address: http://ec.europa.eu/environment/soil/pdf/climsoil.

Schrumpf, M., Kaiser, K., Guggenberger, G., Persson, T., Kögel-Knabner, I., Schulze, E.D., 2013. Storage and stability of organic carbon in soils as related to depth, occlusion within aggregates, and attachment to minerals. Biogeociencias 10, 1675–1691.

Sreenivas, K., Dadhwal, V.K., Kumar, S., Sri Harsha, G., Tarik, M., Sujatha, G., Janaki Rama Suresh, G., Fyzee, M.A., Ravisankar, T., 2016. Digital organic and inorganic carbon mapping of India. Geoderma 269, 160–173.

Stockmann, U., Adams, M.A., Crawford, J.W., Field, D.J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A.B., de Courcelles, V., de Singh, K.R., Wheeler, I., Abbott, L., Angers, D.A., Baldock, J., Bird, M., Brookes, P.C., Chenu, C., Jastrow, J.D., Lal, R., Lehmann, J., O'Donnell, A.G., Parton, W.J., Whitehead, D., Zimmermann, M., 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agric. Ecosyst. Environ. 164, 80–99.

Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P., 2003. Random forest: a classification and regression tool for compound classification and QSAR modeling. J. Chem. Inf. Comput. Sci. 43, 1947–1958.

Székely, G.J., Rizzo, M., Bakirov, N.K., 2007. Measuring and testing dependence by correlation of distances. Ann. Statist. 35 (6), 2769–2794.

Todd-Brown, K.E., Randerson, J.T., Post, W.M., Hoffman, F.M., Tarnocai, C., Schuur, E.A.G., Allison, S.D., 2013. Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. Biogeoscience 10, 1717–1736.

Tóh, G., Jones, A., Montanarella, L. (Eds.), 2013. LUCAS Soil: methodology, data and results. JRC Technical Reports. European Coommission. Report EUR 26102 EN, Luxembourg.

Tomlinson, R.W., 2005. Soil carbon stocks and changes in the Republic of Ireland. J. Environ. Manage. 76, 77–93.

Torn, M.S., Trumbore, S.E., Chadwick, O.A., Vitousek, P.M., Hendricks, D.M., 1997. Mineral control of soil organic carbon storage and turnover. Nature 389 (6647), 170–173.

Vågen, T.G., Winowiecki, L.A., Abegaz, A., Hadgu, K.M., 2013. Landsat-based approaches formapping of land degradation prevalence and soil functional properties in Ethiopia. Remote Sens. Environ. 134, 266–275.

Van Ranst, E., Thomasson, A.J., Daroussin, J., Hollis, J.M., Jones, R.J.A., Jamagne, M., King, D., Vanmechelen, L., 1995. Elaboration of an extended knowledge database to interpret the 1:1,000,000 EU Soil Map for environmental purposes, in: King, D., Jones, R.J.A., Thomasson, A.J. (Eds.), European Land Information Systems for Agro-environmental Monitoring. European Soil Bureau Research Report. Office for Official Publications of the European Communities, Luxembourg, pp. 71–84.

Verde, J.R., 2009. Influencia de los factores de formación y uso del suelo en las propiedades ándicas. Formas y mecanismos de estabilización de carbono en suelos del NO de la península ibérica (PhD thesis). University of Santiago de Compostela, Spain.

Wada, K.J., 1985. The Distinctive Properties of Andosols. Springer, New York.

Wiesmeier, M., Barthold, F., Blank, B., Kögel-Knabner, I., 2011. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. Plant Soil 340, 7–24.