

Determination of minimum number of samples allowing to detect long term soil organic carbon changes in Mediterranean arable lands using paired-sites

Calogero Schillaci (✉ calogero.schillaci@unimi.it)

University of Milan <https://orcid.org/0000-0001-7689-5697>

Sergio Saia

University of Pisa

Aldo Lipani

UCL: University College London

Alessia Perego

University of Milan

Claudio Zaccone

Università degli Studi di Verona

Marco Acutis

University of Milan

Research

Keywords: Arable crops, Paired sites, Power analysis, Semi-arid, SOC dynamics

Posted Date: January 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-150726/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published on June 7th, 2021. See the published version at <https://doi.org/10.1186/s13021-021-00182-7>.

Abstract

Background

Legacy data are unique occasions for estimating soil organic carbon (SOC) concentration changes and spatial variability, but their use can pose limitations due to the sampling schemes adopted and improvements may be needed in the analysis methodologies. When SOC changes is estimated with legacy data, the use of soil samples collected in different plots (i.e., non-aligned data) may lead to biased results. In the present work, N=302 georeferenced soil samples were selected from a regional (Sicily, south of Italy) soil database. An operational sampling approach was developed to spot SOC concentration changes from 1994 to 2017 in the same plots at the 0-30 cm soil depth and tested.

Results

The measurements were conducted after computing the minimum number of samples needed to have a reliable estimate of SOC variation after 23 years. By applying an effect size based methodology, 30 out of 302 sites were resampled in 2017 to achieve a power of 80%, and an $\alpha=0.05$.

A Wilcoxon *test* applied to the variation of SOC from 1994 to 2017 suggested that there was not a statistical difference in SOC concentration after 23 years ($Z = -0.556$; 2-tailed asymptotic significance = 0.578). In particular, only 40% of resampled sites showed a higher SOC concentration than in 2017.

Conclusions

This finding contrasts with a previous SOC concentration increase that was found in 2008 (75.8% increase when estimated as differences of 2 models built with non-aligned data), when compared to 1994 observed data ($Z = -9.119$; 2-tailed asymptotic significance < 0.001).

Such a result implies that the use of legacy data to estimate SOC concentration dynamics requires soil resampling in the same locations to overcome the stochastic model errors. Further experiment is needed to identify the percentage of the sites to resample in order to align two legacy datasets in the same area.

1 Introduction

Soil organic carbon (SOC) is a main contributor to fertility in agricultural soils, which also guarantee water accumulation and biodiversity (Lal, 2008). Baseline SOC estimates and maps are generally built on legacy data (Odeh et al., 2012), whereas any new collection of soil samples in the same legacy locations are often scarce. Up to date SOC data and assessments are on the global agenda (Vermeulen et al., 2019) and are necessary to evaluate many ecosystem characteristics such as resilience, productivity, ability of soil to provide a wide range of ecosystem services (Williams et al., 2020), and to gain precious insight into policy measures for soil preservation (Chenu et al., 2019).

Prior to a SOC assessment, a sampling campaign would be needed, and the number of samples would affect obtained results. Many sampling size determination strategies have been proposed in the last decades to spot SOC changes (Biswas and Zhang, 2018; Köhl et al., 2011; Lark, 2009; Parras-Alcántara et al., 2015; Wadoux and Brus, 2020). Most of these have been suggested for new data acquisition and to overcome the problems of difficult terrain. There is now a strong and growing need to utilize legacy soil data sources for monitoring SOC changes (Boubehziz et al., 2020; Chen et al., 2019; Francaviglia et al., 2012). Regions that have soil monitoring networks need periodic recollection of soil samples to evaluate changes over time. Such resampling could be minimized to contain costs, but should be large enough to produce reliable estimates.

The need for long-term temporal paired sites is essential when aiming to depict SOC changes (Bellamy et al., 2005; Zhang and McGrath, 2004). Few countries have such monitoring schemes. The European Union (EU) started around 10 years ago with a Pan-European monitoring network to improve sustainable farming solutions and monitor soil pollution (Ballabio et al., 2014; Lugato et al., 2014).

Soil and SOC conservation is critically important in semi-arid areas, where desertification risk is increasing (Plaza et al., 2018). For these areas, a recent increase in the output of literature regarding SOC accounting and spatial modelling (Smith et al., 2005), legacy databases (Schillaci et al. 2018), and digital soil mapping (Minasny and McBratney, 2016) has been noted.

Legacy soil data and soil maps could be integrated into a unified database. This would provide special insight into hard-to-sample areas, past and present trends, and insight into the application of proper modelling procedures.

Recently, the development of digital soil mapping, or pedometrics, and the presence of an ample archive of historical soil data has allowed for the assessment of country scale SOC patterns with relatively high accuracy (Krol, 2008; McBratney et al., 2014). However, models can amplify uncertainty when the assessment is based on multiple predictors (Grunwald et al., 2018; Shi et al., 2018).

In some areas of the world, a lack of recent SOC measurements is prompting a rediscovery of legacy data which is in the process of being fully integrated into mapping methods at an operational level (Arrouays et al., 2017).

SOC distribution is determined by multiple factors (Lacoste et al., 2014), the importance of which vary mainly with bioclimatic conditions. It is therefore hard to delineate general functions that explain the world SOC distribution using only geographical position, although a general inverse correlation was found with average annual air temperature on a regional scale between 52° N and 40° S and a direct correlation beyond this region (Huang et al., 2018).

Land use and land use change is also a main driver of SOC stocks, although mechanisms of SOC dynamics seem to be independent of the ecosystem type or land use (Giannetta et al., 2018). Guo and Gifford (2002) showed that around 50% of SOC is gained in the transition from cropland to secondary vegetation communities with a meta-analytic approach, and recent papers confirmed such a trend (Wei et al., 2014; Zhou et al., 2015).

Sommer and Bossio (2014) hypothesized that SOC sequestration in arable land can show a 0.012–0.027% annual increase in the first two decades after the establishment of SOC preservation practices, after which a saturation occurs and the increase ceases. Following the same hypotheses, Zomer et al. (2017) presented a global assessment of cropland SOC under the aforementioned scenarios and found that the potential SOC sequestration in cropland is below 53% of the 4p1000 target (Minasny et al., 2017).

High spatial variability and temporal trends induced spatial modellers to design reliable sampling strategies (Biswas and Zhang, 2018; Schmidt et al., 2008; Shiwen et al., 2017) and develop efficient methods to compare intra-field and inter-field variations (Pezzuolo et al., 2017) with similar agro-ecological conditions over the course of two decades (Kühnel et al., 2019). Application of this technique was carried out to determine the effects of sampling density on interpolation accuracy (Long et al., 2018) and uncertainty assessment (Szatmári et al., 2018; Veronesi and Schillaci, 2019).

Cropland covers 12.6% of the world's surface (FAOSTAT data, accessed 2019). Cropland SOC has been mapped on a global scale using the WoSiS database (Batjes et al., 2017; Hengl et al., 2014; Zomer et al., 2017) and SOC maps were obtained by applying Generalized Additive Models (GAM) and machine learning methods.

At the European level, cropland plays an important role. Due to the large area covered, cropland acts as a potential carbon (C) sink. If considering a biomass return of up to 45 Mg C per year in raw materials, the biological potential of cropland for C storage is on the order of 90–120 Mg C per year (Freibauer et al., 2004; Smith, 2004). In particular, Smith (2004)

demonstrated that models of SOC changes should be used with statistical power analyses for planning sample design to determine density and time of sampling during experiments.

Little is known about long-term SOC changes in Mediterranean semi-arid arable lands, which are frequently dominated by winter-growing species (mostly cereals and legumes) in rotation with fallow periods characterized by various crop residue management practices. Field crop production in these areas can cause SOC depletion (Rodríguez Martín et al., 2019) and soil loss by erosion, especially when conventional tillage (e.g., ploughing) is continuously applied. Conversely, no-till has been shown to be strongly beneficial compared to conventional tillage in semi-arid climates with an aridity index lower than 0.52 ± 0.03 (Gristina et al., 2018). Land management practices such as reduction of tillage intensity (Sperow, 2020), addition of manure and sowing cover crops could help to increase SOC contents and keep significant amounts of nutrients (Álvaro-Fuentes et al., 2012b; Barbera et al., 2012).

In addition, in Mediterranean areas, frequent fire events can burn tons of biomass; this lowers the yearly C input derived from crop residues utilization while increasing SOC permanence, and affect the cycle of several nutrients including nitrogen (Knicker, 2007). In these areas, short-term SOC changes due to management practices (and especially land use) can temporarily override background changes (Lou et al., 2011) since length of cultivation is a main driver of SOC variation (Wang et al., 2019). In particular, time-space expected variation in soil traits can be differentially modelled if a priori information is available. Such information can allow for the determination of a minimum sample size to test a hypothesis effectively. Defining the sample size and location is required to enhance the power analysis while reducing laboratory costs and maximizing the accuracy of the representation (Confalonieri et al., 2009).

This experiment aimed at verifying whether or not a legacy estimation of SOC changes (1994–2008 model results) from non-paired data (Schillaci et al., 2017) matches the SOC variation measured in paired soil samples after 23 years (1994–2017). Thirty temporal paired sites from Sicily (South of Italy) and under continuous crop cover were resampled (Schillaci et al., 2019) and included in the present study. The analysis focused on arable land as it represents the main land use in the study area. The land cover of these sites was verified using historical remote sensing imagery to confirm that each site was continuously cultivated during the intervening period. Minimum sample size was determined and locations were randomly selected. Topsoil SOC contents were determined using the same laboratory method as in 1994 and 2008 (Walkley and Black, 1934).

2 Materials And Methods

2.1 Study area

Sicily (25,286 km²) is a semi-arid to arid region of Italy. Sicily is characterised by prolonged droughts from mid or late spring to early or mid-fall, with the addition of high energy storms in fall and winter. Rainfed arable land selected using the CORINE code 211 is the most common land cover class in the area under study with roughly 300,000 ha yearly under cultivation with durum wheat (Istat, 2013).

Rainfed arable lands represented the target land cover in the study area as they represent 60% of the surface in this region. Thus, it is a primary candidate for C sequestration and mitigation of the anthropogenic impact on the landscape. The land is predominantly under private ownership and the average farm size is around 6 ha, in general family-run businesses (all farm types), with approximately 10% of foreign labour (Istat, 2013).

2.2 Sampling campaign

In the rainfed arable lands of Sicily, the seedbed is generally prepared by soil ploughing during late summer and one or two harrowing in early fall. The amount of nitrogen (N) applied in non-legume field crops is usually between 80–100 kg N ha⁻¹ year⁻¹ and durum wheat yield is between 2–4 Mg ha⁻¹ (with a harvest index ranging from 45 to 55%, and therefore with a

similar straw yield). Land cover data was derived from the Geographic information system of Sicily (SITR; <http://www.sitr.regione.sicilia.it/>). Availability of a legacy databases and determination of the minimum number of samples and their locations for SOC determination.

In this work, data obtained from a sampling campaign carried out in 1993-94 (SMSLD) were used. In particular, data about the rainfed arable land (CORINE code 211) were identified using the optimization procedure shown in Schillaci et al. (2019). Briefly, the whole SMSLD consists of 6674 checked samples. Samples from the CORINE Land use 2 (agricultural lands) were 5471 from 2886 locations, with 2, 2, and 3 layers sampled at the 1st, 2nd, and 3rd quartile, respectively. Within these 2886 locations, samples from the sole CORINE Land use 2.1 (arable land) were 2162 from 880 locations, with 2, 2, and 3 layers sampled at the 1st, 2nd, and 3rd quartile, respectively.

A *power analysis* (Lakens, 2013) was used to find the minimum number of samples needed to determine the SOC change with time (from 1994 to 2017). The expected change with time was derived from a modelled SOC variation (1994 to 2008) at the regional level, as observed in Schillaci et al. (2019). In particular, samples from the same soil layer, land use and sub-area were used (Fig. 1).

To define the effect size, means and standard deviation of the oldest SOC survey (1993) and the hypothesized change after 15 years were used. The effect size was computed at different degree of confidence (0.1, 0.05, 0.01) using the G-Power software (Faul et al., 2007). The estimated change comes from the difference between the average topsoil SOC concentration (0–30 cm) measured in 1994 and the estimated values at the same location in 2008 (predicted) at a 1-km spatial scale (Schillaci et al., 2017). This allows for six sets of data (Table 1) used as raw or \log_{10} transformed data, none of which was normally distributed after a Shapiro-Wilk test of normality (Table 2).

Table 1
Datasets use in the present experiment. \log_{10} of each of these databases were also computed.

Abbreviation	Number of sites	description
LEG-SOC94	302	Data of measured topsoil SOC collected in the 1994 in land use CORINE 2.1 also used in (Schillaci et al., 2017) after a normalization procedure for depth. See further for explanation of the normalization process.
EST-SOC08	302	Data of estimated topsoil SOC in land use CORINE 2.1 extracted in the coinciding location of the LEG-SOC94 after the BRT modelling built with samples taken in 2008 and provided in (Schillaci et al., 2017)
SOC94	30	Random samples from the LEG-SOC94 in land use CORINE 2.1 after control for stability of land use until the 2017. See further for explanation of the strategy to establish the number of samples and location
SOC17	30	Samples taken in April 2017 in coinciding locations of the C94
EST-SOC08-LEG-SOC94	302	Differences between data estimated in 2008 in the same locations of data collected in 1994.
SOC17-SOC94	30	Differences between data measured in 2017 in the same locations of data selected in 1994

The *expected difference* in fourteen years from 1994 (measured data in LEG-SOC94) to 2008 (estimated data in EST-SOC08) was subjected to the effect size computation (i.e., Cohen's *d*) by means of the Wilcoxon signed-rank test (Wilcoxon, 1945) in IBM SPSS software. Standard deviation of legacy data was computed. Cohen's *d* provided the minimum number of sites to be sampled in the 2017 survey on the original locations of the 1994 (LEG-SOC94) to ascertain if a change in SOC content occurred. When this number was achieved, the sampling consisted in the collection

of topsoil (0–30 cm) in 30 randomly selected sites for which coordinates were reported with up to six digit precision (thus with an error < 30 m; Schillaci et al., 2019).

A *two-tailed paired Wilcoxon-test* was performed because, despite previous findings (Schillaci et al., 2017), there was uncertainty regarding the pattern of differences between SOC2017 and SOC94. Such uncertainty derived from the internal variability of the model used in Schillaci et al. (2017) suffered by the lack of paired locations. Indeed, dynamics of the difference from the data in 1994 (measured in LEG-SOC94) to 2008 (estimated in EST-SOC08) was on average positive for each land use, but with a high variation. Determination of sampling sites was carried out via random selection. This random selection was performed within the set of locations determined in 1994 (SOC94).

2.3 Calculations of the legacy topsoil SOC

Original legacy SOC data were collected at various depths. To compare SOC data sampled in the 2017 at a uniform depth to the ones sampled in 1993, the Hobley and Wilson (2016) method was used to uniform the SOC concentration of the former sampling campaign. Such a method is based on an exponential generalized function, as follows:

$$\text{SOC}(d) = \text{SOC}_0 \times e^{-k \times d} \text{ (Eq. 1)}$$

where d is the depth (expressed in meters), SOC_0 is the concentration of SOC at the soil surface (%), and k is the depletion constant (m^{-1}). The Hobley and Wilson (2016) method first fits this model and finds an optimal k and SOC_0 for each location, then computes SOC at any depth (d). Note that in the original work (Hobley and Wilson, 2016), Eq. (1) also contains an additional term (i.e., SOC_∞) modelling the concentration of residual SOC to a soil depth tending to infinity. In this study, SOC_∞ was assumed to be null.

In this work, the Hobley and Wilson (2016) method was applied as follows: in location (10 locations) where more than 2 layers were sampled, the model was directly fitted applying a $\text{SOC}_{(0.30\text{ m})}$ depth threshold. For those locations where the number of layers sampled was less than or equal to 2 (292 locations) the function could not be fit. In this case, the above model was first fit on all locations in order to find an expected SOC_0 and k values. This resulted in $k = 0.4815$ and $\text{SOC}_0 = 1.4396$. Then, using only the k found, SOC_0 was computed for each location using Eq. (1) as follows:

$$\text{SOC}_0 = \text{SOC}(d') \times [e^{-0.4815 \times d'}]^{-1} \text{ (Eq. 2),}$$

where d' is the depth of one data point of the locations for which SOC concentrations are known. The accuracy of the fitted depth functions was expressed with the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE).

2.4 Historical land use and soil type in resampled sites

Web Mapping Services (WMS) from the regional geodata service (<http://www.sitr.regione.sicilia.it/>), consisting of aerial photograph surveys, were used to check the historical land cover. Sites were checked to ascertain that these sites maintained the same land use during the intervening period. To ensure this, aerial photographs of at least 3 surveys carried out from 1994 to 2017 (Fig. 2.) were consulted. To avoid sites being temporarily converted to short-term grassland or bare soil or abandoned, local farmers were also interviewed.

Soil samples from the selected sites, rainfed arable land (N = 302) derived from the 1994 sampling campaign, and the resampled soils (N = 30) had a consistent distribution among the main soil systems according with the Soil map of Italy (Costantini et al., 2014) (Supplemental materials Fig. 1). Around 80% of sites were Vertic soils or Cambisols or Regosols. All of these soils frequently showed calcic or calcareous subgroups.

2.5 Soil organic carbon determination

SOC17 measurements were carried out using the Walkley-Black method (Walkley and Black, 1934). The SOC values so obtained were within the range of LEG-SOC94. In LEG-SOC94, soil data were georeferenced, and this has allowed for the resampling at exact (field site) location.

2.6 Computation of SOC differences between 1994 and 2017 by means of ad hoc sampling

Since the residues from the means of the \log_{10} of SOC94 and SOC17 were normally distributed (Table 2), the \log_{10} -SOC variation from 1994 to 2017 was tested for difference from zero by means of a paired t -test. This does not assume that observations within each group are normal, but only that their residuals are normal (Hsu and Lachenbruch, 2014; Zar, 1999) (Table 2). Such assumptions were only partly met, since the Log distribution were normal, whereas the differences between log were not normal but its skewness was comprised between -1.96 and $+1.96$. To overcome such problems, a bootstrap ANOVA was performed (Krishnamoorthy et al., 2007). Changes in cropland SOC content have previously been accounted for with the same statistical approach (Wang et al., 2018) where no subsampling was done to assess the SOC change.

3 Results

3.1 Power analysis

Farmer interviews confirmed that the land use of the sampling sites have been continuously cultivated during the intervening period. The crop rotation was generally made with durum wheat (*Triticum durum* Desf.) followed by broad beans (*Vicia faba* L.), or other pulses alternated with fallow. The sampling campaign LEG-SOC94 involved 302 sites, with a SOC concentration in the CORINE land use 2.1 (arable) of $1.01 \pm 0.59\%$ (mean \pm s.d.; Table 3) after normalization to 0.3 m depth. This value was significantly lower than the mean SOC in the 30 sites selected in 1994 (i.e., the SOC94, 1.31%; Table 3), but very close to its median (1.05%). Mean (predicted) SOC of the EST-SOC08 database was $1.38 \pm 0.39\%$ (mean \pm s.d., with a median of 1.25%). SOC was expected to vary in the cropland between the 1994 and the 2008 values when using original data or \log_{10} data (Table 4). Such variation was due to an increase in SOC in 75.8% of the sites (2008). The calculated effect size of the SOC was 0.54 for the original data and 0.69 for the \log_{10} data (Table 5). According to Cohen (1988), such effect sizes correspond to 'medium to high' effect, which needed a minimum sample size ranging from 15 to 45 samples to be able to detect a SOC variation. Given this effect size and the power chosen for the Wilcoxon test, which is by default set to 80%, and a significance level of 5%, the calculated sample size required would be 30 samples. These were identified in LEG-SOC94 and collected in their respective paired locations SOC17. In 2017, only 12 sites showing a SOC concentration higher than 1994 were found, so SOC variation from 1994 to 2017 depended more on the SOC difference within each pair of samples than on the % of samples in 2017 having a SOC higher than in 1994.

Table 3
 Results of the Shapiro-Wilk for the raw and log-transformed data of the 1994, 2008, and 2017. Please see Table 1 for a description of each dataset.

Shapiro-Wilk		
	Stat	Sign.
Raw - LEG-SOC94*	0.903	< 0.001
Raw - EST-SOC08*	0.798	< 0.001
Log-LEG-SOC94*	0.980	< 0.001
Log-EST-SOC08*	0.883	< 0.001
Differences of raw data (EST-SOC08-LEG-SOC94)*	0.961	< 0.001
Differences of log data (EST-SOC08-LEG-SOC94)*	0.960	< 0.001
Raw - SOC94**	0.833	< 0.001
Raw - SOC17**	0.809	< 0.001
Log - SOC94**	0.951	0.183
Log - SOC17**	0.965	0.404
Differences of raw data (SOC17-SOC94)**	0.886	0.004
Differences of log data (SOC17-SOC94)**	0.824	< 0.001
* degrees of freedom = 302; ** degrees of freedom = 30		

Table 3
Descriptive statistics and main quantiles of each of the dataset used in the present study

	<i>SOC94</i>	<i>SOC17</i>	<i>LOG of SOC94</i>	<i>LOG of SOC17</i>	<i>LEG-SOC94 (measured)</i>	<i>EST-SOC08 (estimated)**</i>	<i>LOG of LEG-SOC94 (measured)</i>	<i>LOG of EST-SOC08 (estimated)**</i>
Mean*	1.31	1.23	0.00550	0.03800	1.01	1.38	-0.06983	0.12939
Standard deviation	0.96	0.67	0.33708	0.20808	0.59	0.39	0.26742	0.10394
Variance	0.93	0.45	0.11362	0.04330	0.35	0.15	0.07151	0.01080
Kurtosis	3.55	3.28	1.22	0.52	4.40	3.10	0.57	0.96
Skewness	1.76	1.78	-0.69	0.36	1.53	1.79	-0.55	1.23
n	30	30	30	30	302	302	302	302
quantiles								
Min	0.120	0.400	-0.9208	-0.3979	0.100	0.953	-1.000	-0.021
<i>0.010</i>	<i>0.149</i>	<i>0.429</i>	<i>-0.8445</i>	<i>-0.3698</i>	<i>0.150</i>	<i>1.001</i>	<i>-0.824</i>	<i>0.000</i>
<i>0.025</i>	<i>0.193</i>	<i>0.473</i>	<i>-0.7300</i>	<i>-0.3277</i>	<i>0.215</i>	<i>1.012</i>	<i>-0.667</i>	<i>0.005</i>
<i>0.050</i>	<i>0.270</i>	<i>0.545</i>	<i>-0.5783</i>	<i>-0.2654</i>	<i>0.300</i>	<i>1.027</i>	<i>-0.523</i>	<i>0.012</i>
<i>0.250</i>	<i>0.780</i>	<i>0.803</i>	<i>-0.1079</i>	<i>-0.0956</i>	<i>0.560</i>	<i>1.142</i>	<i>-0.252</i>	<i>0.058</i>
<i>0.500</i>	<i>1.050</i>	<i>1.100</i>	<i>0.0205</i>	<i>0.0414</i>	<i>0.945</i>	<i>1.253</i>	<i>-0.025</i>	<i>0.098</i>
<i>0.750</i>	<i>1.633</i>	<i>1.353</i>	<i>0.2128</i>	<i>0.1310</i>	<i>1.328</i>	<i>1.487</i>	<i>0.123</i>	<i>0.172</i>
<i>0.950</i>	<i>3.343</i>	<i>2.830</i>	<i>0.5240</i>	<i>0.4493</i>	<i>2.099</i>	<i>2.219</i>	<i>0.322</i>	<i>0.346</i>
<i>0.975</i>	<i>3.699</i>	<i>3.128</i>	<i>0.5648</i>	<i>0.4952</i>	<i>2.260</i>	<i>2.477</i>	<i>0.354</i>	<i>0.394</i>
<i>0.990</i>	<i>4.156</i>	<i>3.171</i>	<i>0.6155</i>	<i>0.5012</i>	<i>2.820</i>	<i>2.773</i>	<i>0.450</i>	<i>0.443</i>
Max	4.460	3.200	0.6493	0.5051	4.330	2.969	0.636	0.473
*for logs, mean of log represent the arithmetic mean of the logarithms ₁₀ of the SOC values expressed in %; ** such an estimation is from the 302 sites in which a measure was available 1994 and applying the 1-km resolution estimation process used in the 2008 described in (Schillaci et al., 2017).								

Table 4

Wilcoxon test for the difference between the soil organic carbon (SOC) of the arable lands (CORINE 2.1) in the 0–30 cm layer in 1994 (LEG-SOC94) and estimated SOC in 2008 (EST-SOC08) in the coinciding locations of the samples taken in 1994.

		Raw		Logs	
Ranks	N	Mean ranks	Sum of ranks	Mean ranks	Sum of ranks
Negative ranks (EST-SOC08 < LEG-SOC94)	73	123.6	9026	100.3	7320
Positive ranks (EST-SOC08 > LEG-SOC94)	229	160.4	36727	167.8	38433
Ties (EST-SOC08 = LEG-SOC94)	0				
Test statistics based on positive ranks (Wilcoxon)					
Z based on negative ranks		-9.119		-10.243	
Asymptotic significance (2-tailed)		< 0.001		< 0.001	

Table 5

Output parameters of the *a priori* power analysis computation process at varying the α (0.10; 0.05; or 0.01). The process was carried out through the G-Power software with the Wilcoxon test for non-normal distributed datasets (Faul et al., 2007). Input data were from 302 measured samples of SOC in 1994 and modelled SOC in 2008, each of which expressed as either raw or log₁₀ data. Minimum power was set to 0.80.

α	0.10	0.05	0.01
Raw data			
Noncentrality parameter δ	2.6008000	2.9077828	3.5612921
Critical t	1.7174255	2.0495831	2.6981518
Degrees of freedom	21.9183118	27.6478898	41.9718346
Minimum sample size needed	24	30	45
Actual power	0.8090485	0.8010923	0.8005742
Log ₁₀ data			
Noncentrality parameter δ	2.6384055	3.0465682	3.6685540
Critical t	1.7676466	2.1001020	2.7730806
Degrees of freedom	13.3239449	18.0985932	26.6929601
Minimum sample size needed	15	20	29
Actual power	0.8040250	0.8216418	0.8056966

3.2 Descriptive statistics of SOC and SOC variation from 1994 to 2017

SOC distribution in the datasets used in the present experiment had different means and standard deviations, but similar skewness and kurtosis (Table 3). SOC94 had a mean of $1.31 \pm 0.96\%$ (mean \pm s.d.), whereas in the LEG-SOC94 sites, mean was $1.01 \pm 0.59\%$ (Fig. 3). The SOC17 mean was on average significantly lower than EST-SOC08, but with similar

distribution properties to both the SOC94 and LEG-SOC94. Also, the SOC17 mean was slightly lower (-6.1% relative change, - 0.08% absolute change) than the mean of SOC94, but not than the EST-SOC08. The accuracy of the fitted depth functions was expressed with the RMSE and MAE resulting in 0.36 and 0.23 respectively. Transformation to logarithms improved the distribution properties in term of skewness and kurtosis, especially for SOC17 (Fig. 4). According to the Wilcoxon test, the change in SOC between SOC94 and SOC17 did not appear different from zero (Table 6 and Fig. 5). The t-test applied to the Log_{10} of SOC17 and SOC94, in which residues were normally distributed, provided similar results compared to the Wilcoxon test carried out on the same data (SOC difference = $+0.03300 \pm 0.27659$, mean \pm s.d.; C.I._{95%}: -0.07028 to +0.13628; $t = +0.653$; 2-tails significance = 0.52). The bootstrap ANOVA also provided consistent results compared to the Wilcoxon test (Table 7), but notably, the difference in the log-SOC between the sampling campaigns had 95% confidence intervals marginally overlapping zero.

Table 6

Results of the Wilcoxon tests of the mean difference of 30 samples of soil organic carbon measured in the 2017 (SOC17) and 1994 (SOC94) as difference of the raw data or log-transformed. Positive mean indicates a mean increase in SOC with time.

	Raw			Logs		
	N	Mean ranks	Sum of ranks	N	Mean ranks	Sum of ranks
Ranks						
Negative ranks (SOC17 < SOC94)	18	14.42	259.5	16	12.75	204
Positive ranks (SOC17 > SOC94)	12	17.13	205.5	12	16.83	202
Ties (SOC17 = SOC94)	0			2		
Test statistics based on positive ranks (Wilcoxon)						
Z based on negative ranks		-0.556			-0.023	
Asymptotic significance (2-tailed)		0.578			0.982	

Table 7

Results of the bootstrap ANOVA on the differences between each pair of dataset. N = 30. Bootstrap samples were 10000.

	Bias	Std. Error	Sig. (2-tailed)	Lower 95% C.I.*	Upper 95% C.I.
Differences of raw data (SOC17-SOC94)	0.0013	0.1049	0.4780	-0.2923	0.1167
Differences of log data (SOC17-SOC94)	0.0007	0.0505	0.5400	-0.0583	0.1394
* C.I., confidence interval					

4 Discussion

This work was aimed at assessing the reliability of estimations of topsoil SOC changes with non-paired data using time-paired sampling. Compared to a previous work on temporal variation in the same area (Schillaci et al., 2017), various improvements in methodology were applied. These improvements consisted of the use of temporal paired sites instead of estimates based on different locations. Locations were further checked for land use continuity. The previous analyses based on modelled data over a 15-year span (from 1994 to 2008) predicted a mean relative increase of around the 21% of SOC content in arable lands, but such an increase was affected by a strong variability when survey data was reviewed. By using a monitoring network spanning 30 years, Gubler et al. (2019) found that SOC dynamic is more determined by a change in land use than other predictors in a colder climate (Switzerland).

According to a hypothetical linear growth, the increase expected from 1994 to 2017, which was the target period of the present study, was predicted to be a relative SOC increase of around 30%. In this study, no evidence of this magnitude of increase was found, and no such difference was seen in neither the original data nor in the log transformed data. When using a similar amount of data from meadows, Gubler et al. (2019) found that the minimum detectable change in 10 to 100 years (at a power of the 80%, such as in the present study) spanned from approximately 2 to 6%, respectively. Bellamy et al. (2005) showed that variation in time strongly depended on the initial SOC content, e.g., sites with low SOC had more opportunity to increase their SOC than sites with high SOC. However, both these latter studies were conducted in wetter climates than that of the present study. Using data from 20 regions in the world, Minasny et al. (2017) showed a tendency of a higher C sequestration potential on croplands with low initial SOC stock ($\leq 30 \text{ t C / ha}$ at 0–30 cm) compared to grasslands, which already have a high initial SOC stock, although a general decrease of the stock rate with time was also reported.

In addition, most of the samples in the present study were derived from the thermo-Mediterranean bioclimatic area (Rivas-Martínez et al., 2011), so these results can also reflect the latency of SOC variation in these areas, which even under soil abandonment showed limited increases in SOC when compared to other bioclimates in the same region (Novara et al., 2017).

The lack of an increase found from 1994 to 2017 when compared to the estimated increase from 1994 to 2008 was partly expected given that a SOC mean of the SOC94 extracted from the LEG-SOC94 was higher than that of the complete LEG-SOC94 dataset. In few exceptions (3 samples) there was an increase in SOC concentration that can also be due to soil deposition following erosion from the sites at higher altitudes (Lizaga et al., 2019; Navas et al., 2017; Panagos et al., 2015), changes in tillage depth or soil compaction (Álvaro-Fuentes et al., 2012a). In addition, in this area soil respiration due to increasing temperatures (Viola et al., 2014) may have offset any potential increase in SOC, and the 23-year time span may have not been sufficient to detect a SOC change, as pointed by Saby et al. (2008). Soil management with ploughing and the increasing mean temperature (Viola et al., 2014) are not conducive to SOC accumulation (Goidts and van Wesemael, 2007; Kämpf et al., 2016). In particular, Goidts and van Wesemael (2007) showed that ploughing may override the increase in SOC over time. Also, an increase in SOC in ploughed soils is hard to achieve unless high quantities of organic residues and N are provided (Mazzoncini et al., 2011). These two latter conditions are very limited in Sicily due to the low crop yield, the low amount of residues returned to the soil and scarce fertilization. Other authors showed that slight increases in temperature and water availability ratio may contribute in reducing SOC (Davidson and Janssens, 2006). In the present work, this may have occurred and may be what can be seen when comparing the expected change from 1994 to 2008 to the measured change from 1994 to 2017. In addition, the ratio between water availability and temperature may be increasing in the area under study, despite the fact that no direct report is available (Viola et al., 2014).

The lack of increase in SOC concentration found here (from 1994 to 2017), compared to those estimated in the period from 1994 to 2008, could also be due to a SOC reduction from 2008 to 2017 which cannot be excluded using the present data. Such a reduction from 2008 to 2017 might possibly be connected with the decoupling of EU Common Agricultural Policy (CAP) payments regarding agriculture in 2005 (Regulation EEC 1782/2003). Before this year, wheat was the continuous primary crop for arable lands and after this point crop rotation with legumes or fallow land was encouraged by new regulations. Notably, continuous wheat has been shown to favour SOC accumulation when compared to a high percentage of wheat-legume rotations or wheat-fallow rotations, in the same or similar environments (Barbera et al., 2012; López-Bellido et al., 2010; West and Post, 2002). Lastly, differences between the previous estimation (1994 to 2008) and the present measurements (1994 to 2017) can also depend on the differences between direct and indirect measurement (Smith et al., 2019), or transient changes in cultivation history that may have influenced the previous estimates (Wang et al., 2019), the latter of which was discarded here by an ad-hoc sampling in continuously ploughed soils with field crops.

Variability and confidence intervals regarding SOC suggest that the estimated change in the previous work (Schillaci et al., 2017) could have been affected by outliers or errors in the measurements related to analytical methods used (Gessesse

and Khamzina, 2018), especially when used for highly alkaline soils with scarce SOC (Bisutti et al., 2004; Walkley, 1935), such as in the present study. These issues may have produced an apparent pattern of SOC accumulation from 1994 to 2008 that was not detected between 1994 and 2017. Problems in the estimate due to the sampling strategy (Brus, 2019; Wadoux and Brus, 2020) were excluded, since the selection method used here provided a dataset with similar statistical properties to the original SOC distribution, and thus allowed for a maximum reduction in sampling sites.

Mitigation strategies and international projects have contributed to a debate regarding SOC sequestration. These traits are increasingly being taken into account in EU subsidies to the agricultural sector (Ciliberti and Frascarelli, 2015; European Commission, 2017).

There were different ongoing discussions following the paper by Sommer and Bossio (2014) and more recently after the Soil 4x1000 initiative (Minasny et al., 2017), all of which are above all pivotal to a reliable estimation of SOC dynamics. Zomer et al. (2017) modelled the estimated increase in SOC concentration and stocks at a global scale in croplands and found that an average increase of approximately +26%, which is similar to the estimated value found from 1994 to 2008. However, the Zomer et al. (2017) estimation may not be suitable for small scale assessment and mapping and the use of legacy information can be crucial to confirm these trends.

5 Conclusion

The estimated SOC change in arable lands occurred from 1994 to 2008 in the studied area was not confirmed by a direct measurement using paired sites in the 1994–2017 timespan. This result has a direct implication for the SOC monitoring network in the mid-term (e.g., 15–25 years). Results also suggest that measurements of soil properties based on legacy data need to be supported by reliable information on the land use, land use changes and soil management practices. Both these aspects were indirectly taken into account here by choosing sites with no change in the land use or soil management. These variables could be used to correct the effect of other predictors (e.g., rainfall, slope). Only when many limitations are found, there would be the need for large scale sampling campaigns.

The debate regarding the structuring of subsidies should take into account the current information to assure on the one hand, their competitiveness, and on the other hand, their environmental sustainability and ability to provide ecosystem services. Further works should aim at: i) increasing the number of sites to be resampled which would be derived from both the collection in 1994 and 2008, ii) gathering information regarding land use dynamics in paired sites, and iii) quantifying the effect of soil erosion in the flow of SOC within and among catchments. Lastly, direct information regarding the variation in SOC concentration and stock in topsoil and subsoil in the target areas are urgently needed to drive policy making. The discrepancies between the present data compared to the previously published results may depend on various factors, including: I) possible errors in 1994 measurements and 2008 estimates, II) changes in land use and, III) soil erosion, IV) bias in the comparisons when using soil samples collected at different depth and showing different bulk density. However, SOC concentration reduction from 2008 to 2017 cannot be excluded. A further study is needed to evaluate a) possible errors in the determination of SOC concentrations in soil from semiarid to arid conditions, b) possible changes in bulk density with time and depth, c) SOC changes during the intermediate periods and to the present, and d) the influence of wildfire and arson.

Declarations

Availability of data and materials

The legacy data that support the findings of this study were made available from [the Regional Bureau for Agriculture, Rural Development and Mediterranean Fishery, the Department of Agriculture, ARTA. Palermo]. The new data samples

taken in 2017 can be obtained upon request. The legacy data are subjected to licensing; in this study they were used under an agreement.

Competing interests

The Authors declare that they have no competing interests

Funding

The work is partially founded with the H2020 program Landsupport, ID: 774234

Authors' contributions

CS, MA, SS, AP, conceptualized the methodology used in the study. CS, SS, AP, AL, developed the data investigation, validation and visualization, with the supervision of MA. CS, AP, SS, AL, MA, contributed to the writing of the original and CZ contributed to the manuscript revision. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank M.G Matranga, V. Ferraro and A. Guaitoli from the Regional Bureau for Agriculture, Rural Development and Mediterranean Fishery, the Department of Agriculture, ARTA, Palermo, that have provided the legacy data.

References

- Álvaro-Fuentes, J., Easter, M., Paustian, K., 2012a. Climate change effects on organic carbon storage in agricultural soils of northeastern Spain. *Agric. Ecosyst. Environ.* 155, 87–94. <https://doi.org/10.1016/j.agee.2012.04.001>
- Álvaro-Fuentes, J., Morell, F.J., Plaza-Bonilla, D., Arrúe, J.L., Cantero-Martínez, C., 2012b. Modelling tillage and nitrogen fertilization effects on soil organic carbon dynamics. *Soil Tillage Res.* 120, 32–39. <https://doi.org/10.1016/j.still.2012.01.009>
- Arrouays, D., Lagacherie, P., Hartemink, A.E., 2017. Digital soil mapping across the globe. *Geoderma Reg.* 9, 1–4. <https://doi.org/https://doi.org/10.1016/j.geodrs.2017.03.002>
- Ballabio, C., Panagos, P., Montanarella, L., 2014. Predicting soil organic carbon content in Cyprus using remote sensing and Earth observation data, in: *Proceedings of SPIE - The International Society for Optical Engineering*. European Commission, Joint Research Centre, Institute for Environment and Sustainability, Via E. Fermi 2749, I-21027 Ispra (VA), Italy. <https://doi.org/10.1117/12.2066406>
- Barbera, V., Poma, I., Gristina, L., Novara, A., Egli, M., Egli, S., 2012. Long-term cropping systems and tillage management effects on soil organic carbon stock and steady state level of C sequestration rates in a semiarid environment. *L. Degrad. Dev.* 23, 82–91. <https://doi.org/10.1002/ldr.1055>
- Batjes, N.H., Ribeiro, E., van Oostrum, A., Leenaars, J., Hengl, T., Mendes de Jesus, J., 2017. WoSIS: providing standardised soil profile data for the world. *Earth Syst. Sci. Data* 9, 1–14. <https://doi.org/10.5194/essd-9-1-2017>
- Bellamy, P.H., Loveland, P.J., Bradley, R.I., Lark, R.M., Kirk, G.J.D., 2005. Carbon losses from all soils across England and Wales 1978-2003. *Nature* 437, 245–248. <https://doi.org/10.1038/nature04038>
- Bisutti, I., Hilke, I., Raessler, M., 2004. Determination of total organic carbon – an overview of current methods. *TrAC Trends Anal. Chem.* 23, 716–726. <https://doi.org/10.1016/j.trac.2004.09.003>

- Biswas, A., Zhang, Y., 2018. Sampling Designs for Validating Digital Soil Maps: A Review. *Pedosph. An Int. J.* 28, 1–15. [https://doi.org/10.1016/S1002-0160\(18\)60001-3](https://doi.org/10.1016/S1002-0160(18)60001-3)
- Boubehziz, S., Khanchoul, K., Benslama, M., Benslama, A., Marchetti, A., Francaviglia, R., Piccini, C., 2020. Predictive mapping of soil organic carbon in Northeast Algeria. *Catena* 190, 104539. <https://doi.org/10.1016/j.catena.2020.104539>
- Brus, D.J., 2019. Sampling for digital soil mapping: A tutorial supported by R scripts. *Geoderma* 338, 464–480. <https://doi.org/10.1016/j.geoderma.2018.07.036>
- Chen, S., Angers, D.A., Martin, M.P., Walter, C., 2019. Soil carbon stocks under different land uses and the applicability of the soil carbon saturation concept. *Soil Tillage Res.* 188, 53–58. <https://doi.org/10.1016/J.STILL.2018.11.001>
- Chenu, C., Angers, D.A., Barré, P., Derrien, D., Balesdent, J., 2019. Increasing organic stocks in agricultural soils: Knowledge gaps and potential innovations. *Soil Tillage Res.* 188, 41–52. <https://doi.org/10.1016/J.STILL.2018.04.011>
- Ciliberti, S., Frascarelli, A., 2015. A critical assessment of the implementation of CAP 2014- 2020 direct payments in Italy. *Bio-based Appl. Econ.* 4, 261–277. <https://doi.org/10.13128/BAE-16377>
- Confalonieri, R., Perego, A., Chiodini, M.E., Scaglia, B., Rosenmund, A.S., Acutis, M., 2009. Analysis of sample size for variables related to plant, soil, and soil microbial respiration in a paddy rice field. *F. Crop. Res.* 113, 125–130. <https://doi.org/10.1016/j.fcr.2009.04.014>
- Costantini, E.A.C., Lorenzetti, R., Malorgio, G., 2016. A multivariate approach for the study of environmental drivers of wine economic structure. *Land use policy* 57, 53–63. <https://doi.org/10.1016/j.landusepol.2016.05.015>
- Davidson, E.A., Janssens, I.A., 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*. <https://doi.org/10.1038/nature04514>
- Deng, X., Chen, X., Ma, W., Ren, Z., Zhang, M., Grieneisen, M.L., Long, W., Ni, Z., Zhan, Y., Lv, X., 2018. Baseline map of organic carbon stock in farmland topsoil in East China. *Agric. Ecosyst. Environ.* 254, 213–223. <https://doi.org/10.1016/J.AGEE.2017.11.022>
- European Commission, 2017. Cap context indicators 2014-2020., available at: https://ec.europa.eu/info/files/context-indicator-fiches_en.
- Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. <https://doi.org/10.3758/BF03193146>
- Francaviglia, R., Coleman, K., Whitmore, A.P., Doro, L., Urracci, G., Rubino, M., Ledda, L., 2012. Changes in soil organic carbon and climate change - Application of the RothC model in agro-silvo-pastoral Mediterranean systems. *Agric. Syst.* 112, 48–54. <https://doi.org/10.1016/j.agsy.2012.07.001>
- Freibauer, A., Rounsevell, M.D., Smith, P., Verhagen, J., 2004. Carbon sequestration in the agricultural soils of Europe. *Geoderma* 122, 1–23. <https://doi.org/10.1016/j.geoderma.2004.01.021>
- Gessesse, T.A., Khamzina, A., 2018. How reliable is the Walkley-Black method for analyzing carbon-poor, semi-arid soils in Ethiopia? *J. Arid Environ.* 153, 98–101. <https://doi.org/10.1016/J.JARIDENV.2018.01.008>
- Giannetta, B., Plaza, C., Vischetti, C., Cotrufo, M.F., Zaccone, C., 2018. Distribution and thermal stability of physically and chemically protected organic matter fractions in soils across different ecosystems. *Biol. Fertil. Soils* 54, 671–681. <https://doi.org/10.1007/s00374-018-1290-9>

- Goidts, E., van Wesemael, B., 2007. Regional assessment of soil organic carbon changes under agriculture in Southern Belgium (1955–2005). *Geoderma* 141, 341–354. <https://doi.org/10.1016/j.geoderma.2007.06.013>
- Gristina, L., Keesstra, S., Novara, A., 2018. No-till durum wheat yield success probability in semi arid climate: A methodological framework. *Soil Tillage Res.* 181, 29–36. <https://doi.org/10.1016/J.STILL.2018.04.002>
- Grunwald, S., Yu, C., Xiong, X., 2018. Transferability and Scalability of Soil Total Carbon Prediction Models in Florida, USA. *Pedosphere* 28, 856–872. [https://doi.org/10.1016/S1002-0160\(18\)60048-7](https://doi.org/10.1016/S1002-0160(18)60048-7)
- Gubler, A., Wächter, D., Schwab, P., Müller, M., Keller, A., 2019. Twenty-five years of observations of soil organic carbon in Swiss croplands showing stability overall but with some divergent trends. *Environ. Monit. Assess.* 191. <https://doi.org/10.1007/s10661-019-7435-y>
- Guo, L.B., Gifford, R.M., 2002. Soil carbon stocks and land use change: A meta analysis. *Glob. Chang. Biol.* 8, 345–360. <https://doi.org/10.1046/j.1354-1013.2002.00486.x>
- Hengl, T., De Jesus, J.M., MacMillan, R.A., Batjes, N.H., Heuvelink, G.B.M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Gonzalez, M.R., 2014. SoilGrids1km - Global soil information based on automated mapping. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0105992>
- Hobley, E.U., Wilson, B., 2016. The depth distribution of organic carbon in the soils of eastern Australia. *Ecosphere* 7, e01214. <https://doi.org/10.1002/ecs2.1214>
- Hsu, H., Lachenbruch, P.A., 2014. Paired t-Test, in: *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd, Chichester, UK. <https://doi.org/10.1002/9781118445112.stat05929>
- Huang, J., Minasny, B., McBratney, A.B., Padarian, J., Triantafyllis, J., 2018. The location- and scale- specific correlation between temperature and soil carbon sequestration across the globe. *Sci. Total Environ.* 615, 540–548. <https://doi.org/10.1016/j.scitotenv.2017.09.136>
- Istat, 2013. 6° censimento generale dell'agricoltura in Sicilia. https://doi.org/https://www4.istat.it/it/files/2012/12/6%C2%B0_censimento_agricoltura_in_Sicilia_Risultati_definitivi.pdf?title=Censimento+dell%E2%80%99Agricoltura+in+Sicilia++05%2Fdic%2F2012++Testo+integrale.pdf
- Kämpf, I., Hölzel, N., Störle, M., Broll, G., Kiehl, K., 2016. Potential of temperate agricultural soils for carbon sequestration: A meta-analysis of land-use effects. *Sci. Total Environ.* 566, 428–435. <https://doi.org/10.1016/j.scitotenv.2016.05.067>
- Knicker, H., 2007. How does fire affect the nature and stability of soil organic nitrogen and carbon? A review. *Biogeochemistry* 85, 91–118. <https://doi.org/10.1007/s10533-007-9104-4>
- Köhl, M., Lister, A., Scott, C.T., Baldauf, T., Plugge, D., 2011. Implications of sampling design and sample size for national carbon accounting systems. *Carbon Balance Manag.* 6, 10. <https://doi.org/10.1186/1750-0680-6-10>
- Krishnamoorthy, K., Lu, F., Mathew, T., 2007. A parametric bootstrap approach for ANOVA with unequal variances: Fixed and random models. *Comput. Stat. Data Anal.* 51, 5731–5742. <https://doi.org/10.1016/j.csda.2006.09.039>
- Krol, B.G.C.M., 2008. Towards a Data Quality Management Framework for Digital Soil Mapping with Limited Data BT - Digital Soil Mapping with Limited Data, in: Hartemink, A.E., McBratney, A., Mendonça-Santos, M. de L. (Eds.), . Springer Netherlands, Dordrecht, pp. 137–149. https://doi.org/10.1007/978-1-4020-8592-5_11

- Kühnel, A., Garcia-Franco, N., Wiesmeier, M., Burmeister, J., Hobley, E., Kiese, R., Dannenmann, M., Kögel-Knabner, I., 2019. Controlling factors of carbon dynamics in grassland soils of Bavaria between 1989 and 2016. *Agric. Ecosyst. Environ.* 280, 118–128. <https://doi.org/10.1016/j.agee.2019.04.036>
- Lacoste, M., Minasny, B., McBratney, A., Michot, D., Viaud, V., Walter, C., 2014. High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. *Geoderma* 213, 296–311. <https://doi.org/10.1016/j.geoderma.2013.07.002>
- Lakens, D., 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lal, R., 2008. Carbon sequestration. *Philos. Trans. R. Soc. B Biol. Sci.* <https://doi.org/10.1098/rstb.2007.2185>
- Lark, R.M., 2009. Estimating the regional mean status and change of soil properties: Two distinct objectives for soil survey. *Eur. J. Soil Sci.* 60, 748–756. <https://doi.org/10.1111/j.1365-2389.2009.01156.x>
- Lizaga, I., Gaspar, L., Quijano, L., Dercon, G., Navas, A., 2019. NDVI, 137Cs and nutrients for tracking soil and vegetation development on glacial landforms in the Lake Parón Catchment (Cordillera Blanca, Perú). *Sci. Total Environ.* 651, 250–260. <https://doi.org/10.1016/j.scitotenv.2018.09.075>
- Long, J., Liu, Y., Xing, S., Qiu, L., Huang, Q., Zhou, B., Shen, J., Zhang, L., 2018. Effects of sampling density on interpolation accuracy for farmland soil organic matter concentration in a large region of complex topography. *Ecol. Indic.* 93, 562–571. <https://doi.org/10.1016/J.ECOLIND.2018.05.044>
- López-Bellido, R.J., Fontán, J.M., López-Bellido, F.J., López-Bellido, L., 2010. Carbon Sequestration by Tillage, Rotation, and Nitrogen Fertilization in a Mediterranean Vertisol. *Agron. J.* 102, 310. <https://doi.org/10.2134/agronj2009.0165>
- Lou, Y., Xu, M., Wang, W., Sun, X., Zhao, K., 2011. Return rate of straw residue affects soil organic C sequestration by chemical fertilization. *Soil Tillage Res.* 113, 70–73. <https://doi.org/10.1016/j.still.2011.01.007>
- Lugato, E., Panagos, P., Bampa, F., Jones, A., Montanarella, L., 2014. A new baseline of organic carbon stock in European agricultural soils using a modelling approach. *Glob. Chang. Biol.* 20, 313–326. <https://doi.org/10.1111/gcb.12292>
- Mazzoncini, M., Sapkota, T.B., Bàrberi, P., Antichi, D., Risaliti, R., Bahadur, T., Ba, P., 2011. Long-term effect of tillage, nitrogen fertilization and cover crops on soil organic carbon and total nitrogen content. *Soil Tillage Res.* 114, 165–174. <https://doi.org/10.1016/j.still.2011.05.001>
- McBratney, A., Field, D.J., Koch, A., 2014. The dimensions of soil security. *Geoderma* 213, 203–213. <https://doi.org/10.1016/J.GEODERMA.2013.08.013>
- Minasny, B., Malone, B.P., McBratney, A.B., Angers, D.A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.-S., Cheng, K., Das, B.S., Field, D.J., Gimona, A., Hedley, C.B., Hong, S.Y., Mandal, B., Marchant, B.P., Martin, M., McConkey, B.G., Mulder, V.L., O'Rourke, S., Richer-de-Forges, A.C., Odeh, I., Padarian, J., Paustian, K., Pan, G., Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C.-C., Vågen, T.-G., van Wesemael, B., Winowiecki, L., 2017. Soil carbon 4 per mille. *Geoderma* 292, 59–86. <https://doi.org/10.1016/j.geoderma.2017.01.002>
- Minasny, B., McBratney, A.B., 2016. Digital soil mapping: A brief history and some lessons. *Geoderma* 264, 301–311. <https://doi.org/10.1016/j.geoderma.2015.07.017>
- Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I., 2013. Chapter One – Digital Mapping of Soil Carbon, in: *Advances in Agronomy*. pp. 1–47. <https://doi.org/10.1016/B978-0-12-405942-9.00001-3>

- Navas, A., Quine, T.A., Walling, D.E., Gaspar, L., Quijano, L., Lizaga, I., 2017. Relating Intensity of Soil Redistribution to Land Use Changes in Abandoned Pyrenean Fields Using Fallout Caesium-137. *L. Degrad. Dev.* 28, 2017–2029. <https://doi.org/10.1002/ldr.2724>
- Novara, A., Gristina, L., Sala, G., Galati, A., Crescimanno, M., Cerdà, A., Badalamenti, E., La Mantia, T., 2017. Agricultural land abandonment in Mediterranean environment provides ecosystem services via soil carbon sequestration. *Sci. Total Environ.* 576, 420–429. <https://doi.org/10.1016/j.scitotenv.2016.10.123>
- Odeh, I.O.A., Leenaars, J., Hartemink, A., 2012. The challenges of collating legacy data for digital mapping of Nigerian soils. *Digit. Soil Assessments Beyond* 453–458. <https://doi.org/10.1201/b12728-88>
- Panagos, P., Borrelli, P., Meusburger, K., Alewell, C., Lugato, E., Montanarella, L., 2015. Estimating the soil erosion cover-management factor at the European scale. *Land use policy* 48, 38–50. <https://doi.org/10.1016/j.landusepol.2015.05.021>
- Parras-Alcántara, L., Lozano-García, B., Brevik, E.C., Cerdá, A., 2015. Soil organic carbon stocks assessment in Mediterranean natural areas: A comparison of entire soil profiles and soil control sections. *J. Environ. Manage.* 155, 219–228. <https://doi.org/10.1016/j.jenvman.2015.03.039>
- Pezzuolo, A., Dumont, B., Sartori, L., Marinello, F., De Antoni Migliorati, M., Basso, B., 2017. Evaluating the impact of soil conservation measures on soil organic carbon at the farm scale. *Comput. Electron. Agric.* 135, 175–182. <https://doi.org/10.1016/j.compag.2017.02.004>
- Plaza, C., Zaccone, C., Sawicka, K., Méndez, A.M., Tarquis, A., Gascó, G., Heuvelink, G.B.M., Schuur, E.A.G., Maestre, F.T., 2018. Soil resources and element stocks in drylands to face global issues. *Sci. Rep.* 8, 13788. <https://doi.org/10.1038/s41598-018-32229-0>
- Rivas-Martínez, S., Rivas-Sáenz, S., Penas-Merino, A., 2011. Worldwide bioclimatic classification system. *Glob. Geobot.* 1, 1–638. <https://doi.org/10.5616/gg110001>
- Rodríguez Martín, J.A., Álvaro-Fuentes, J., Gabriel, J.L., Gutiérrez, C., Nanos, N., Escuer, M., Ramos-Miras, J.J., Gil, C., Martín-Lammerding, D., Boluda, R., 2019. Soil organic carbon stock on the Majorca Island: Temporal change in agricultural soil over the last 10 years. *CATENA* 181, 104087. <https://doi.org/10.1016/J.CATENA.2019.104087>
- Rodríguez 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. <https://doi.org/10.1016/j.geoderma.2015.10.010>
- Saby, N.P.A., Bellamy, P.H., Morvan, X., Arrouays, D., Jones, R.J.A., Verheijen, F.G.A., Kibblewhite, M.G., Verdoodt, A., Üveges, J.B., Freudenschuß, A., Simota, C., 2008. Will European soil-monitoring networks be able to detect changes in topsoil organic carbon content? *Glob. Chang. Biol.* 14, 2432–2442. <https://doi.org/10.1111/j.1365-2486.2008.01658.x>
- Schillaci, C., Acutis, M., Lombardo, L., Lipani, A., Fantappiè, M., Märker, M., Saia, S., 2017. Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: The role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. *Sci. Total Environ.* 601–602, 821–832. <https://doi.org/10.1016/j.scitotenv.2017.05.239>
- Schillaci, C., Acutis, M., Vesely, F., Saia, S., 2019. A simple pipeline for the assessment of legacy soil datasets: An example and test with soil organic carbon from a highly variable area. *Catena* 175, 110–122. <https://doi.org/10.1016/j.catena.2018.12.015>

- Schmidt, K., Behrens, T., Scholten, T., 2008. Instance selection and classification tree analysis for large spatial datasets in digital soil mapping. *Geoderma* 146, 138–146. <https://doi.org/https://doi.org/10.1016/j.geoderma.2008.05.010>
- Shi, Z., Crowell, S., Luo, Y., Moore, B., 2018. Model structures amplify uncertainty in predicted soil carbon responses to climate change. *Nat. Commun.* 9, 2171. <https://doi.org/10.1038/s41467-018-04526-9>
- Shiwen, Z., Lanlan, Z., Zishuang, L., Qingyun, W., Hongbiao, C., Zhongxiang, S., Chang, G., Huiling, L., Yuanfang, H., 2017. Three-dimensional stochastic simulations of soil clay and its response to sampling density. *Comput. Electron. Agric.* 142, 273–282. <https://doi.org/10.1016/J.COMPAG.2017.08.031>
- Smith, J., Smith, P., Wattenbach, M., Zaehle, S., Hiederer, R., Jones, R.J. a, Montanarella, L., Rounsevell, M.D. a, Reginster, I., Ewert, F., 2005. Projected changes in mineral soil carbon of European croplands and grasslands, 1990-2080. *Glob. Chang. Biol.* 11, 2141–2152. <https://doi.org/10.1111/j.1365-2486.2005.001075.x>
- Smith, P., 2004. Carbon sequestration in croplands: the potential in Europe and the global context. *Eur. J. Agron.* 20, 229–236. <https://doi.org/10.1016/J.EJA.2003.08.002>
- Smith, P., Soussana, J., Angers, D., Schipper, L., Chenu, C., Rasse, D.P., Batjes, N.H., van Egmond, F., McNeill, S., Kuhnert, M., Arias-Navarro, C., Olesen, J.E., Chirinda, N., Fornara, D., Wollenberg, E., Álvaro-Fuentes, J., Sanz-Cobena, A., Klumpp, K., 2019. How to measure, report and verify soil carbon change to realise the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Glob. Chang. Biol.* gcb.14815. <https://doi.org/10.1111/gcb.14815>
- Sommer, R., Bossio, D., 2014. Dynamics and climate change mitigation potential of soil organic carbon sequestration. *J. Environ. Manage.* 144, 83–87. <https://doi.org/10.1016/J.JENVMAN.2014.05.017>
- Sperow, M., 2020. What might it cost to increase soil organic carbon using no-till on U.S. cropland? *Carbon Balance Manag.* 15, 26. <https://doi.org/10.1186/s13021-020-00162-3>
- Szatmári, G., László, P., Takács, K., Szabó, J., Bakacsi, Z., Koós, S., Pásztor, L., 2018. Optimization of second-phase sampling for multivariate soil mapping purposes: Case study from a wine region, Hungary. *Geoderma*. <https://doi.org/10.1016/J.GEODERMA.2018.02.030>
- Vermeulen, S., Bossio, D., Lehmann, J., Luu, P., Paustian, K., Webb, C., Augé, F., Bacudo, I., Baedeker, T., Havemann, T., Jones, C., King, R., Reddy, M., Sunga, I., Von Unger, M., Warnken, M., 2019. A global agenda for collective action on soil carbon. *Nat. Sustain.* 2, 2–4. <https://doi.org/10.1038/s41893-018-0212-z>
- Veronesi, F., Schillaci, C., 2019. Comparison between geostatistical and machine learning models as predictors of topsoil organic carbon with a focus on local uncertainty estimation. *Ecol. Indic.* 101, 1032–1044. <https://doi.org/10.1016/J.ECOLIND.2019.02.026>
- Viola, F., Liuzzo, L., Noto, L. V., Lo Conti, F., La Loggia, G., 2014. Spatial distribution of temperature trends in Sicily. *Int. J. Climatol.* 34, 1–17. <https://doi.org/10.1002/joc.3657>
- Wadoux, A.M.J.-C., Brus, D.J., 2020. How to compare sampling designs for mapping? *Eur. J. Soil Sci.* <https://doi.org/10.1111/ejss.12962>
- Walkley, A., 1935. An Examination of Methods for Determining Organic Carbon and Nitrogen in Soils. *J. Agric. Sci.* 25, 598–609. <https://doi.org/10.1017/S0021859600019687>
- WALKLEY, A., BLACK, I., 1934. AN EXAMINATION OF THE DEGTJAREFF METHOD FOR DETERMINING SOIL ORGANIC MATTER, AND A PROPOSED MODIFICATION OF THE CHROMIC ACID TITRATION METHOD. *Soil Sci.* 37, 29–38.

<https://doi.org/10.1097/00010694-193401000-00003>

Wang, X., Yu, D., Wang, C., Pan, Y., Pan, J., Shi, X., 2018. Variations in cropland soil organic carbon fractions in the black soil region of China. *Soil Tillage Res.* 184, 93–99. <https://doi.org/10.1016/J.STILL.2018.07.010>

Wang, Y., Wang, S., Adhikari, K., Wang, Q., Sui, Y., Xin, G., 2019. Effect of cultivation history on soil organic carbon status of arable land in northeastern China. *Geoderma* 342, 55–64. <https://doi.org/10.1016/j.geoderma.2019.02.007>

Wei, X., Shao, M., Gale, W., Li, L., 2014. Global pattern of soil carbon losses due to the conversion of forests to agricultural land. *Sci. Rep.* 4, 4062. <https://doi.org/10.1038/srep04062>

West, T.O., Post, W.M., 2002. Soil Organic Carbon Sequestration Rates by Tillage and Crop Rotation. *Soil Sci. Soc. Am. J.* 66, 1930. <https://doi.org/10.2136/sssaj2002.1930>

Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. *Biometrics Bull.* 1, 80. <https://doi.org/10.2307/3001968>

Williams, J.N., Morandé, J.A., Vaghti, M.G., Medellín-Azuara, J., Viers, J.H., 2020. Ecosystem services in vineyard landscapes: a focus on aboveground carbon storage and accumulation. *Carbon Balance Manag.* 15, 23. <https://doi.org/10.1186/s13021-020-00158-z>

Zar, J.H., 1999. *Biostatistical analysis*. Prentice Hall.

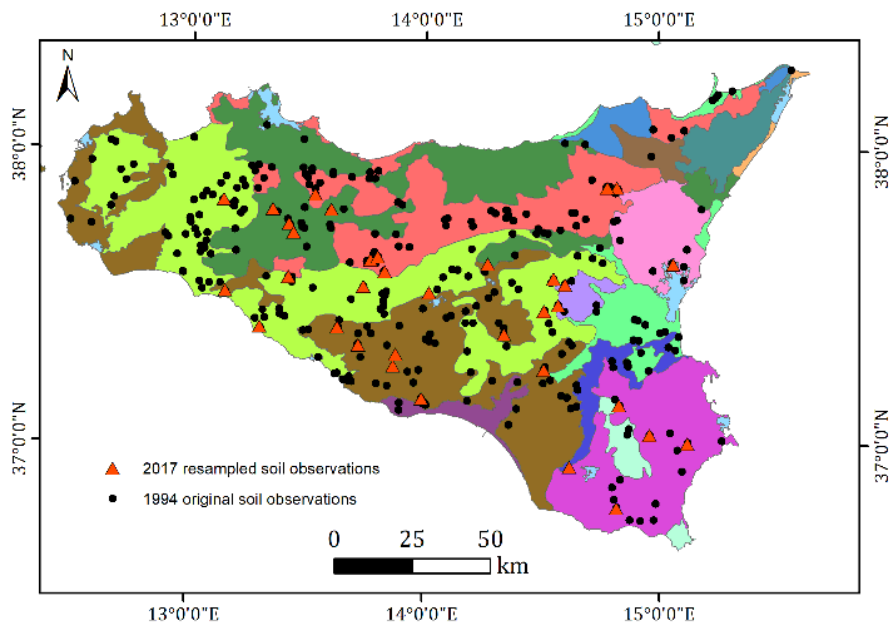
Zhang, C., McGrath, D., 2004. Geostatistical and GIS analyses on soil organic carbon concentrations in grassland of southeastern Ireland from two different periods. *Geoderma* 119, 261–275. <https://doi.org/10.1016/j.geoderma.2003.08.004>

Zhi, J., Jing, C., Lin, S., Zhang, C., Liu, Q., DeGloria, S.D., Wu, J., 2014. Estimating Soil Organic Carbon Stocks and Spatial Patterns with Statistical and GIS-Based Methods. *PLoS One* 9, e97757. <https://doi.org/10.1371/journal.pone.0097757>

Zhou, Z., Zhang, X., Gan, Z., 2015. Changes in soil organic carbon and nitrogen after 26 years of farmland management on the Loess Plateau of China. *J. Arid Land* 7, 806–813. <https://doi.org/10.1007/s40333-015-0051-y>

Zomer, R.J., Bossio, D.A., Sommer, R., Verchot, L. V, 2017. Global Sequestration Potential of Increased Organic Carbon in Cropland Soils. *Sci. Rep.* 7. <https://doi.org/10.1038/s41598-017-15794-8>

Figures



**Soil systems
(Soilmaps.it)**

Map units

- | | | | |
|--|--|---|---|
| <ul style="list-style-type: none"> Calcaric, Eutric, Calcaric Gleyic, Calcaric Endoleptic, and Vertic Cambisol; Calcic Chernozem; Haplic, Leptic, Vertic, and Calcaric Phaeozem; Calcaric Regosol; Haplic Calcisol; Calcic Kastanozem Calcic, Sodic, and Haplic Vertisol; Fluvic and Calcaric Cambisol; Calcic Luvisol, Gypsic Regosol Chromic, Calcic, and Haplic Luvisol; Haplic, Calcic, Chromic, and Hyposodic Vertisol; Haplic Calcisol; Calcaric and Eutric Cambisol; Calcaric Regosol, Calcaric Phaeozem | <ul style="list-style-type: none"> Eutric Leptosol; Andic, Eutric, Thaptoandic, and Calcaric Cambisol; Haplic Luvisol (Vitric); Vitric and Mollic Vitric Andosol; Tephric and Eutric Regosol Eutric Planosol (Sodic); Brunic Arenosol; Gleyic Solonchak; Luvic Phaeozem; Chromic and Leptic Luvisol; Eutric Fluvisol (Arenic); Eutric and Sapric Histosol; Mollic and Calcaric Gleysol; Gleyic Vertic CM; Chromic VR (Grumic); Salic Sodic VR; Calcaric AR Eutric, Calcaric, Dystric, Stagnic, Fluvic, Vertic, and Leptic Cambisol; Calcaric Regosol; Calcaric Leptosol; Haplic Luvisol (Profondic) | <ul style="list-style-type: none"> Eutric, Calcaric, Vertic, and Fluvic Cambisol; Haplic Calcisol; Calcaric Regosol; Haplic, Luvic, Endoleptic, Leptic, and Skeletic Phaeozem; Luvic Kastanozem Haplic Calcisol (Hypercalcic); Mollic Vitric Andosol; Haplic Luvisol (Vitric) Haplic and Leptic Umbrisol (Arenic, Humic); Dystric Cambisol; Umbric Leptosol; Silandic Andosol Haplic and Petric Calcisol; Calcic, Chromic, and Skeletic Luvisol; Calcaric and Luvic Phaeozem; Calcaric Fluvisol; Haplic and Calcic Vertisol; Calcic Kastanozem; Eutric, Fluvic, Endogleyic, and Calcaric Cambisol; Vitric Andosol; Calcaric RG; Calcaric AR | <ul style="list-style-type: none"> Leptic Umbrisol; Dystric Leptic Cambisol; Eutric Regosol; Mollic Leptosol (Vitric); Mollic Vitric and Silandic Andosol Leptic and Luvic Phaeozem; Leptic and Chromic Luvisol; Haplic Calcisol; Calcic Chernozem; Calcaric Regosol; Calcaric Cambisol; Calcic Kastanozem; Calcaric Leptosol; Calcaric Arenosol Mollic, Eutrisilic, Vitric, and Silandic Andosol; Rendzic Leptosol; Eutric Leptic Cambisol (Humic); Eutrisilic Leptic Andosol (Mollic); Haplic Luvisol (Andic); Calcaric Fluvic Cambisol Rhodic, Leptic, Rendzic, Calcic, and Chromic Luvisol Urban areas |
|--|--|---|---|

Figure 1

Study area and main soil systems from the Italian soil information system <http://www.soilmaps.it/>. The samples collected only in 1994 (black dots) and both in 1994 and 2017 (red triangles) were projected on the map.

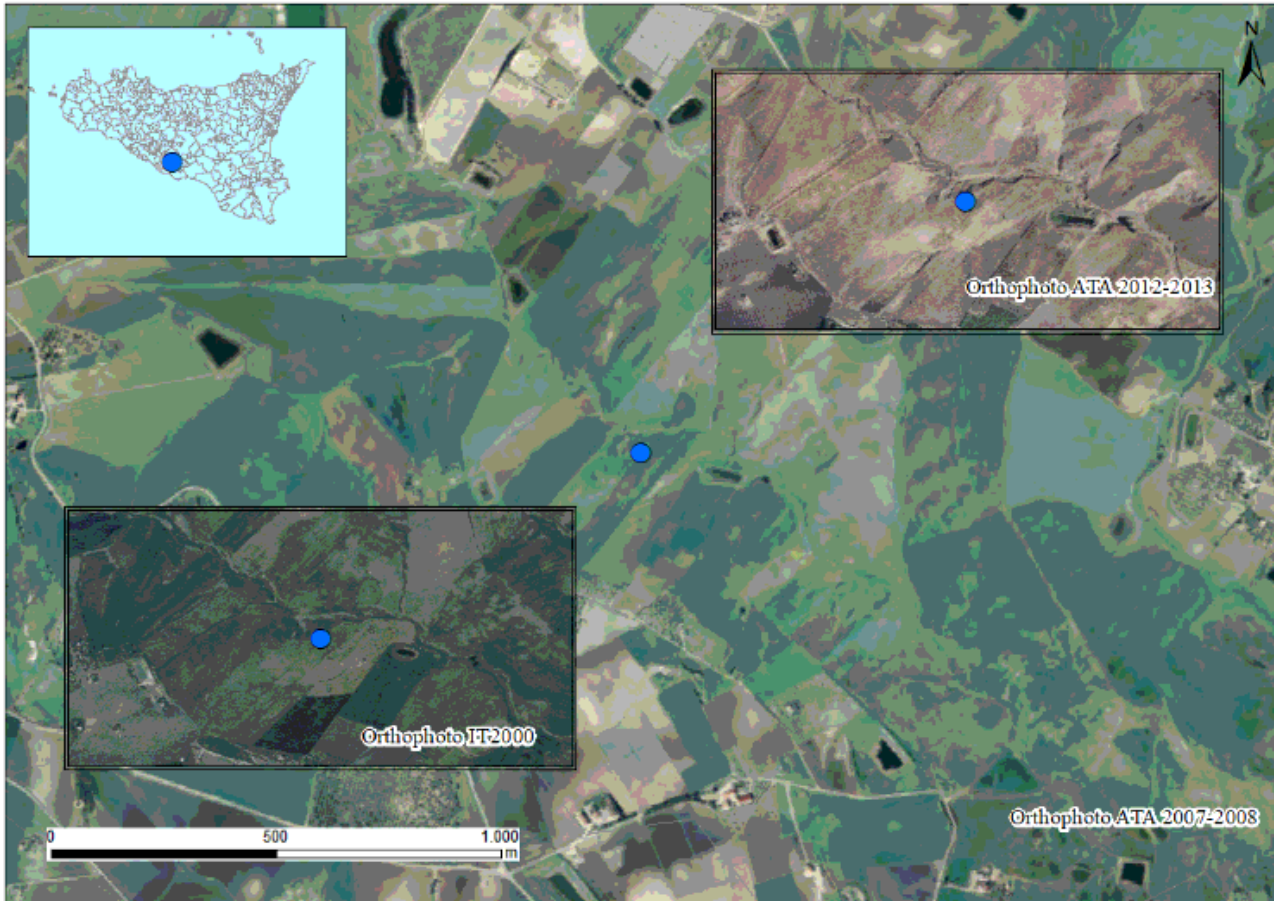


Figure 2

Example of checking of the actual land use along the course from the 1994 to the 2017 in each sampling location n=30) by means of a visual interpretation of the Orthophoto taken from a Web Mapping Services (WMS).

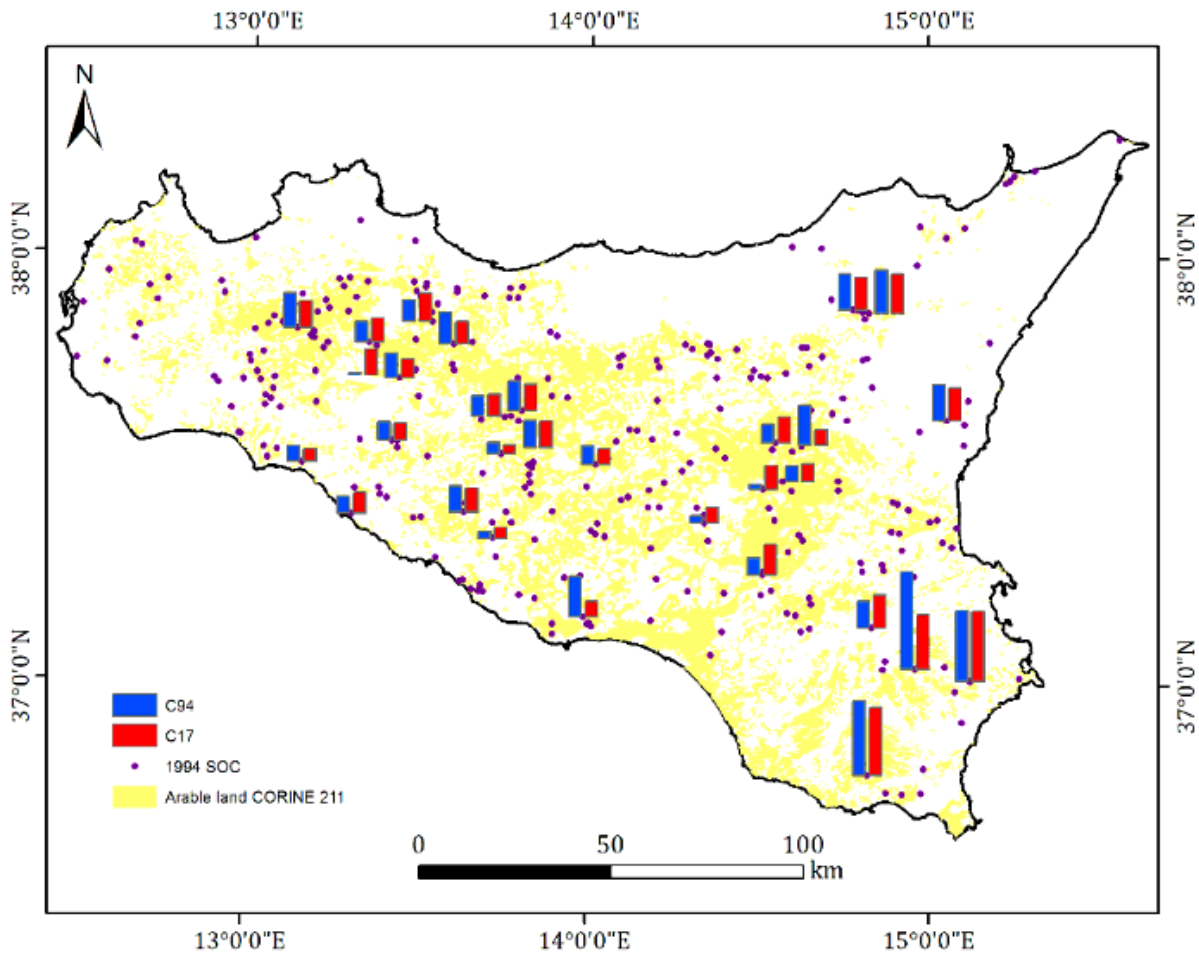


Figure 3

Study area, original soil data (dots) and original compared to new soil samples (histograms, background mask the Arable land cover from CORINE 2000, code 211 (yellow). Blue bars for SOC94, red bars for SOC17.

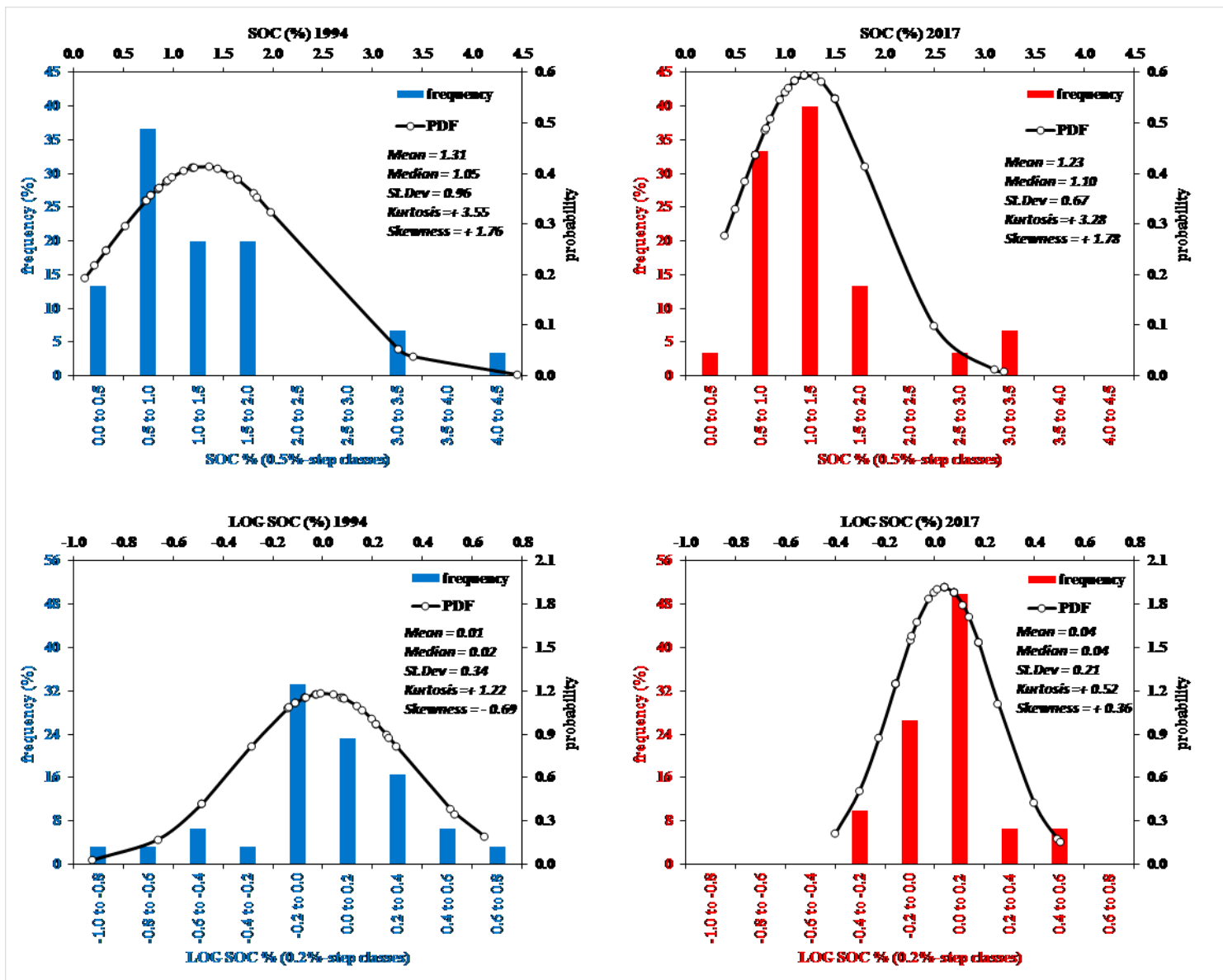


Figure 4

Distribution of the 30 samples of SOC 94 (blue bars) and SOC17 (red bars) and relative probability distribution function. Upper panels are for raw data, lower panels for Logs

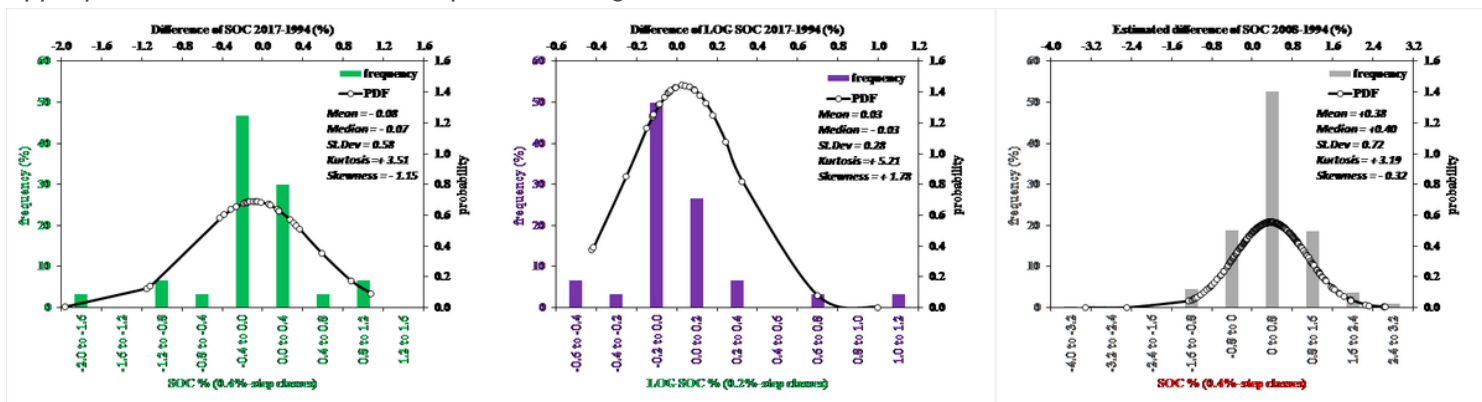


Figure 5

Distribution of the difference between pairs of the 30 samples of SOC17-SOC94 (expressed as raw data, green bars; or their logs, purple bars) and relative probability distribution function. Differences between the 302 original data (LEG-SOC94) and the 302 estimated data of 2008 (EST-SOC08) is also shown (grey bars). Positive values indicate an increase in SOC with time.