Diachronic Mapping of Soil Organic Matter in Eastern Croatia Croplands

Sebastiano Trevisani 1 and Igor Bogunovic 2, *

1 Dipartimento di Culture del Progetto, University IUAV of Venice, Dorsoduro 2206, 30123 Venezia, Italy; strevisani@iuav.it
2 Department of General Agronomy, Faculty of Agriculture, University of Zagreb, 10000 Zagreb, Croatia

* Correspondence: ibogunovic@ag.hr; Tel.: +385-12393815

Abstract: The spatiotemporal analysis and mapping of soil organic matter (SOM) play a pivotal role for evaluating soil health and for implementing preservation and restoration actions. In this context, the first aim of the study is to furnish a high-resolution mapping of current SOM content in eastern Croatia. The second aim is to perform a diachronic analysis of SOM content, comparing two datasets characterized by an extreme data imbalance. The more recent dataset (SOM2010), representative of 2010s, comprises 19,386 samples and the older dataset (SOM1970), representative of the 1970s, comprises 152 samples. The marked data imbalance and the different modalities in soil sampling and laboratory analysis of the two datasets are taken into consideration in performing the comparison. The study reveals a general depletion trend of SOM from the 1970s to the 2010s, more evident in with regard to Fluvisols and Gleysols. At a regional scale, the SOM2010 is characterized by lower variability compared to SOM1970, indicating a process of homogenization of SOM spatial distribution in recent years. Considering the local scale, there is limited information for the 1970s; for the 2010s the SOM spatial distribution is characterized by a high short-range spatial variability, with a characteristic spotty appearance, likely related to agricultural practices.

Keywords: agroecosystem; geostatistics; kriging; soil quality; land degradation

1. Introduction

Soils are essential for life on Earth since they are responsible for the majority (>95%) of food production [1] and play a vital role in climate regulation by carbon sequestration [2]. Nowadays, soils are under extreme pressure due to a growing demand for food, which increases the risk of land degradation due to soil compaction and erosion [3], both closely associated with loss of soil organic matter (SOM) content in soils [4]. Therefore, the spatiotemporal analysis of soil properties in areas of intensive agriculture is crucial for multiple purposes. First, it is essential for the evaluation of the current state of natural resources and for recognizing evolutionary trends. Then, this kind of analysis can be useful to highlight anthropic impacts as well as possible geomorphological [5,6] forcing on soil properties. Moreover, the spatiotemporal analysis and mapping of soil properties play a pivotal role in the context of landscape management, e.g., for designing and implementing preservation/restoration actions [7,8]. Unfortunately, the spatiotemporal analysis of soil properties is a difficult task to achieve, in particular when there is the need to compare the current spatial distribution of soil properties with past conditions. In fact, very often the soil-related information on past conditions is characterized by much lower spatial sampling density and, potentially, higher uncertainty compared to current conditions [4,9–11]. The comparison of present environmental dynamics with older ones is a commonly ill posed problem that characterizes many disciplines, both in the Sciences (e.g., geology, climatology, etc.) as well as in the Humanities (e.g., historical studies). In general terms, spatiotemporal comparative analysis is characterized by a massive “data
imbalance” between current and past conditions [12]. The “imbalance” is relative both to data sampling spatiotemporal density as well as to data typology and quality. Nowadays, accurate information on soil properties can be derived by means of multiple direct and indirect methodologies, permitting spatially dense sampling networks [5,13,14]. For example, soil properties can be measured both by means of collection of samples in the field and further laboratory analysis [10,15,16], as well as by means of proximal/remote sensing technologies and geophysical approaches [17–19]. A different situation is found for soil data information related to past conditions (legacy data), which tendentially become sparser, both in relation to available monitoring technologies as well as in relation to the goals, which differ from current goals, that may underlie sampling campaigns. It is clear that any comparison between present and past conditions should be carried out taking into account the data imbalance to avoid any analysis bias.

To observe a strong data imbalance, it is not necessary to move too far in the past. For example, even considering relatively recent past soil conditions, e.g., moving 50 years back, it is possible to observe a remarkably different spatiotemporal density between current and past sampling campaigns. Such differences are in part related to available technologies, which affect costs and efficiency in soil monitoring; however, these are also related to the different goals underlying soil sampling in past and current time. Generally, legacy data from recent past are derived from soil profiles excavated for pedological soil mapping in the context of national spatial and agricultural planning; differently, current soil datasets are often collected in relation to precision agriculture goals, for example to characterize soil nutrients status, and to provide management recommendations to farmers [13], avoiding diffuse pollution or over-fertilization [20]. However, regardless of the mentioned differences, the methodologies, the instrumentations, and the related accuracies of data from of the 60’s/70’s of the past century are comparable with nowadays standards. However, when the diachronic analysis of soil properties is long-term, e.g., moving back in time for centuries, the information related to past conditions becomes highly fragmentary, uncertain, and fuzzy. In this setting, information on soil or geoenvironmental factors may only be provided by various environmental proxies, for example from “primitive” instrumental equipment or even documental sources [21–24].

The present research, focusing on the evaluation of the state of agricultural land in terms of SOM content in a Croatian region (Slavonia and Baranja), encounters the above cited issues of data imbalance. The current SOM content of agricultural soils of the studied area, covering 6609 km², is sampled with a relatively spatially dense sampling network (19,386 data, collected between 2006 and 2012). This dataset (hereinafter referred to as “SOM2010”) extends significantly preceding studies, which focused on smaller and spatially disconnected areas [5,13]. These studies revealed a general depletion of SOM in the study domain with a main anthropic forcing on SOM spatial characteristics. The geostatistical analysis of the updated dataset permits exploration of a much larger area, furnishing an unprecedented perspective on SOM for the whole region. In addition, a legacy dataset is available, reporting SOM content of soil samples collected in the 1970s (hereinafter referred as “SOM1970”), with a total of 162 samples. The SOM1970 dataset is extremely interesting for retrieving information on past conditions of agricultural land in the region, permitting a diachronic analysis of SOM content. However, this represents a challenging task, considering that SOM1970 has a much lower spatial sampling density than SOM2010, having 1% of the amount of data compared to SOM2010.

This case study provides the possibility of studying the current state of SOM content with high spatial resolution and for a wide area, in a highly anthropized environment, constituted by agricultural land, characterized by a sharp increase of intensive agricultural practices in last 70 years. The Pannonia region is one of the most productive agricultural areas in Central Europe and is under extreme ecological pressure due to intensive cropping practices. The majority of farmers use conventional tillage as practice [25–27], leading to serious soil degradation such as depletion of SOM, sealing, and erosion, as reported in previous works [28–30]. It is well-known that SOM concentrations generally depend on
soil properties, environmental factors, and management practices, such as residue return or tillage intensity [31]. The knowledge of SOM spatial patterns makes it feasible to delineate areas of SOM depletion and to evaluate the proportion of degraded land. Soil organic matter content is an indicator of soil quality and fertility [32] and plays a crucial role in soil aggregation, water retention, nutrient cycling, regulation of greenhouse gases, and susceptibility to soil structural degradation and land degradation [33].

The possibility of mapping SOM spatial distribution at high resolution in such a wide area, highlighting areas of SOM depletion, is a novelty for Croatia and can be relevant for other European countries, especially in the context of soil exploitation. Moreover, the comparison of the current situation with the one depicted in 1960s and 1970s is intriguing, both from the methodological perspective, in regard to the cited data imbalance, as well as for the diachronic analysis of SOM content in a highly exploited area.

The key objectives of the study are hence twofold. The first is to furnish a mapping of SOM content for the study area to evaluate the state of depletion of SOM in the Pannonia region. The second is to perform a diachronic analysis of SOM content, comparing the two datasets characterized by a strong data imbalance. A key question, related to the second objective, is whether this analysis should be limited to simple summary statistics or if it can also shed light on the spatial characteristics of the two datasets.

2. Materials and Methods

2.1. Study Site

The study area covers the territory of two counties in eastern Croatia: Osjeckobaranjska and Vukovarsko-srijemska counties (Figure 1). The counties are located between the latitudes of 19°26′ E to 17°53′ E and the longitudes of 44°51′ N to 45°55′ N, covering 6609 km². According to Köppen’s climate classification, the area is characterized by a moderately warm, wet climate [34]. More precisely, moderate continental climate conditions prevail towards the east, which is characterized by relatively warm summers (mean daily temperature in July of 22 °C) and relatively cold winters (mean daily temperature in January of −3 °C) [35]. The lowest annual rainfall occurs in the eastern part of the area (average annual value of 650 mm), which gradually increases to 900 mm towards the western hilly part of the research area [35]. Most precipitation falls in spring and the mid-summer period.

The study area, located between the Drava, Sava, and Danube rivers, is characterized by a gentle topography, with elevations up to 200 m a.s.l.; mountains are present only in the western portions of the study area (Figure 1) (approximately 600 m a.s.l.). The morphology of the region, with the main valleys and ridges aligned in the NW-SE and WNW-ESE directions, is strongly influenced by the geostructural setting, such as tectonic lineaments marking the scarps at the foothills. The area is characterized by a high heterogeneity of soil types and lithology [5,36–39]. Terrains proximal to the rivers and located at the lowest elevations are mainly composed of Holocene sediments, representative of alluvial deposits of the Drava, Sava, and Danube rivers, and to a lesser extent of lake deposits [5]. The rest of the plain and of the hills are mainly composed of Pleistocene sediments, such as the loess deposits in correspondence of terraces, and to a lesser extent the alluvial and colluvial deposits. Mountains cover a marginal proportion of the study area and are characterized by a prevalence of Paleozoic and Mesozoic rocks.

The geological and geomorphological heterogeneity influences the diversity of soils in the studied area (Figure 1b). The dominant soil types are Gleysols, Luvisols, Cambisols, Chernozem, and Stagnosols [40]. The other soil groups include Albeluvisols and Fluvisols, while Leptosols, Anthrosols, Regosols, and Histosols [40] cover a small proportion of the study area. This region of Croatia is one of the most productive for agriculture. Lowlands are usually dedicated to annual crops; the dominant species used in crop rotation include winter wheat, maize, sunflower, soybean, and sugar beet. Gently hilly areas are covered by permanent plantations, with a prevalence of vineyards and orchards.
The dataset comprises 19,386 georeferenced soil samples (Figure 2) with an average nearest neighbor distance between samples of 147 m (median 164 m). The spatial support [41] of the samples is 900 m², considering that the analysis is conducted on a composite sample, built from the homogenization of 20–25 individual samples, collected in an area of 30 × 30 m. Samples were taken from at depths of 0–30 cm. Soil samples were collected soon after the harvest, before application of fertilizers and other soil amendments. During the field campaign, individual soil samples were homogenized and stored in plastic bags, then were dried, milled, and analyzed in the laboratory, according to the modified Walkley-Black method [42].
2.2.2. SOM1970

The legacy dataset used in this study has been made available from a soil survey related to a soil mapping national project ("Basic Soil Map"), designed for a cartographic scale of 1:50,000, covering all Croatian territory [43]. The "Basic Soil Map" project lasted 23 years, from 1964 to 1986, and several investigators worked on different sections. At the national scale, the sampling is diachronic, and the quality of the data is heterogeneous, considering that different experts worked on the project, sometimes adopting different inventory criteria, classification approaches, and methods. Differently, the study area considered has been mainly sampled in the 1970s, according to a common standard regarding sampling and laboratory analysis. In fact, after 1970, in the area considered in this study, the project coordinators harmonized and standardized survey methodologies and criteria, instructing all involved investigators. The soil samples considered in this study have been selected after a careful analysis of the legacy database, discarding those data lacking a clear description of sampling practices and soil profiles. A further control on legacy data reliability was performed comparing the reported pedogenetic and morphological characteristics with evidence coming from field surveys and laboratory analysis of soil samples. For example, profiles for which field and laboratory analysis findings differ were discarded. Moreover, soil profiles with insufficient location description, which might have reflected in inaccurate georeferencing, were also removed. It is worth noting that the legacy samples are practically representative of a punctual support of measurement, being samples obtained from individual soil profile pits, sampling horizon by horizon. Soil organic matter concentration was determined by Tjurin method [44] and only the concentrations from the topsoil layer were considered. The depth of topsoil layer differs but generally follows the depth of plowing, approximately 25–30 cm.

After the reported analyses and validation of the legacy database, 152 soil profile pits have been retained for further mapping [45]. The spatial distribution of the soil samples is highly irregular and characterized by low spatial density (Figure 2), with an average nearest neighbor distance between the samples of 2595.6 m (median 2955.4 m).

2.3. Geostatistical Analysis and Mapping

Geostatistical analysis of SOM has been conducted, adopting standard approaches for exploratory data analysis (ESDA), variogram modelling, and interpolation. The Gstat library [46] for the R statistical programming environment and the “Geostatistical Analyst” in the ArcMap (version 10.81) GIS environment (ESRI) were used for the analysis. ESDA has been conducted with classed post maps, variogram cloud [47], moving windows statistics [41], directional variograms, and variogram maps [48]. ESDA is fundamental for data error detection, identification of possible local and global outliers, and the presence of spatial non-stationarity, for example related to trends and heteroscedasticity, such as a proportional effect [49]. Moreover, in this phase of analysis it is also worth exploring possible relationships between primary and secondary variables (e.g., soil types, and elevation) both for interpretative reasons as well as for possible use in interpolation [49].

The spatial continuity structure of SOM was analyzed by means of experimental variograms, using variogram maps and directional variograms for the detection of anisotropies. The base lags used in the calculation of experimental variograms of SOM2010 and SOM1970 have been selected considering the average nearest neighbor sampling distance, approximately 150 m and 2500 m. Variogram model parameters were inferred using a weighted least squares approach as implemented in the Gstat software [46]. The selection of the variogram models have been done by expert judgment and, in case of the presence of equally compatible models, using the one with the best-fit diagnostics.

The interpolation has been conducted by means of ordinary kriging algorithm, given the characteristics of the data. The interpolation parameters, including a variogram model and neighboring search criteria, have been evaluated by means of cross-validation [41,49]. In particular, standard leave-one-out cross validation has been used to evaluate basic diagnostics, including root mean square error (RMSE), average standard error (ASE),
and statistical and spatial distributions of the errors. The resolution and accuracy of the interpolated maps are dependent on the spatial variability and sampling geometry of the two datasets.

2.4. Comparison between SOM2010 and SOM1970 Datasets

The statistical and spatial statistical comparison between SOM2010 and SOM1970 have been conducted by means of various approaches, taking into consideration the extreme difference in the number of data between the two datasets, with SOM1970 totaling 0.8% of the data of the SOM2010 dataset. All the tests have been performed using the R statistical programming environment. Regarding the differences in the mean and overall distributions, usual nonparametric tests as well as a resampling approach have been considered.

One-sided and two-sided two-sample Wilcoxon test (i.e., Mann-Whitney) and Kolmogorov-Smirnov tests have been performed, using the R algorithms of the “Stats” package. Then, the much higher numerosity of SOM2010 dataset permitted us to adopt a resampling approach to test various statistics and spatial statistics. The SOM2010 dataset can be considered almost exhaustive compared to SOM1970, and its statistics can be considered as the reference. Accordingly, it is reasonable to explore how likely it is to draw a subsample (with an amount of data comparable to SOM1970) from SOM2010 with similar statistical characteristics observed in SOM1970s dataset. This can be achieved easily by means of resampling methods, randomly drawing subsamples from SOM2010 with a subsampling ratio of 0.8% and calculating from these subsamples various statistical and spatial statistical indices. The median, the variance, and the interquartile range have been tested; concerning spatial-statistics, sample variogram values for a set of lags have been tested.

To take into account the impact on variability due to the different spatial support, the usual approach for variogram regularization has been adopted. In particular, the regularization of the variogram has been derived by calculating a corrective factor represented by the integral of the within-block punctual variogram.

3. Results and Discussion

3.1. SOM Content

The summary statistics of SOM content are reported in Table 1 and their frequency distribution in Figure 3. The SOM1970 and SOM2010 have a wide range of variation, ranging respectively from 0.50% to 11.70% and from 0.32% to 11.00%. The mean and median SOM1970 concentration is 3.25% and 2.4%, while for SOM2010 is 2.11% and 1.97%. Usually, a concentration of 3.4% of SOM is considered the threshold below which a soil can be considered degraded in a temperate environment [31], negatively affecting yields in different textured soils [52]. Accordingly, this can indicate that already in the 1970s soil degradation processes impacted the studied pedosphere. Regardless of the soil type and geomorphology, the study region is recognized to be vulnerable to soil degradation [53]. Several locations are characterized by soil water erosion [30,54] and compaction [25,28,55], and are characterized by decreasing yields [56] due to unappropriated land use and soil management.

Table 1. Descriptive statistics of the SOM datasets. Min—minimum, Max—maximum. Different letters represent significant differences at \( p < 0.05 \).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>n</th>
<th>Spatial Spacing (m)</th>
<th>Mean</th>
<th>Median</th>
<th>Min</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOM1970</td>
<td>152</td>
<td>2595.6</td>
<td>2955.4</td>
<td>0.50</td>
<td>1.87</td>
<td>2.40</td>
<td>3.25 ( \text{a} )</td>
<td>4.07</td>
<td>11.70</td>
<td></td>
</tr>
<tr>
<td>SOM2010</td>
<td>19,386</td>
<td>147</td>
<td>164</td>
<td>0.32</td>
<td>1.66</td>
<td>1.97</td>
<td>2.11 ( \text{b} )</td>
<td>2.40</td>
<td>11.00</td>
<td></td>
</tr>
</tbody>
</table>

Regarding current SOM content (SOM2010) in the different soil types, the highest median values correspond with Chernozems and Gleysoils, while the lowest ones are correspond with Luvisols and Albeluvisols (Figure 4). On one hand, this result is expected considering that Chernozems are among the most fertile soils [57] and are usually managed
intensively, with the adoption of manure and mineral fertilizers. On the other hand, higher SOM in Gleysoils, and to a minor extent in Fluvisols, could be explained by the fact that these soils are mainly located in lowlands near the rivers [38]. Although a significant part of these soils are drained [5], their general proximity to the rivers (Drava, Sava, and Danube) (Figure 1b) corresponds to a shallow water table, inhibiting the decomposition rate and supporting the accumulation of SOM. Also, Cambisols have relatively high SOM, likely due to intensive soil mineral fertilization. Cambisols naturally continue from the Chernozem zone and are recognized as the fertile soil, usually widely exploited for intensive plant production [58]. Luvisols, Albeluvisols, and Stagnosols have low SOM concentration, confirming the results of previous works [36,59] in the same region. Bogunovic et al. [5] explain their low SOM content by natural process, since they are still under formation.

Concerning SOM content from the 1970s (SOM1970) in the different soil types, some soil types have marked differences compared to the present status (Figure 4). In particular, the Fluvisols and Gleysoils appear to be the soils with highest SOM concentration, with remarkably higher concentration compared to the other soil types. This could be related to the fact that in the 1960s and 1970s these soils were not highly exploited, because soil drainage was not implemented in those areas. Relatively high SOM is encountered, in Chernozems, Cambisols, and Albeluvisols. As for SOM2010, the lowest concentrations are in Luvisols and Stagnosols.

The difference of the SOM concentrations between SOM2010 and SOM1970 datasets could be partly attributed to the lower number of samples in the SOM1970 dataset and possibly to the different spatial support. However, the differences could be explained by other factors, and could be attributed to the different management practices, and socioeconomic changes that occurred in the last several decades. During the 1970s the study area was a part of the former federal State of Yugoslavia, and the government had a centralized and rigid government system. Croatia was one of the federations with reduced agricultural development. In the socialist period the productive sectors in agriculture comprised state (social) and private entities (mainly family-owned farms). The productive chain was mainly based on traditional agricultural production methods, characterized by low productivity and low marketability potentials [60]. On the other hand, the former

Figure 3. Frequency distributions of SOM2010 (a) and SOM1970 (b) datasets, and their box-plots (c) and Q-Q plots (d).
socialist state incentivized state enterprises with relevant investments, to highlight the supposed advantage of so-called socialistic model of agriculture over the private one. This did not lead to a significant improvement in productivity at the state level. In this regard, it is interesting to highlight that the average parcel size in 1969 was 3.3 ha [61], and that only in the 2010s did the average size increased to 5.5 ha [62]. Although this average size is still low compared to the EU average, the increased parcel size reveals a recent shift in the productive chain. Moreover, agricultural practices in the end of 20th and in the beginning of the 21st century improved. Private farmers invested in equipment, fertilizers, and protection agents, which increased the biomass production and yields and, at the same time, the overexploitation of soils.

Figure 4. Boxplots of SOM content for SOM2010 (red boxplots) and SOM1970 (grey boxplots), grouped according to soil types. The boxplots depict the minimum, maximum, median, as well as the upper and lower quartiles of the data for each soil class. The horizontal line is in correspondence of the threshold of 3.4% in SOM content.

In general, the statistical differences between SOM1970 and SOM2010 dataset are quite evident. From the analysis of statistical summaries and graphics, such as histograms and boxplots (Figures 3 and 4), SOM1970, with respect to SOM2010, had higher mean/median values and higher variability. The two distributions are likely to be from a different population, as confirmed both by two-sided Kolmogorov-Smirnov test ($D = 0.29418$, $p$-value $= 9.202 \times 10^{-12}$) as well as by the two-sided Wilcoxon rank sum test ($W = 974,532$, $p$-value $= 5.964 \times 10^{-13}$). When the tests are performed as one-sided tests, these confirm the general depletion of SOM2010 with respect to SOM1970.

A further statistical comparison between the two datasets has been performed via a resampling approach, drawing 10,000 subsamples with a subsampling ratio of 0.8% from the SOM2010s dataset. For each subsample, statistical indices of central tendency (median value) and dispersion (interquartile range and variance) have been calculated for each sample. As reported in Figure 5, the results confirm the differences in the statistical structures between SOM2010 and SOM1970 datasets. It is quite interesting to note the much lower variability of all SOM2010 subsamples respect to SOM1970 dataset. The general depletion in SOM induces a homogenization of SOM content in current times, not justified by the different support size.
Figure 5. Boxplots report the statistical distribution (including extremes) of the statistical indices (median, interquartile range, and variance) calculated from each of the 10,000 subsamples (subsampling ratio 0.8%) generated from the SOM2010 dataset. The horizontal thick black lines represent the respective statistical index (i.e., median, interquartile range, and variance) calculated from the SOM1970 dataset.

To analyze this topic further and to consider the different support size of measurements, experimental variograms have been calculated from the subsamples, permitting us to analyze the differences in spatial variability. Omnidirectional variograms (Figure 6) have been computed with a base lag of 2500 m, corresponding to the average spacing of SOM1970 dataset and of subsamples, up to a distance of 25 km. Given the differences in support size between the two datasets, punctual for SOM1970 and with a spatial support of 900 m² for SOM2010, the potential impact of variogram regularization [51] has been taken into account. Considering an exponential model fitted to the experimental variogram calculated from SOM1970 (see next section) and a spatial support of 30 × 30 m, the integral of the within-block punctual variogram is approximately 0.6%². Accordingly, the differences in variances and variograms values between the two datasets cannot be ascribed only to differences in the spatial support of measurement. Only for the first two lags, considering the corrective factor, could there be some SOM2010 subsample with comparable variogram values to SOM1970. For all the remaining lags there is no overlap. It is worth noting that the estimation of the correction factor for regularization is dependent on the estimation of punctual variogram and that the impact of nugget effect could be quite high.

Figure 6. Boxplots reporting the statistical distribution (including extremes) of variogram values computed for each lag from the 10,000 subsamples (subsampling ratio 0.8%) drawn from SOM2010 dataset. In red variogram values for SOM1970, including the regularization factor for SOM1970 variogram.
3.2. Mapping SOM2010 and SOM1970

For both the datasets, at least for the distances considered, there is no evidence of trend or of the necessity to adopt data transformation procedures such as box-cox or logarithmic procedures. The variogram analysis for SOM2010 reveals, for distances beyond 5000 m, the presence of a slight zonal anisotropy with a direction of maximum continuity approximately along the 112° direction. This anisotropy in the experimental variogram is interesting from the interpretative point of view, because it reflects the morphological and geological structure of the area, with clear lithological and morphological features elongated in the NW-SE direction (Figure 1). However, from the perspective of interpolation, for this dataset the modeling of the anisotropy does not improve interpolation diagnostics, and conversely it induces some interpolation artifacts, represented by elongated features, in areas characterized by low data density. The negligible impact of anisotropy in interpolation is related to the high sampling density and to the fact that most of the interpolation weights are given to data near the interpolation location, and consequently at distances slightly affected by the anisotropy. Accordingly, the final decision was to model the experimental variogram with an isotropic exponential model (Figure 7). Given the absence of trend, the ordinary Kriging algorithm has been adopted, with a search radius of 7000 m. The interpolation has been conducted with two block size dimensions, 250 m (Figure 8) and 2500 m (Figure 9). The former, at the price of higher interpolation uncertainty and variability in interpolation variance, permits us to highlight fine-scale spatial patterns; the latter permits us to map at a lower resolution but with lower and more spatially homogeneous interpolation variance. Moreover, at this resolution, it is possible to perform a comparison with the map obtained for SOM1970 (see next in the text).

Figure 7. Variogram cloud, variogram map, and best fit of omnidirectional variogram models for SOM2010 and SOM1970.
Figure 8. Spatial distribution map of SOM2010 via ordinary block kriging with a spatial support of 250 m (areas with a prediction standard deviation >0.5% have been blanked).

The experimental variogram analysis for SOM1970 dataset (Figure 7) reveals the possible presence of an anisotropy in NW-SE direction, analogously to SOM2010; however, the few data available do not permit us to model this anisotropy. The absence of trend permits the use of an Ordinary block Kriging approach (search radius 20 km). Given the low data density it is reliable only to map SOM using the larger block dimensions of 2500 m (Figure 9). The cross-validation diagnostics are reported in Table 2, indicating, as expected, a much higher interpolation uncertainty compared to SOM2010.

<table>
<thead>
<tr>
<th></th>
<th>SOM2010</th>
<th>SOM1970</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>0.41</td>
<td>1.66</td>
</tr>
<tr>
<td>Absolute Error</td>
<td>1.03</td>
<td>1.06</td>
</tr>
<tr>
<td>Minimum</td>
<td>−8.30</td>
<td>−5.63</td>
</tr>
<tr>
<td>1st Qu.</td>
<td>−0.16</td>
<td>−0.68</td>
</tr>
<tr>
<td>Median</td>
<td>0.03</td>
<td>0.18</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>3rd Qu.</td>
<td>0.21</td>
<td>0.85</td>
</tr>
<tr>
<td>Maximum</td>
<td>6.10</td>
<td>4.47</td>
</tr>
</tbody>
</table>

Table 2. Cross-validation diagnostics for SOM2010 and SOM1970 datasets.
Figure 9. Spatial distribution map of SOM1970 and SOM2010 data using block kriging, with blocks of 2500 m (areas with a prediction standard deviation >0.5% have been blanked).

Due to the marked differences in interpolation standard deviations, it is not feasible to quantitatively compare the two maps; in fact, in many areas the differences between the two maps would be comparable with the interpolation standard deviations. However, the visual analysis of the maps highlights the areas with stronger depletion of SOM in current times. Moreover, the SOM1970 map is characterized by more structured spatial patterns, with a clear tendency of high SOM values near the forested areas. A quantitative spatial comparison between present and older data can be performed, interpolating by means of ordinary kriging SOM2010 data in the same locations as SOM1970s samples (Figure 10). This permits us to obtain, in correspondence of SOM1970 locations, an estimate of current SOM values with a relatively low prediction standard deviation, given the high spatial density of SOM2010 dataset. The correlation between the SOM2010 interpolated data and SOM1970 data is quite low (slope 0.28, standard error for slope 0.02, R² 0.2) but confirms the general depletion of SOM for the current dataset. The map of the differences between SOM1970 and SOM2010 interpolated values, filtering out the data with a too high interpolation standard deviation, permits us to spatially analyze the differences. The data are too sparse to provide a clear spatial pattern of depletion; however, visible signs of intensive soil management during past decades clearly negatively affect the SOM. Farmers mostly use narrow crop rotation, favoring the cash crops over the grasses and legumes [53]. Simultaneously, despite the fact that crops need for tillage intervention, more than 90% of farmers perform conventional tillage [26], involving plowing at least once per year, negatively affecting the aggregate stability and oxidation of SOM. Other lands suffer from SOM depletion during performed amelioration activities. These are mainly adopted in lowland areas, which partly correspond with the evident areas of SOM depletion in Gleysoils and Luvisols (Figure 4). These two soil types are mainly situated in lowlands, which during the 1970s still suffered from high water tables. In the decades since, a great part of these lands have been subject to hydro—and agro—amelioration interventions, which enabled intensive plant production. The water table was lowered by drainage pipes and tillage management appeared, annually accelerating land use changes from natural grasslands to croplands, favoring SOM losses after transition [63,64]. Other soil types show a less marked decrease of SOM compared to Gleysoils and Fluvisols. This is partially related to the overexploitation of these soil types over millennia in the studied region [65]. As reported, farmers still use conventional agricultural practices negatively affecting SOM
content; moreover, pressure for increasing farmer’s incomes leads to the abandoning of proper crop rotations or annual use of cover crops. Decrease of SOM in conventionally tilled soils is a continuous but slow process; the intensive cultivation of virgin soils leads to a marked drop in SOM concentration [66]. This can explain the difference in SOM depletion of Gleysoils and Fluvisols compared to other soil types.

Figure 10. Map of differences between SOM1970 and SOM2010 interpolated values with classes according to quantile classification. Interpolated SOM2010 values with a prediction standard deviation higher than 50% of computed differences have been excluded.

4. Conclusions

The high spatial density of SOM2010 dataset permitted us to map at relatively high resolution and with confidence the SOM concentration in the studied area, highlighting SOM depletion patterns in the Pannonia region. The comparison between current soil conditions and the ones of the 1970s is challenging, given the strong data imbalance between the two datasets. In term of statistical indexes, including spatial variability indexes, it is possible to assess, with a reasonable level of confidence, a general depletion of SOM and a reduction in spatial variability from the 1970s to current times. It is also possible to highlight that the stronger depletion trends are observed in Gleysoils and Fluvisols. The low spatial density of SOM1970 dataset limits the information that can be extracted for the diachronic mapping of SOM1970 and SOM2010. Notwithstanding, the spatial interpolation of the two datasets permits us to shed some light on the depletion patterns in the region and complements the information provided by statistical and spatial statistical indexes. At the regional scale the current SOM spatial distribution, with respect to the 1970s, appears more homogenous. At the local scale little can be said about the SOM1970s spatial patterns; the current ones are characterized by a high short-range spatial variability, with a characteristic spotty appearance, likely related to anthropic processes. Future developments of this study could be based on the analysis of land use changes and historical data concerning agricultural practices in the area. In a context with limited amount of soil data representative
of past conditions, a multidisciplinary approach based on historical analysis could shed further light on the spatiotemporal patterns of soil degradation in the area.

**Author Contributions:** Conceptualization, S.T. and I.B.; methodology, S.T.; software, S.T.; investigation and collection of samples I.B.; writing—review and editing, S.T. and I.B.; funding acquisition, I.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Croatian Science Foundation through the project “Soil erosion and degradation in Croatia” (UIP-2017-05-7834) (SEDCRO).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to their current usage for the preparation of another draft for publication.

**Acknowledgments:** Danijel Rašić is acknowledged for his cooperation for providing and preparation the SOM2010 dataset.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**References**

3. FAO; ITPS. *Status of the World’s Soil Resources (Main Report)*; No. 608; Food and Agriculture Organization of the United Nations: Rome, Italy, 2015.


54. Telak, L.J.; Pereira, P.; Bogunovic, I. Management and seasonal impacts on vineyard soil properties and the hydrological response in continental Croatia. *Catena* 2021, 202, 105267. [CrossRef]


64. Wang, X. Changes in Cultivated Land Loss and Landscape Fragmentation in China from 2000 to 2020. *Land* 2022, 11, 684. [CrossRef]
