Journal Pre-proof

Farmland use data and remote sensing for ex-post assessment of CAP environmental performances: An application to soil quality dynamics in Lombardy

Davide Fugazza, Giacomo Aletti, Danilo Bertoni, Daniele Cavicchioli

PII: S2352-9385(22)00031-3

DOI: https://doi.org/10.1016/j.rsase.2022.100723

Reference: RSASE 100723

To appear in: *Remote Sensing Applications: Society and Environment*

Received Date: 21 September 2021

Revised Date: 8 February 2022

Accepted Date: 5 March 2022

Please cite this article as: Fugazza, D., Aletti, G., Bertoni, D., Cavicchioli, D., Farmland use data and remote sensing for ex-post assessment of CAP environmental performances: An application to soil quality dynamics in Lombardy, *Remote Sensing Applications: Society and Environment* (2022), doi: https://doi.org/10.1016/j.rsase.2022.100723.

Conceptualization: all authors

Methodology: all authors

Software: D.F. and G.A.

Formal Analysis: all authors

Writing: original draft: D.F., D.B. and D.C.

Writing: review and editing: D.F., D.B.and D.C.

Project administration: D.B. and D.C.

Funding acquisition: D.B. and D.C.

Farmland use data and remote sensing for ex-post assessment of CAP environmental performances: an application to soil quality dynamics in Lombardy

Davide Fugazza[1*], Giacomo Aletti[1], Danilo Bertoni[1], Daniele Cavicchioli[1]

1 Department of Environmental Science and Policy, Via Celoria 2, Milano, Italy

*Corresponding author: davide.fugazza@unimi.it

Abstract

With the reform of the European Common Agricultural Policy (CAP) in 2013, subsidies to farms are now bound to the fulfilment of environmentally friendly measures, such as crop diversification and allocation of a share of their farmland to Ecological Focus Areas, the so-called "Greening requirements". Research on the effects of these policy changes so far have focused mainly on land use transition; however, a detailed investigation of how CAP greening affects the properties of agricultural land is required to assess the actual environmental benefits of the reform. In this study, we present a first attempt to assess the impacts of CAP greening on selected soil quality indices in Lombardy, a populated region in northern Italy where high-intensity agriculture is widespread. We combine high resolution (10/30 m) soil indices from remote sensing based on Landsat-8 and Sentinel-2 data with a regional administrative database covering all agricultural parcels of the region. We then perform a correlation analysis to investigate whether and to what extent greening prescriptions affect the soil quality indices from 2014, representing pre-greening conditions, to 2017, representing greening conditions after 3 years of implementation. Our analysis indicates a high persistence of soil quality indicators and suggests that some crops might have a significant impact on soil quality dynamics, along with farm's compliance with CAP greening. Although we identified some uncertainties in the soil indices, by integrating a large volume of data and an efficient processing algorithm our method paves the way for ex-post environmental performance assessment of agricultural policies.

1. Introduction

Soil properties are affected by agricultural activity through farming practices and land use choices. Farmer's decisions on crop allocation and land management may be influenced by various factors such as selling prices of agricultural products and cost of inputs (Glenk et al., 2017), pedoclimatic conditions (Leteinturier et al., 2006) along with risk preferences of the farmer itself (Latawiec et al., 2017; Paut et al., 2020). Many of the agricultural drivers that impact, directly or indirectly, on soil management and properties depend, in turn, on external factors, such as agricultural policies (Kremmydas et al., 2018).

In Europe, in particular, the Common Agricultural Policy (CAP) has affected and shaped land use choices and land management practices (Posthumus and Morris, 2010, Viaggi et al., 2013). In general, the CAP has shifted over time from productivity-incentive measures toward a more environmentally-friendly regulatory framework. Along with this evolution, there has been a growing interest in measuring and monitoring its effects on land use, coverage and soil properties (Tóth et al., 2013; Orgiazzi et al., 2018; JRC, 2020) and on soil conservation (Borrelli et al., 2016; Borrelli and Panagos, 2020). Referring to soil quality and management, the recent switch of CAP monitoring and assessment, "from compliance to performance", calls for a stronger integration of available data sources. The term "from compliance to performance" refers to a shift in policy tools monitoring. The former (compliance) targets the adoption of farming practices or technologies as a proxy for desired environmental outcomes, while the latter (performance) targets those outcomes directly (Rossi, 2018; Carey, 2019; Ehlers et al., 2021).

While soil properties can be monitored through field sampling and laboratory analysis (see e.g. Paz-Kagan et al., 2014; Orgiazzi et al., 2018), this is a laborious, costly and time-consuming process (Mulder et al., 2011). A more efficient alternative is to combine field sampling with remote sensing to produce spatial maps of soil properties, exploiting the wide area coverage of optical and radar sensors and their ability to monitor the topsoil (Shoshany et al., 2013). Usually, the soil properties of interest are measured in the laboratory and a prediction model is built to relate field and remote sensing data, using multivariate regression (Douaoui et al., 2006; Gholizadeh et al., 2018) or machine learning techniques (Bachofer et al., 2015; Castaldi et al., 2019). While this approach can greatly extend the scope of traditional soil surveys, it still relies on accurate field data collection (Castaldi et al., 2018), and results in a static map of soil properties. Remote sensing has also been used directly in the context of CAP monitoring. However, most research has only approached the issue of assessing compliance with the CAP requirements, either by investigating different supervised classification schemes for the identification of land use types at the parcel scale (Sitokostantinou et al., 2018), or by addressing the automatic extraction of satellite data and their integration into a spatiotemporal framework (Rousi et al., 2020). Estimating the effects of the CAP reform however is a more complex issue, and one that has been relatively overlooked from a remote sensing perspective. Previous efforts have been made particularly in the field of soil erosion monitoring. For instance, Borrelli and Panagos (2020) developed an indicator to estimate soil erosion in the EU, which could be used to evaluate the effects of policy changes; however, their approach is based on coarse resolution satellite data (MODIS, 500 m resolution). In summary, the evolution of soil properties, and the link with agricultural policy changes, are seldom investigated using remote sensing at a high level of detail (Sheffield and Morse-McNabb, 2015; Dube et al., 2017).

While in this context the applications of remote sensing have so far been limited, the environmental impacts of CAP instruments have been widely investigated from different points of view (Ehlers et al., 2021). In some studies, detailed scale assessments of CAP environmental effects have been carried out based on field experiments (Kleijn and Sutherland, 2003; Walker et al., 2007; Ansell et al., 2016), or ad-hoc surveys on a limited number of farms (Zahm et al., 2008; Paracchini et al., 2015). This kind of evaluation suffers from the difficulty of extending the results to wider territorial areas. Other studies instead rely on datasets provided by public administrations (Pufahl and Weiss, 2009; Chabé-Ferret and Subervie, 2013; Bertoni et al., 2020; Bertoni et al., 2021). In particular, they represent ex-post assessments on the effects of participation in the agri-environmental measures of the CAP (Pufahl and Weiss, 2009; Chabé-Ferret and Subervie, 2013; Bertoni et al., 2020), or on the effect of greening and CAP direct payments on farmland allocation, GHG emissions and use of farm inputs (Arata and Sckokai, 2016; Coderoni and Esposti, 2018; Bertoni et al., 2021). In these latter studies, the Farm Accountancy Data Network (FADN)[1] is usually exploited in the assessments of CAP effects (Mary, 2013; Arata and Sckokai, 2016; Coderoni and Esposti, 2018). However, being conceived to evaluate the economic performance of farms, it reports a limited amount of information on environmental performance, usually limited to the inputs (pesticides and fertilizers) and water consumption (Kelly et al., 2018).

The main bottleneck of such assessment studies is the lack of data having, at the same time, a reasonable detail (at least farm level) and full territorial coverage (Primdahl et al., 2010). Indeed, the availability of indicators measured at a detailed spatial scale represents one of the main challenges in assessing the sustainability of the agricultural sector (Burkhard et al., 2009). A further challenge is represented by the extensibility of the results to a larger scale (Kelly et al., 2018). This condition is

[1] The FADN is a sample dataset of about 80,000 farms, representing about 5 million EU farms that concentrate 90% of EU agricultural production (Kelly et al., 2018)

guaranteed when, as in our case, available data cover the universe of observations on a given territory. For the above-mentioned reasons, ex-post CAP impact assessments could be usefully complemented by using geo-referenced data with wide spatial coverage, deriving from remote sensing.

In line with the performance-measurement of policy tools, this paper describes a first attempt to assess whether and to what extent the crop-mix change induced by the greening measures (Bertoni et al., 2018; Micheletti et al., 2020; Bertoni et al., 2021) may be associated with a change in soil quality indicators. These indicators were obtained from spectral indices obtained from high resolution satellite data (Landsat 8 OLI and Sentinel-2, at 30 m and 10 m spatial resolution), in view of their simplicity and low sensitivity to atmospheric conditions. Such indicators are chosen as proxies for specific soil health policy targets (European Commission, 2021a). Our analysis further seeks to demonstrate the feasibility and benefits of integrating administrative and remote sensing data into agricultural policy ex-post evaluation processes. Administrative datasets ensure full territorial coverage and a high level of detailed information (down to the parcel level). On the other hand, they contain a limited amount of information, especially on environmental indicators (often limited to land use or participation in voluntary agro-environmental measures), which can be filled by remote sensing.

With our analysis, we aim at testing how different factors, including farmland use type, eligibility and compliance with respect to greening policies, contribute to explaining the dynamic variability in soil properties derived from remote sensing indices. Our aims are: 1) To develop a methodology to combine high resolution remote sensing and administrative agricultural datasets in an efficient manner, to provide dynamics maps of soil properties in a wide agricultural region of the EU without the need for field surveys; 2) To undertake a preliminary assessment of the evolution of soil properties before and after the implementation of greening policies; 3) To perform a correlation analysis to evaluate the possible influences on the observed changes in soil properties in the context of CAP greening.

2. Common Agricultural Policy and greening elements

This paragraph describes evolution and functioning of CAP prescriptions (in particular greening payments), whose effects are tested in the present article.

In a first stage, which lasted about 30 years, the CAP was mainly based on interventions aimed at supporting selling prices of agricultural products, in order to stimulate farm productivity and ensure food self-sufficiency of European Countries. Subsequently, agricultural support moved from selling prices to per-hectare payments decoupled from farmland use choices (Garzon, 2006; Folmer et al., 2013). Such payments have been tied to the fulfilment by farms of practices respectful of the land and the environment. According to the so-called cross-compliance, farms were obliged to keep their land in Good Agronomic and Environmental Conditions (GAEC) in order to receive CAP payments (Bennett et al., 2006; JRC, 2019a; JRC, 2019b). With the current CAP reform (2015-2020) the requirements to obtain farm payments have increased, with the introduction of the so-called greening payments (European Commission, 2020). Greening payments represent the main part of decoupled payments for farms, provided with the purpose of encouraging farmers to make a sustainable use of their land. In order to receive greening payments, farmers are required to respect cross compliance (GAEC), to adopt crop diversification and to devote a portion of their farmland to Ecological Focus Areas[2].

Although the European Union has been approaching the issue of sustainable soil management for some years[3], until now this has not been an explicit objective of the CAP. However, many elements indicating attention to the issue of soil protection also emerge in the current CAP 2015-2022 framework, with particular reference to the wide set of policy instruments directly, or indirectly, aimed at addressing sustainable soil management. In fact, many of GAECs and voluntary agri-environmental measures in the second pillar are directly related to soil quality issues. The Reg. (EU) 1307/2013, establishing the greening payment, states that "the obligations relating to crop diversification should be applied in a way that takes into account the difficulty for smaller farms to diversify, while continuing to make progress towards enhanced environmental benefit, and in particular the improvement of soil quality" (recital 41, Reg. EU 1307/2013).

Definitely, the future CAP reform (2023-2027) addresses the issue of soil health more incisively than in the past. In fact, efficient soil management is formally stated as a CAP specific objective. The new CAP framework addresses soil protection to face threats to the health of European soils, such as erosion, organic matter decline, biodiversity loss, soil compaction, and contamination, salinisation, and desertification (European Commission, 2021a). New instruments have been put in place to tackle sustainable soil management. The main one is the shift from crop diversification (a static concept) to crop rotation, as a mandatory requirement to receive farm payments. Soil cover, tillage management,

---

[2] Some examples of Ecological Focus Areas are: land lying fallow, landscape features, buffer strips, agroforestry, areas with short rotation coppice, areas with catch crops and green cover, and areas with nitrogen-fixing crops.

[3] Since 2006 the EU, by the Soil Thematic Strategy [COM(2006) 231], set for the first time its general objectives to preserve soil functions . This should have been followed by a soil framework directive, which has never been approved because of the high implementation costs and excessive difference in national legislation (European Commission, 2021b). More recently, the 7th Environmental Action Programme of the EU [Decision 1386/2013/EU] stated that the general objective is to 'promote a sustainable use of soil', establishing potential causes of degradation and suggesting actions to contrast them (Ronchi et al., 2019).

and precision farming are indicated as best practices to protect the soil (European Commission, 2021a), potentially incentivised by voluntary eco-schemes and agri-environmental measures.

While so far CAP subsidies have been linked to the compliance to certain practices (GAEC, greening) the reform process binds them to the attainment of certain evidence-based targets, defined in the CAP National Strategic Plans (NPS) presented by each MS. Given the emphasis on soil management, it is plausible that such targets will encompass, among others, the effect of sustainable farming practices on soil properties. For this reason, our analysis uses one or more soil indices (described in chapter 3.3) as a proxy for specific soil health policy targets: organic matter decline, salinisation, and desertification. It is important to clarify that such proxies represent, for a variety of reasons, a simplification of the underlying policy target.

As mentioned above, the introduction of the greening payment represents one of the main novelties of the 2015-2022 CAP programming period. Greening ensures that the allocation of CAP direct payments to the farmers is bound by their fulfilment of some ‘agricultural practices beneficial for the climate and environment’ (Reg. (EU) No 1307/2013).

The practices are as follows:

a) diversification of arable crops;

b) maintaining existing permanent grassland;

c) having an ecological focus area on the agricultural area.

In Italy, only the first and the third requirements had to be accomplished at farm level. The second practice stated that the national grassland area could not decrease by more than 5% compared to a reference area. The consequence, at the level of individual farms, is that an authorisation was required if a farmer intended to convert a permanent grassland parcel to another use. However, the Italian national register of permanent grassland was established only in 2018, that is out of our period of investigation. For this reason, only the two greening practices ‘diversification of arable crops’ and ‘ecological focus area’ are relevant to our analysis.

‘Diversification of arable crops’ mandates that farms having more than 10 (30) hectares of arable land should cultivate at least 2 (3) arable crops, allocating a minimum share of cropland to the less represented crop(s). The third rule establishes that in farms with above 15 hectares of arable land, at least 5% of such surface should be devoted to Ecological Focus Areas (for instance nitrogen-fixing crops, fallow land, landscape features, buffer strips, etc.). Organic farms and farms with a large share of forage crops and/or leguminous crops and/or fallow land and/or flooded crops (rice) were exempted from the compliance with greening rules, regardless of their dimension. For further details, see Regulation (EU) n. 1307/2013, articles 43-47.

It is quite evident that these practices potentially affect farmland uses, particularly in a region like Lombardy, traditionally characterized by the presence of large farms and a widespread maize monoculture. Given such features, some studies have attempted to estimate how greening introduction impacted on farmland allocation choices in Lombardy. Most of these analyses were ex-ante simulations, based on small-scale sample data (e.g. Cortignani et al., 2017). Conversely, Bertoni et al. (2018), Micheletti et al. (2020) and Bertoni et al. (2021) performed a detailed ex-post analysis by observing the actual behaviour of all farmland parcels in the Lombardy region for some years before and after greening implementation. Results of Bertoni et al. (2018) and Micheletti et al. (2020) claim for a change in transition probabilities of the main arable crops. Such changes consist in a decrease of transitions probabilities toward maize and in an increase of those toward nitrogen-fixing crops (soybean, alfalfa), herbages and other cereals. Bertoni et. al. (2021) estimated the net effect of greening on farmland allocation between 2014 and 2015. Their results show a reduction of 10% of maize area in eligible and not initially compliant farms, counterbalanced by increases of other crops (mainly soybean, alfalfa, wheat, barley, and fallow land). However, all these estimations focus on the effects of the policy on farmland allocation, which is only the visible outcome of the greening policy, and not on its environmental impact. In this sense, our study aims at testing whether and to what extent level and variations in soil indices may be somehow associated to observable outcomes (land use change) induced by the CAP greening.

3. Datasets and methods

3.1. Study area

The area examined in this study is a large portion (covering about 2,150 km^2) of the flatland of Lombardy region, in Northern Italy (see Figure 1). Lombardy is the most populated Italian region with more than 10 million inhabitants, corresponding to 16.4% of the Italian population, concentrated in only 7.9% of the national territory (ISTAT, 2021). Over 53% of the area, mainly in the north, is mountainous or hilly, while the southern part of the region is a plain, where 68.5% of the land is managed by the agricultural sector (ISTAT, 2021; DUSAF, 2015). Here, arable crops – in particular maize and forage crops - represent the main farmland utilization type (92.9%), while permanent grassland is limited to only 6.4% of the farmland (DUSAF, 2015). Water flood irrigation is widespread. Agriculture in Lombardy is characterized by high-intensity farming systems (Fumagalli et al., 2011). The average value of production per hectare of farmland is 3.9 times that of the EU-28, while the average economic dimension of the farms is 5.7 times (Eurostat Farm Structure Survey, 2013). Livestock products, especially milk and pigs, provide two thirds of the agricultural value, with the livestock density (measured as the number of Livestock Standard Units – LSU – per hectare), being 3.7 times that of the EU-28 (Eurostat Farm Structure Survey, 2013). Given the high livestock density, both nitrate leaching in water, and its dispersion in air, are considered serious problems (Perego et al., 2012; Acutis et al., 2014; Paracchini et al., 2015), while there is no particular evidence in the literature of critical issues related to the management of phosphorus.

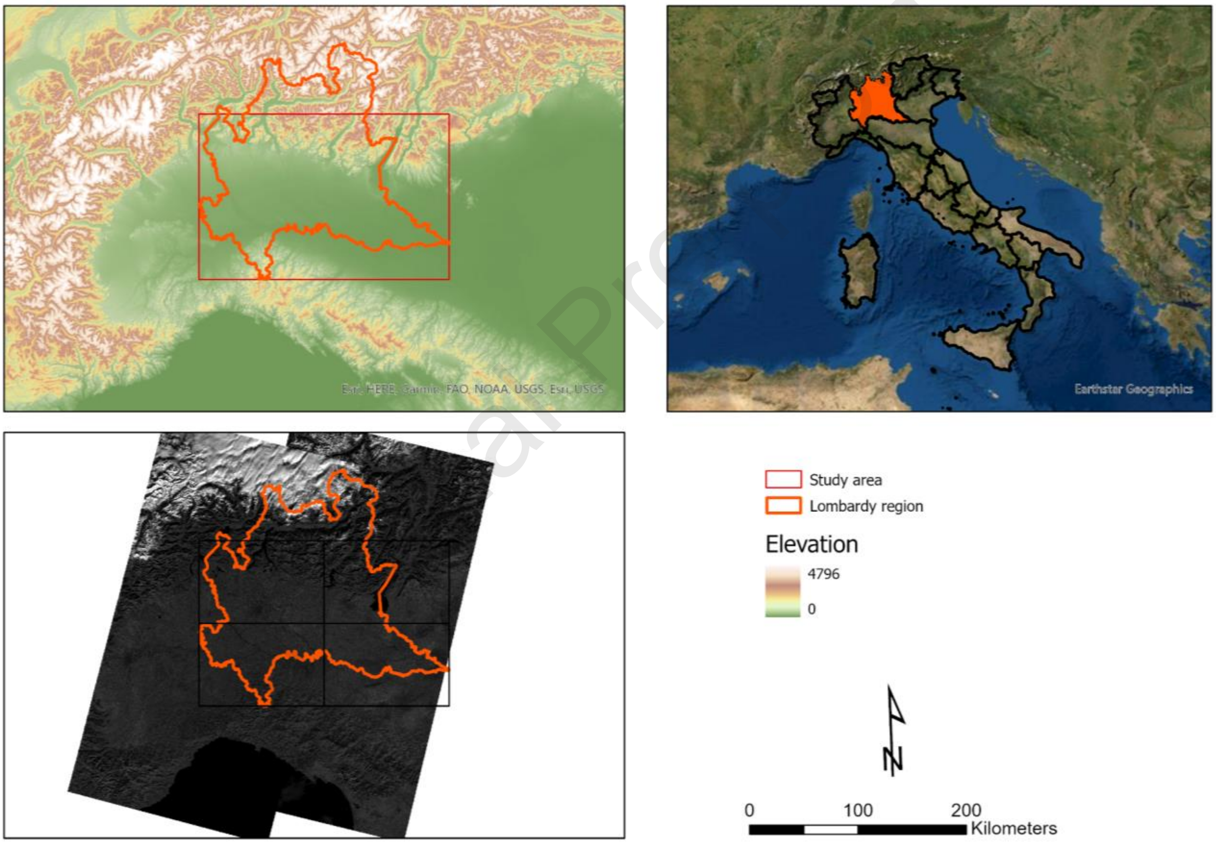


Figure 1: Top left: elevation of Lombardy and the study area; top right: location of Lombardy region within Italy. Bottom left: the study area and example of the BI index calculated using Landsat OLI data from 2014.

3.2. Remote sensing data

Remote sensing data were acquired for 2014 (pre-greening conditions) and 2017 (greening conditions). To map soil properties through spectral indices, soils are required to be bare to exclude the influence of vegetation (Wang and Qu, 2007). Thus, we selected satellite images from October/ early November, as this time of year marks the period of rotation between summer and winter crops.

For 2014, we acquired Landsat-8 OLI data at 30 m spatial resolution, at the L1T processing level. These data are freely available from USGS and were downloaded at http://earthexplorer.usgs.gov/; altogether we selected four cloud-free tiles (see Table 1), acquired between 23/10 and 01/11, covering the low-lying areas of Lombardy. Individual tiles were merged together using the mean value for overlapping portions and top-of-atmosphere percent reflectance between 0 and 1 was calculated using the coefficients found in the Landsat tile metadata. Soil indices were calculated on the image mosaic obtained from the four tiles.

For 2017, we downloaded six cloud-free 100 km by 100 km tiles (see Table 1) from the Sentinel-2A and –B satellites, acquired on 14/10 and 16/10, at the L1C processing level. The tiles were downloaded from the ESA Copernicus open access hub (https://scihub.copernicus.eu/) and were merged together using the same approach described for Landsat before calculating the indices on the image mosaic. For consistency with the other bands, we resampled bands 8A, 11 and 12, originally at 20 m spatial resolution, to 10 m resolution. These bands are necessary for calculation of the NMDI. We did not resample the other 20 m or 60 m bands of Sentinel-2, as they were not needed for calculation of other indices. Sentinel-2 digital numbers are provided as scaled reflectance between 0 and 10,000. We transformed this value to percent reflectance between 0 and 1.

Table 1: satellite images used to calculate indices for 2014 and 2017. Tile refers to the tiling system of Landsat 8 (path_row) and Sentinel-2 (UTM Military Grid Reference System).

Satellite	**Tile**	**Date**	**Resolution (m)**
Landsat 8 OLI	193_28	01/11/2014	30
Landsat 8 OLI	193_29	01/11/2014	30
Landsat 8 OLI	194_28	23/10/2014	30
Landsat 8 OLI	194_29	23/10/2014	30
Sentinel-2	32TMQ	14/10/2017	10
Sentinel-2	32TMR	14/10/2017	10
Sentinel-2	32TNQ	14/10/2017	10
Sentinel-2	32TNR	14/10/2017	10
Sentinel-2	32TPQ	16/10/2017	10
Sentinel-2	32TPR	16/10/2017	10

3.3 Administrative data

In our analysis, remote sensing data were combined with georeferenced administrative data coming from the SIARL dataset of Lombardy Region. SIARL is an administrative dataset by which Lombardy Region administration manages farmer requests for CAP payments. For each of about 50,000 farms located in the Region, it contains information about farm structures, crops, livestock, family and hired labour, and CAP payments received. Particularly, for our purposes we gathered data about the use of each of the two millions of farmland parcels of the Lombardy region from 2011 to 2017. These data came from yearly declarations of farmers applying for CAP payments. Parcels are typically rectangular in shape, with an average size of 1.2 ha, so they cover at least 10 Landsat 8 pixels and more than 100 Sentinel-2 pixels, which we deem adequate for our analysis.

Given the occasional contemporary presence of more crops on the same parcel in the same year, not being able to georeference which part of the parcel is occupied by each crop, we adopted a methodology to assign each georeferenced parcel to only one use per year. In any case, this eventuality was quite rare, with the notable exception of the intra-annual succession between ryegrass and maize for silage. In this particular case we created a separate category of farmland use (ryegrass + maize for silage). For further details about the adopted procedure see Bertoni et al. (2018). Consequently, the resulting number of potential farmland uses was 23. Subsequently, we created a balanced panel 2011-2017, eliminating those parcels not declared in all the observed years. In any case, the resulting balanced panel of farmland parcels represents 95% of the total agricultural area. Finally, we reduced the georeferenced balanced panel of parcels, considering only those parcels which overlap with remote sensing data. The result was a balanced panel 2011-2017 of 153,954 farmland parcels covering altogether 214,968 hectares. Since some farmland uses result in a permanent land cover, leading to distortions in remote sensing indices, and/or are not affected by greening rules at all, we decided to eliminate parcels with those land use types. More specifically, we excluded from the dataset parcels with the following uses for at least one year between 2014 and 2017: permanent crops, permanent grassland, flowers, woods and rice. After cleaning these parcels, the balanced panel 2011-2017 shrinks to 129,166 parcels for 191,645 hectares. The final number of farmland uses included in the model was 17.

Based on Bertoni et al. (2021), each parcel belonging to the dataset was assigned to one of three groups based on criteria of both eligibility and compliance with the greening policy of the farm to which they belonged. These conditions were verified by referring to the year 2014, which was the last year before greening introduction. Specifically, we separated parcels into three groups: 1) Group 1 (not eligible). This group includes parcels belonging to small farms and those that are exempted from greening payments because of their characteristics; 2) Group 2 (eligible and compliant). This group includes parcels belonging to those farms that are eligible for greening practices, but were already compliant with these rules in 2014; 3) Group 3 (eligible, not compliant). This group includes parcels pertaining to farms that were eligible for greening payments, but that before the implementation of greening policies did not satisfy greening requirements. Following this framework, Bertoni et al. (2021) demonstrated that after the introduction of greening, farmland parcels belonging to Group 3 changed their probability of transition from one use to another, compared to the pre-greening period, with a penalization of maize in favour of soybean, alfalfa, herbages and other cereals. Conversely, parcels of both Group 1 and Group 2 maintained the same transition dynamics.

Classification of farmland parcels into three greening groups resulted in the creation of three sub-samples of respectively 32,155 hectares (Group 1), 74,810 hectares (Group 2), and 84,680 hectares (Group 3). In annex 1 we show the details about the evolution of the farmland use, before and after greening introduction, in each of the three above mentioned groups.

3.4 Calculation of remote sensing indices for each parcel.

Remote sensing indices described here represent a proxy for specific soil health policy targets such as organic matter decline, salinisation, and desertification (see chapter 2 for more reference). For a variety of reasons, such proxies represent a simplification of the underlying policy target. Four types of indices were calculated based on the merged Landsat 8 and Sentinel-2 data, including the normalized multiband drought index (NMDI, Wang and Qu, 2007), which is related to the soil moisture content, the salinity and brightness indices (SI and BI, Dehni and Lounis, 2012), which are related to the soil salinity, and the Normalized Difference Vegetation Index (NDVI). The NDVI, although originally developed and most often used to monitor the health or phenology of vegetation (Primi et al., 2016), has been linked to the organic matter content in soils (Guo et al., 2017; Gholizadeh et al., 2018).

The NMDI, developed by Wang and Qu (2007), has an inverse relationship with the moisture content of the soil: higher values (> 0.7) represent a moisture content lower than 0.1, values around 0.6 intermediate wetness conditions and < 0.6 wet soil conditions (Wang and Qu, 2007). It is calculated following equation (1)

$$NMDI = \frac{NIR-(SWIR1-SWIR2)}{NIR+(SWIR1+SWIR2)} \tag{1}$$

Where NIR stands for near infrared and SWIR stands for shortwave infrared. The equation proposed by Wang and Qu (2007) is based on the MODIS sensor, where NIR is a band centered at 860 nm, SWIR1 is centered at 1640 nm and SWIR2 at 2130 nm. We selected the closest bands of Landsat 8 (USGS, 2018) and Sentinel-2 (ESA, 2015). For Landsat, these are band 5 (865 nm), band 6 (1608.5 nm) and band 7 (2200 nm). For Sentinel-2, we used band 8A (20 m resolution, resampled to 10 m) which is narrower than band 8 and closer to the original center wavelength, at 865 nm. Bands 11 (1610 nm) and 12 (2190 mn) were used as SWIR1 and SWIR2, respectively.

The three other indices only require a red and NIR band, combined in different ways. For the three indices, NIR (near infrared) corresponds to Landsat 8 OLI band 5 and Sentinel-2 MSI band 8, while the red band corresponds to band 4 in both Landsat 8 OLI and Sentinel-2 MSI mosaics.

The SI and BI (sometimes called SI-1 and SI-3) are both sensitive to the salinity content of the soil, and are directly proportional to it, i.e. higher values represent an increase in salinity compared to lower ones (Bouaziz et al., 2011); these indices range generally between 0-0.30 (Nguyen et al., 2020). The indices were calculated following equations (2) and (3), see Dehni and Lounis (2013):

$$SI = \sqrt{RED \times NIR} \quad (2)$$

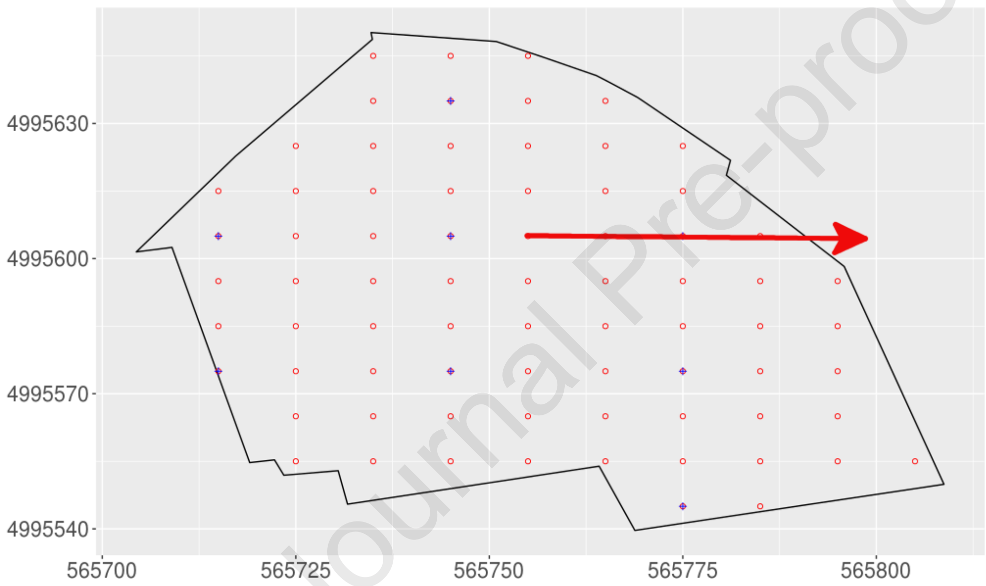
$$BI = \sqrt{RED^2 + NIR^2} \quad (3)$$

The NDVI, while generally employed to retrieve vegetation properties, has also been linked to the organic carbon content of soils (Zhang et al., 2019). This index generally ranges between -1 and +1, with values < 0.2 identifying bare soils, and is directly related to the organic content (Guo et al., 2017; Gholizadeh et al., 2018). We followed Rouse et al. (1974) to calculate the NDVI according to equation (4).

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (4)$$

We further extracted the average value of each index for all agricultural parcels. In view of the large number of parcels in the administrative database (153,954) and high resolution of raster data, using conventional GIS software proved too time consuming to perform these calculations; thus, we developed a computationally efficient version of the ray casting algorithm (Hormann and Agathos, 2001) in the C programming language. The ray casting algorithm is used to detect whether a point lies inside a polygon and is based on the concept of ray crossing (see Fig. 2). A point is deemed inside a polygon if a horizontal ray starting from the point in a fixed direction crosses one side of the polygon an odd number of times. Otherwise, the point lies outside the polygon. The center of each raster cell was used as a starting point and each parcel was attributed the average of all raster cells lying inside it for each remote sensing index.

Figure 2: The principles of the ray casting algorithm. Blue dots represent the center of Landsat 8 OLI cells (30 m resolution), while red dots represent the centers of Sentinel-2 cells (10 m resolution). The horizontal ray (red arrow) from one dot inside the polygon in a fixed direction crosses a polygon side an odd number of times.



3.5 Statistical tests on remote sensing indices and their variation

Based on the ray casting algorithm, we calculated for each parcel the value of each index and the difference between its value in 2017 and 2014, that is the last year before greening implementation (Δ_{index}=index$_{2017}$-index$_{2014}$). Recalling that both BI and SI are directly related to the salinity degree of the soil, for these two indexes a positive variation of Δ_{index} means that salinity increased between 2014 and 2017. Conversely, being NMDI inversely correlated with the content of water in the soil, an increase of Δ_{index} results in a reduction of water in the soil, and vice-versa. As for the NDVI, as the index is directly linked to organic matter content when soil is bare, an increase of Δ_{index} means an increase in soil organic matter content over the period 2014-2017.

To evaluate the degree of association of different, competing factors in affecting values or variations in the indices, we performed five tests using stepwise selection for a linear regression based on the

Akaike information criterion (AIC) and weighting the observations by parcel size (see Venables and Ripley, 2002). Stepwise selection (or sequential replacement) is a combination of forward and backward selections. It starts with no predictors, then sequentially adds the predictor that most increases the AIC, like in a pure forward selection. After adding each new variable, it removes any variables that no longer provide an improvement in the model fit (like in a pure backward selection). When the procedure reaches an equilibrium, it stops. Depending on the test, in turn we used each index, or Δ_{index}, as the dependent variable. It is worth pointing out that the tests performed aim to detect a statistically significant effect of the independent variables on the soil indices (in absolute or in variations), keeping in mind that those dependent variables may be affected by other, unobserved factors. Particularly, in the five tests we evaluated:

1) Variation of soil indices with respect to farmland use. For each parcel, we determined the farmland use type in the three years post-greening (2015, 2016 and 2017) and summed the number of years for each type of land use as the variable coefficient. Δ_{index} was selected as the dependent variable (see table 2 in the results section). As an example, if the parcel use in 2015 and 2016 was wheat, while in 2017 it was soybean, the model evaluates $\Delta_{index} = 2\alpha_{wheat} + 1\alpha_{soybean}$. Maize farmland use was used as a reference. This kind of additive model is a second-best strategy to control, in part, for the kind of rotations over the period examined. In fact, the possible rotations would be too numerous to include them as dependent variables in the regression.
2) Variation of soil indices with respect to crop rotation. As in step (1), we identified the farmland use type for every year after 2014 and summed the number of changes in farmland use type from one year to another, which was used as the independent variable in the regression model. Δ_{index} was used as the dependent variable (see table 3 in the results section). In the previous example, the model would evaluate $\Delta_{index} = 1\alpha_{rot}$, while in case of no changes the number of rotations would be 0.
3) Pre-greening soil conditions with respect to eligibility and compliance with greening policies before they came into effect. The aim of this test was to understand whether soil indices were significantly different in the farms/parcels targeted by greening policies (group 3) compared to those which already complied with the policies (group 2) and those which were not eligible (group 1, chosen as reference). Here, the absolute value of each index, calculated in 2014, was used as the dependent variable (see table 4 in the results section).
4) Variation of soil indices with respect to eligibility for and compliance with greening policies before they came into effect. We used the same approach described at point (3) to determine independent variables. In this case Δ_{index} was used as the dependent variable (see table 5 in

the results section). The aim of this test was to determine whether farms/parcels that had to change cultivation practices to comply with greening rules actually underwent an improvement in soil quality indices as seen by remote sensing indicators. Group 3 (eligible, not compliant) was the reference.

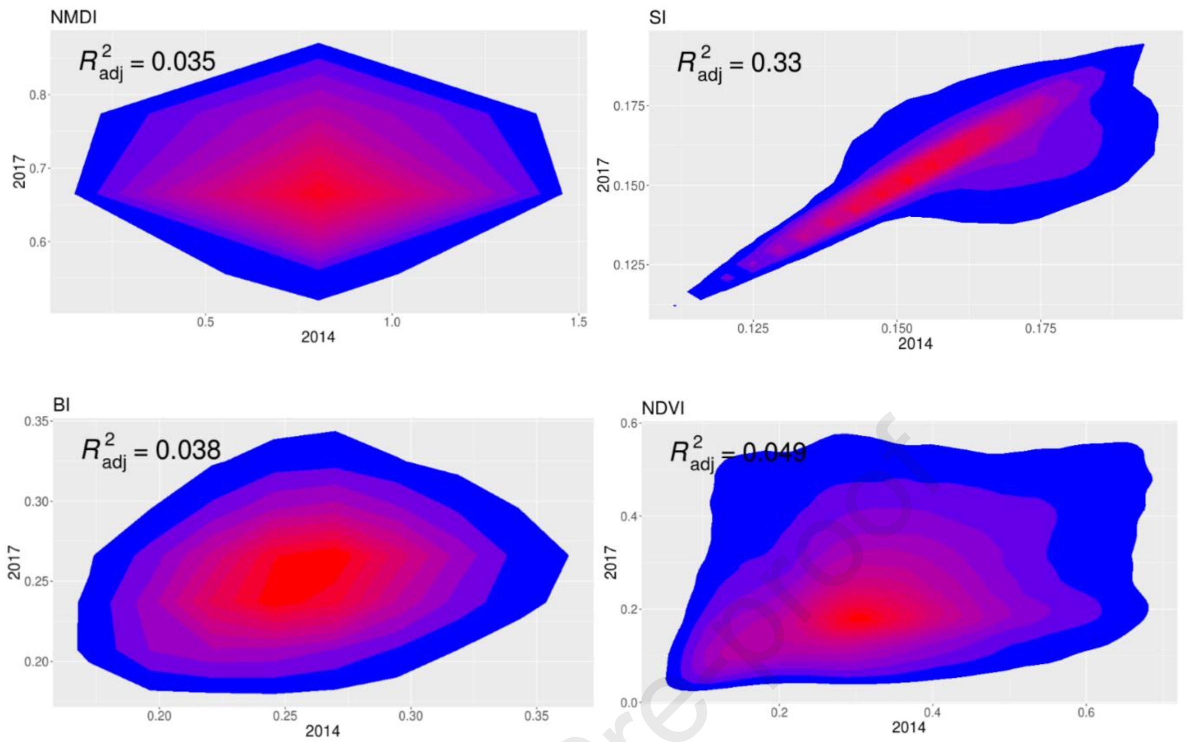
5) Variation of soil indices with respect to eligibility for and compliance with greening policies before they came into effect, and with respect to the farmland use and number of rotations. This last model is aimed at testing whether or not the estimated effects in model 4 are affected by the changes in farmland uses and number of rotations induced by the greening policy. If so, it would be possible to state that the differences between the greening groups are attributable to the variation in the crop mix. If not, these variations could largely depend on unobserved factors linked to the group to which each farmland parcel belongs.

4. Results

4.1 Variation in soil indices

We first investigated changes in remote sensing based soil indices between 2014 and 2017, to understand how they evolved over time and which factors could contribute to variations aside from actual changes in soil properties. Density plots in figure 3 show the relationship between the value of each index in 2014 and 2017 for all agricultural parcels considered in the analysis. Such preliminary analysis is aimed at exploring possible impacts of starting values of each index on its variation. Low values of adjusted r-squared point to an independence between starting and end values, implying lack of correlation between absolute index and its variation. Among the indices, the SI shows a very high persistence, confirmed by a relatively high adjusted r-squared. Lower relative values of this index in 2014 were followed by low values in 2017, while for higher values (> 0.15) in 2014 there is a higher spread and a tendency towards lower values in 2017. In comparison the BI, which should be linked to the same soil characteristics as the SI, i.e. salinity, has a much larger spread, and so do the NDVI and NMDI. The former shows slightly lower values in 2014 than in 2017 while the latter shows very little persistence, as the value of the index in 2014 and 2017 appear independent from each other.

Figure 3: density plots of soil index values in 2014 and 2017 for all parcels considered in the analysis. Red represents a greater density of observations than blue. The value of adjusted R-squared is also reported.



4.2 Regression Analysis

While the variations in soil indices appear shaped by various parameters, we performed subsequent statistical tests to provide insights into the factors directly related to the implementation of the greening policy. In the stepwise regression analysis, we found that the constant term is negative for all models where Δindex was used as the dependent variable, suggesting an average decrease in the value of all soil indices between 2014 and 2017 (see tables 2, 3, 5). Given the exploratory nature of this analysis, results are presented even if adjusted r-squared values are generally low in all five predictive models. For this reason, we chose not to focus on effect size. However, several independent variables included in the models attain high statistical significance. Here, we describe only those values that were significant in stepwise regression at the 99% confidence level.

In the first test, we evaluated the relationship between farmland use type and Δindex (Table 2). The final number of land use types included in the model was highest for the ΔNDVI (14, 12 with $p < 0.01$) and lowest for the ΔBI (8, of which 2 with $p < 0.01$). At the same time, 11 (7) and 9 (8) variables were included in the models for the ΔNMDI and ΔSI, and found significant for their variations at the 99% confidence level, respectively. Variations in BI had a negative correlation with rotation ryegrass + maize for silage and horticulture, but a positive correlation with alfalfa and other temporary grassland. ΔNDVI was negatively correlated with rotation ryegrass + maize for silage, triticale, legume herbages, temporary grassland and fallow land. It was positively correlated with wheat, soybean, grass and mixed herbages, horticulture, ryegrass and alfalfa. Variations in NMDI were positively correlated with wheat, soybean, horticulture and alfalfa. They were negatively correlated with triticale, pulses (highest coefficient), grass herbages and fallow land. Finally, variations in SI were positively correlated with rotation ryegrass + maize for silage, triticale, temporary grassland and fallow land. They were negatively correlated with wheat, soybean, horticulture, other arable crops and alfalfa. Alfalfa and horticulture were significant in predictive models for all indices with concordant effect (with the exception of ΔBI).

In the second test, we looked at the influence of rotation practices (see Table 3). The number of rotations at parcel level was positively correlated with ΔNDVI and ΔNMDI and negatively correlated with ΔSI, with ΔNDVI attaining a slightly higher coefficient. The coefficients suggest a decrease in salinity and an increase in organic matter with increased number of crop rotations, but also decreased soil water content.

In the third and fourth test, we investigated the relationship between soil indices and compliance with greening policies, both before and after the implementation of greening policies. The estimated coefficients indicate the effect of the membership of a farmland parcel to group 1 (not eligible) and to group 2 (eligible and compliant) vs group 3 (eligible and not initially compliant to greening policies). We first checked whether membership in predefined groups had a relationship with the value of soil indices in 2014, that is the last year before greening implementation (see Table 4). In this way we assessed possible different starting points among groups of parcels that will or will not be called upon to adapt to the greening policy. Both group 1 and group 2 had a positive relationship with the SI and a negative correlation both with the NDVI and the NMDI compared to group 3. These findings could point to relatively high salinity in both groups in 2014, but also higher water content, particularly for members of group 1. BI coefficient confirms that the salinity was higher for group 1, at least. For the NDVI, the initial content of organic matter in the soil appears lower in groups 1 and 2, compared to group 3. It is worth mentioning that such results are based on cross-sectional data referred to a single year.

The analysis of the variation in soil indices with respect to eligibility and compliance to greening policies (see Table 5) shows that membership in group 1 (not eligible) and group 2 (eligible and compliant) has a negative relationship with variations in the SI, and positive relationships with variation in the BI, NDVI and NMDI. This suggests group membership could play a role by leading to decreased water content and increased organic matter for groups not expected to change their farmland allocation after the implementation of greening policies, while the changes in salinity are less clear.

Finally, in Table 6, we performed a more complete estimation of membership in greening groups, controlling for both number of rotations and farmland uses. This last elaboration permits us to evaluate how much of the variability between greening groups depends on increased crop variability or on other not controlled factors. Estimated coefficients for greening groups do not change sign and level of significance (with the exception of ΔBI for group 1 which lost significance). However, the magnitude of the coefficients is reduced compared to the model in Table 5, particularly for those of group 2.

Table 2: predictive models of changes in soil remote sensing indices by farmland use type. Land use types are the independent variables, Δ_{index} is the dependent variable. Only variables significant at the 90% confidence level or above are included.

Farmland uses	Dependent variables			
	ΔBI	ΔNDVI	ΔNMDI	ΔSI
Maize for silage			0.0009	
Ryegrass + maize for silage	-0.0009***	-0.0057***		0.0006***
Wheat		0.0106***	0.0023***	-0.0009***
Barley	0.0007			
Triticale	0.0011*	-0.0170***	-0.0041***	0.0013***
Soybean		0.0132***	0.0035***	-0.0016***
Pulses		-0.0162**	-0.0110**	
Horticulture	-0.0008*	0.0133***	0.0038***	-0.0016***
Other arable crops		0.0046**	0.0020	-0.0007***
Ryegrass	0.0013*	0.0077***		
Grass herbages		0.0089***	-0.0098***	
Legume herbages		-0.0176***		
Mixed herbages	0.0015**	0.0110***		
Alfalfa	0.0008***	0.0129***	0.0032***	-0.0006***
Other temporary grassland	0.0005**	-0.0049***	-0.0008*	0.0004***
Fallow land		-0.0061***	-0.0040***	0.0004**
Intercept	-0.0090***	-0.0425***	-0.0130***	-0.0034***
Observations	129,166	129,166	129,166	129,166
R-squared	0.0003	0.0091	0.0016	0.0048
Adjusted R-squared	0.0003	0.0090	0.0015	0.0047
Residual Std. Error	0.0629	0.1921	0.1315	0.0222
F Statistic	5.4539***	85.1864***	18.5115***	69.2321***
	(df = 8; 129157)	(df = 14; 129151)	(df = 11; 129154)	(df = 9; 129156)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3: predictive models of changes in soil remote sensing indices by rotation practices. The number of rotations from one year to the other is considered as independent variable, Δ_{index} is the dependent variable.

N. rotations	ΔBI	ΔNDVI	ΔNMDI	ΔSI
	Dependent variables			
N. rotations		0.0101***	0.0025***	-0.0012***
Constant	-0.0086***	-0.0503***	-0.0151***	-0.0021***
Observations	129,166	129,166	129,166	129,166
R-squared	0.0000	0.0014	0.0002	0.0014
Adjusted R-squared	0.0000	0.0014	0.0002	0.0013
Residual Std. Error	0.0629	0.1929	0.1316	0.0222
F Statistic		176.9489***	22.5275***	175.3862***
		(df = 1; 129164)	(df = 1; 129164)	(df = 1; 129164)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: Predictive model of soil indices in 2014 with respect to eligibility and compliance with greening policies. The index value in 2014 is the dependent variable. Group membership represents the independent variable.

Eligibility and compliance to greening	BI	NDVI	NMDI	SI
	Dependent variables			
Not Eligible (gr.1) vs Eligible and not Compliant (gr.3)	0.0013***	-0.0105***	-0.0043***	0.0011***
Eligible and Compliant (gr.2) vs Eligible and not Compliant (gr.3)	-0.00005	-0.0221***	-0.0020**	0.0013***
Constant	0.2660***	0.3373***	0.5977***	0.1626***
Observations	129.166	129.166	129.166	129.166
R-squared	0.0001	0.0034	0.0002	0.0007
Adjusted R-squared	0.0001	0.0034	0.0002	0.0007
Residual Std. Error	0.0564	0.1640	0.1249	0.0225
F Statistic	6.6760***	223.3371***	11.2833***	45.2999***
	(df = 2; 129163)	(df = 2; 129163)	(df = 2; 129163)	(df = 2; 129163)

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 5: Predictive model of changes in soil remote sensing indices by eligibility and compliance with greening policies. Δindex is the dependent variable; group membership is the independent variable.

Eligibility and compliance to greening	ΔBI	ΔNDVI	ΔNMDI	ΔSI
	Dependent variables			
Not Eligible (gr.1) vs Eligible and not Compliant (gr.3)	0.0013***	0.0191***	0.0044***	-0.0012***
Eligible and Compliant (gr.2) vs Eligible and not Compliant (gr.3)	0.0017***	0.0242***	0.0051***	-0.0015***
Constant	-0.0095***	-0.0462***	-0.0137***	-0.0033***
Observations	129.166	129.166	129.166	129.166
R-squared	0.0001	0.0032	0.0003	0.0009
Adjusted R-squared	0.0001	0.0032	0.0003	0.0009
Residual Std. Error	0.0629	0.1927	0.1316	0.0222
F Statistic	9.1570***	209.7070***	20.7218***	59.0731***
	(df = 2; 129163)	(df = 2; 129163)	(df = 2; 129163)	(df = 2; 129163)

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 6: Predictive model of changes in soil remote sensing indices by eligibility and compliance with greening policies, number of rotations and farmland use. Δindex is the dependent variable; group membership, number of rotations and farmland uses are the independent variables.

Eligibility and compliance to greening	Dependent variables			
	ΔBI	ΔNDVI	ΔNMDI	ΔSI
Not Eligible (gr.1) vs Eligible and not Compliant (gr.3)	0.0008	0.0158***	0.0042***	-0.0010***
Eligible and Compliant (gr.2) vs Eligible and not Compliant (gr.3)	0.0013***	0.0169***	0.0035***	-0.0010***
Constant	-0.0097***	-0.0620***	-0.0200***	-0.0016***
Number of rotation controlled	yes	yes	yes	yes
Farmland uses controlled	yes	yes	yes	yes
Observations	129.166	129.166	129.166	129.166
R-squared	0.0004	0.0113	0.0019	0.0058
Adjusted R-squared	0.0003	0.0112	0.0018	0.0057
Residual Std. Error	0.0629	0.1919	0.1315	0.0221
F Statistic	5.4366*** (df = 10; 129155)	86.8755*** (df = 17; 129148)	20.5504*** (df = 12; 129153)	53.4458*** (df = 14; 129151)

*Note: *p<0.1; **p<0.05; ***p<0.01*

5 Discussion

5.1 Assessing the effects of policy changes through remote sensing: advantages and limitations

In our study, we combine remote sensing-based spectral indices, geospatial administrative data, a fast algorithm for data processing and a robust statistical approach to provide a first attempt to detect the effects of the CAP reform on soil quality indices in the Lombardy region. The novelty of the work lies in the combination of these sources to assess the effects of policy changes, and particularly in the use of spectral indices from satellite data. These indices have the advantage of simplicity over other methods, making use of models (e.g. Borrelli et al., 2016), and the availability of data from Landsat-8 and Sentinel-2 before and after the implementation of the CAP reform offers a unique possibility to look at the changes over wide areas while also providing a great level of detail. Our approach has further advantages: firstly, the combination of remote sensing techniques with administrative datasets ensures the almost total coverage of the territory and, at the same time, an extremely detailed level of analysis (farmland parcel level). Secondly, it represents a direct assessment of parameters related to the environmental sustainability of agriculture, and not derived from other parameters. Thirdly, it aims to offer an assessment of environmental performance parameters, linked to the quality of the soils, not investigated until now in previous studies. Finally, our methodology can be replicated over time and space, provided administrative data are available.

The use of satellite sources for the assessment of soil properties however also comes with a number of uncertainties. In fact, the estimation of remote sensing indices is related to seasonality and a requirement that limits the availability of data for this approach is the choice of the best period to obtain a homogeneous bare soil. Recommendations by Bartholomeus et al. (2008) suggest that vegetation cover should be lower than 20% to be able to estimate bare soil indices; the value might also be lower for the NDVI, which in this study was used for soil organic carbon and which is sensitive to the vegetation content of pixels observed by the satellite. As we cannot exclude that residual vegetation was present at the time of image acquisition, there is a chance that our approach to assess changes in soil organic carbon was affected. To exclude this possibility, and obtain a more robust indication of changes in this soil variable, the relationship between the NDVI and soil organic carbon should be tested in detail through dedicated field studies; alternatively, other indices or individual bands might be tested, including the reflectance in the visible band of the electromagnetic spectrum or the indices proposed by Bartholomeus et al. (2008), which should be the subject of further research. Additionally, the interaction between soil properties could be further evaluated, as organic matter and water content can be correlated (Kerr and Ochsner, 2020). Further issues still to be addressed include the calibration of indicators based on the degree of land cover and weather conditions, which deserve to be carefully investigated on a second step of improvement of such methodology, using e.g. field data as done by Yuzugullu et al. (2020).

Concerning weather conditions, another issue which limits data availability from satellites is the presence of cloud cover, which can be rather high in the plains of Lombardy especially during the winter season, as low clouds can persist for several days (Kästner and Kriebel, 2001). Cloud cover particularly influences data availability before the implementation of greening policies, as in 2014 only data from Landsat 8 were available; with the launch of Sentinel-2A in 2015 and its twin satellite Sentinel-2B in 2017, the number of remote sensing images for the greening period have greatly increased, and with them the chance to obtain high quality cloud-free data. Thus, future changes in remote sensing indices of soil quality might be investigated over a longer period and used to assess the effects of further changes in the agricultural policy if those are implemented. Further still, it might be possible to use a time series as opposed to single images to produce more robust estimates of variations in soil indices.

In this study, we did not perform an atmospheric correction on Landsat and Sentinel-2 data before calculating the spectral indices. The effects of this lack of atmospheric correction should nevertheless be limited, as normalized indices show a low sensitivity to atmospheric conditions. This was confirmed by Rumora et al. (2021), who found high statistically significant correlations between indices computed from Sentinel-2 and Landsat-8 imagery acquired on the same day and processed with three separate atmospheric schemes and without correction. Although non-normalized indices (e.g. SI e BI) might be more affected by the atmosphere, the strength of our analysis is that it focuses on the relative variation in indices compared to a reference, and the absolute value of indices is unimportant. To further increase the reliability of our approach, the integration with radar remote sensing, with its ability to provide estimates of soil moisture (Ezzahar et al., 2020), might prove useful. In summary, and given these limitations, the present analysis is therefore characterized above all from a methodological point of view and aspires to indicate a path for a new line of research, which could be refined in future evaluations.

5.2 Assessment of soil indexes dynamics after the introduction of CAP greening

CAP Greening mainly affects the allocation of agricultural land. As previously pointed out, in the area under investigation, greening has resulted in a reallocation of maize and maize for silage areas towards other crops such as soy, alfalfa, other cereals and set-aside (Bertoni et al., 2021). This occurred especially in group 3 (eligible and not compliant farms). For this reason, our first two elaborations have been related to the variations of the soil quality indices in relation to the farmland use (Tab. 2) and to the number of rotations in the reference period (Tab. 3). The coefficients calculated for the individual crops express a variation with respect to maize, which is the crop mainly affected by the introduction of the new policy. In general, the coefficients are very low, as is the R-square, a confirmation that the conditions of the soils tend to persist over time (Noormets et al., 2006) and are not particularly affected by crop changes, especially in a relatively short time interval such as that analysed. In any case, the coefficients of different crops were significant. Reminding that a positive coefficient indicates for BI and SI an increase in salinity, for NDVI an increase in the content of organic matter, for NMDI a reduction in the water content in soils, with reference to the main crops that have replaced maize, it is observed that soy has a positive effect on the organic substance content (NDVI) and contributes to the reduction of salinity (SI), on the contrary the NMDI coefficient goes in the direction of a reduction of the content of water in the soil. The trend of wheat is similar. Alfalfa would cause an increase in organic matter and a reduction in water content, while the results on salinity are contrasting between the two indices. The number of rotations over the period analysed is positively correlated with the organic matter content and negatively with the salinity and water content.

We then tested the changes in the indices based on the group to which the farmland parcels belong, in terms of eligibility and compliance with greening. The regression in Tab. 5 shows the coefficients of the two groups that have not been affected by greening (group 1 - not eligible farms and group 2 - eligible and compliant farms) vs the reference group (group - 3 eligible and not compliant farms), that has been affected by the introduction of the new policy. Finally, in Tab. 6 we repeated this previous regression by groups, controlling at the same time for farmland use and the number of rotations. The sign of the coefficients does not change, pointing to a time persistence of the dynamics of such indicators. Evidently, these are mainly due to specific factors of the groups analysed for which we do not have available control variables at the level of detail necessary for our analysis. This is also confirmed by the regression in Table 4, which highlights different starting conditions between the various groups. Therefore, the regression in Tab 6 suffers from omitted variable bias due to the exclusion of some control variables. These are cultivation and irrigation practices (e.g. fertilisation, fungicides, seed rates and timings, by-cropping, cover cropping, weed management, harvest dates) for which there is no georeferenced information in the territory considered, and soil type that is mapped at a scale not usable for our detail of analysis. Nevertheless, part of the effect of cultivation practices on soil indexes is controlled by including the type of crop. What is missing is the heterogeneity in agricultural practices for the same crop within the study area. The availability of this kind of information would have certainly improved the quality of our analysis, like the possibility of conducting the analysis over a longer period.

In any case, it is interesting to note that by controlling for farmland uses and number of crops, the coefficient magnitude of the greening groups decreases. This applies, in particular, to the soil organic matter content (NDVI) and the salinity index (SI), the latter in relation to group 2. Membership in group 3 (farms eligible and non compliant) would worsen both of these indicators, compared to groups 1 and 2. However, such an effect is less pronounced when controls are added. Such improvement in the two indexes is attributable to the change in crop mix and to the increased rotations in farms of group 3, as a consequence of the adaptation to greening prescriptions.

In general, the results should be considered as an outcome of a preliminary and innovative research study. The aim of such a tool of analysis is to combine available big data (from administrative and remote sensing sources) in the attempt to lay the foundations for an ex-post assessment of CAP greening on soil properties. Such analytical effort is in line with the green architecture on incoming CAP, focussed on assessing environmental outcomes of its policy tools. Keeping this in mind, its efficacy and accuracy can and should be improved in different ways. The first one has been already mentioned and relates to additional data and information to consolidate the results. The second one pertains to improvements on the calibration of the soil quality indicators. In particular, the indicator of organic matter content (NDVI) presents the most critical issues. Our choice of excluding perennial crops, rice and permanent grassland, as well as that of calculating the indicators in the period with the greatest probability of having bare soil, goes in the direction of a refinement of the analysis. However, we cannot exclude confounding effects due to the presence of crop residues in the field (e.g. maize stalks). Therefore, such aspects should be improved in order to make such a combined methodology more reliable for ex-post assessment of environmental effects of agricultural policy.

Conclusions

The present analysis is mainly methodological, aimed at developing a potential tool for ex-post assessment of CAP policy instruments, by combining administrative and remote sensing data with a high level of detail. As such, it could lay the foundations for the possible use of digital technologies in order to analyse environmental benefits, that are objectives desired by agro-environmental policies. In particular, we assessed the dynamics of agricultural land soil quality indices in the Lombardy region of Italy, over the period subsequent to the introduction of CAP greening in 2015. Soil quality indices considered in this study were obtained from remote sensing imagery, and combined with an administrative database of all agricultural parcels of the region; based on these two datasets, we performed a regression analysis to assess the impact of different factors, including farmland use type, eligibility and compliance with respect to greening policies, on the dynamic variability in soil properties in the region. Our preliminary results indicate a high persistence of the soil quality indicators before and after the introduction of CAP greening, but also significant correlations between variations in the indices and membership in specific groups pertaining to farmland use, eligibility and compliance with greening policies. In particular, greening group membership appears to lead to decreased water content and increased organic matter for groups not expected to change their farmland allocation after the implementation of greening policies.

Nevertheless, our preliminary results come with a number of limitations that should be acknowledged to identify further possible improvements. First of all, the remote sensing indices adopted represent only approximations for specific soil health policy targets such as organic matter decline, salinisation, and desertification. Among them, some uncertainties are related to the choice of the NDVI, which might particularly be impacted by residuals of vegetation in the parcel. Furthermore, the attempt to isolate the effect of greening policy on change of soil property may suffer from the omission of potentially relevant variables such as agronomic practices (for instance, fertilisation, tillage, irrigation) and soil pedological properties. Such variables have been omitted because they were not available in the study area at the parcel scale of the analysis. However, part of such omitted variability may have been controlled for by the inclusion of crop choices. The omitted variable bias is confirmed by the low degree of variability explained by our models. Further improvements are therefore necessary before our methodology can be fully used practically. These might stem from an increased availability of both remote sensing imagery (ideally at higher resolution and increased frequency) and data on agronomic practices and soil composition. Concerning remote sensing imagery, the two satellites of the Sentinel-2 family already partly satisfy the requirement for increased data, and the advancements foreseen by the Landsat programme also go in this direction. Another crucial aspect is to better control for the types of rotations, while the current model considers only the number of rotations. Such improvements would allow performing ex-post assessments of policy effects on soil property changes over longer time scales.

Acknowledgments

This study has been supported by Fondazione Cariplo, within the research project "Evaluation of CAP 2015-2020 and taking action – CAPTION" (Project Id: 2017-2513)

Word count: 11197 words

Reference List

Acutis, M., Alfieri, L., Giussani, A., Provolo, G., Di Guardo, A., Colombini, S., Bertoncini, G., Castelnuovo, M., Sali, G., Moschini, M., Sanna, M., Perego, A., Carozzi, M., Chiodini, E., Fumagalli, M. 2014. ValorE: An integrated and GIS-based decision support system for livestock manure management in the Lombardy region (northern Italy). *Land use policy*, 41, 149-162. https://doi.org/10.1016/j.landusepol.2014.05.007

Ansell, D., Freudenberger, D., Munro, N., Gibbons, P., 2016. The cost-effectiveness of agri-environment schemes for biodiversity conservation: A quantitative review. *Agriculture, Ecosystems & Environment*, 225, 184-191. https://doi.org/10.1016/j.agee.2016.04.008

Arata, L., Sckokai, P., 2016. The impact of agri-environmental schemes on farm performance in five EU member states: a DID-matching approach. *Land Economics*, 92(1), 167-186. https://doi.org/10.3368/le.92.1.167

Bachofer, F., Quénéhervé, G., Hochschild, V., Maerker, M., 2015. Multisensoral Topsoil Mapping in the Semiarid Lake Manyara Region, Northern Tanzania. *Remote Sensing* 7, 9563–9586. https://doi.org/10.3390/rs70809563

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., Spaargaren, O.S.P., 2008. Spectral reflectance based indices for soil organic carbon quantification. *Geoderma*, 145(1-2), 28-36. https://doi.org/10.1016/j.geoderma.2008.01.010

Bennett, H., Osterburg, B., Nitsch, H., Kristensen, L., Primdahl, J., Verschuur, G., 2006. Strengths and Weaknesses of Crosscompliance in the CAP. *EuroChoices*, 5(2), 50-57. https://doi.org/10.1111/j.1746-692X.2006.00034.x

Bertoni, D., Aletti, G., Ferrandi, G., Micheletti, A., Cavicchioli, D., Pretolani, R., 2018. Farmland use transitions after the cap greening: a preliminary analysis using markov chains approach. *Land Use Policy*, 79, 789-800. https://doi.org/10.1016/j.landusepol.2018.09.012

Bertoni, D., Curzi, D., Aletti, G., Olper, A., 2020. Estimating the effects of agri-environmental measures using difference-in-difference coarsened exact matching. *Food Policy*, 90, 101790. https://doi.org/10.1016/j.foodpol.2019.101790

Bertoni, D., Aletti, G., Cavicchioli, D., Micheletti, A., Pretolani, R., 2021. Estimating the CAP greening effect by machine learning techniques: A big data ex post analysis. *Environmental Science & Policy*, *119*, 44-53. https://doi.org/10.1016/j.envsci.2021.01.008

Borrelli, P., Panagos, P., 2020. An indicator to reflect the mitigating effect of Common Agricultural Policy on soil erosion. *Land Use Policy*, 92, 104467. https://doi.org/10.1016/j.landusepol.2020.104467

Borrelli, P., Paustian, K., Panagos, P., Jones, A., Schütt, B., Lugato, E., 2016. Effect of Good Agricultural and Environmental Conditions on erosion and soil organic carbon balance: a national case study. *Land Use Policy*, 50, 408–421. https://doi.org/10.1016/j.landusepol.2015.09.033

Bouaziz, M., Matschullat, J., Gloaguen, R., 2011. Improved remote sensing detection of soil salinity from a semi-arid climate in Northeast Brazil. Comptes Rendus *Geoscience*, 343, 795–803. https://doi.org/10.1016/j.crte.2011.09.003

Burkhard, B., Kroll, F., Müller, F., Windhorst, W., 2009. Landscapes' capacities to provide ecosystem services-a concept for land-cover based assessments. *Landscape online*, 15, 1-22. https://doi.org/10.3097/LO.200915

Carey, M., 2019. The Common Agricultural Policy's New Delivery Model post-2020: National administration perspective. *EuroChoices*, 18(1): 11–17. https://doi.org/10.1111/1746-692X.12218

Castaldi, F., Chabrillat, S., Jones, A., Vreys, K., Bomans, B., van Wesemael, B., 2018. Soil Organic Carbon Estimation in Croplands by Hyperspectral Remote APEX Data Using the LUCAS Topsoil Database. *Remote Sensing*, 10, 153. https://doi.org/10.3390/rs10020153

Castaldi, F., Hueni, A., Chabrillat, S., Ward, K., Buttafuoco, G., Bomans, B., Vreys, K., Brell, M., van Wesemael, B., 2019. Evaluating the capability of the Sentinel 2 data for soil organic carbon prediction in croplands. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, 267–282. https://doi.org/10.1016/j.isprsjprs.2018.11.026

Chabé-Ferret, S., Subervie, J., 2013. How much green for the buck? Estimating additional and windfall effects of French agro-environmental schemes by DID-matching. *Journal of Environmental Economics and Management*, 65(1), 12-27. https://doi.org/10.1016/j.jeem.2012.09.003

Coderoni, S., Esposti, R., 2018. CAP payments and agricultural GHG emissions in Italy. A farm-level assessment. *Science of the Total Environment*, 627, 427-437. https://doi.org/10.1016/j.scitotenv.2018.01.197

Cortignani, R., Severini, S., Dono, G., 2017. Complying with greening practices in the new CAP direct payments: An application on Italian specialized arable farms. *Land Use Policy*, 61, 265-275. https://doi.org/10.1016/j.landusepol.2016.11.026

Dehni, A., Lounis, M., 2012. Remote Sensing Techniques for Salt Affected Soil Mapping: Application to the Oran Region of Algeria. *Procedia Engineering*, SWEE'11 33, 188–198. https://doi.org/10.1016/j.proeng.2012.01.1193

Douaoui, A.E.K., Nicolas, H., Walter, C., 2006. Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. *Geoderma*, 134, 217–230. https://doi.org/10.1016/j.geoderma.2005.10.009

Dube, T., Mutanga, O., Sibanda, M., Seutloali, K., Shoko, C., 2017. Use of Landsat series data to analyse the spatial and temporal variations of land degradation in a dispersive soil environment: A case of King Sabata Dalindyebo local municipality in the Eastern Cape Province, South Africa. *Physics and Chemistry of the Earth,* Parts A/B/C 100, 112–120. https://doi.org/10.1016/j.pce.2017.01.023

DUSAF, 2015. Online at: https://www.regione.lombardia.it/wps/portal/istituzionale/HP/DettaglioServizio/servizi-e-informazioni/Enti-e-Operatori/Territorio/sistema-informativo-territoriale-sit/uso-suolo-dusaf/uso-suolo-dusaf. Last accessed 16/12/2021. In Italian.

Ehlers, M.H., Huber, R., Finger, R., 2021. Agricultural policy in the era of digitalisation. *Food Policy*, 100, 2019. https://doi.org/10.1016/j.foodpol.2020.102019

European Commission, 2020. *Sustainable land use (greening) - Sustainable use of farmland, how farmers benefit financially* https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/income-support/greening_en last access: 19th February 2020

European Commission, 2021a. *CAP Specific Objectives… explained. Brief No 5. Efficient Soil Management*. https://ec.europa.eu/info/sites/default/files/food-farming-fisheries/key_policies/documents/cap-specific-objectives-brief-5-soil_en.pdf last access: 21th December 2021

European Commission, 2021b. *Evaluation support study on the impact of the CAP on sustainable management of the soil. Final Report*. https://doi.org/10.2762/799605

ESA, 2015. *Sentinel-2 user handbook.*

Eurostat Farm Structure Survey, 2013. Agricultural holdings by economic size of the farm. Online at: https://ec.europa.eu/eurostat/databrowser/view/tag00123/default/table?lang=en. Last accessed 16/12/2021.

Ezzahar, J., Ouaadi, N., Zribi, M., Elfarkh, J., Aouade, G., Khabba, S., Er-Raki, S., Chehbouni, A., Jarlan, L., 2020. Evaluation of backscattering models and support vector machine for the retrieval of bare soil moisture from Sentinel-1 data. *Remote Sensing*, 12(1), 72. https://doi.org/10.3390/rs12010072

Folmer, C., Keyzer, M.A., Merbis, M.D., Stolwijk, H.J., Veenendaal, P.J., 2013. *The common agricultural policy beyond the MacSharry reform* (Vol. 230). Elsevier.

Fumagalli, M., Acutis, M., Mazzetto, F., Vidotto, F., Sali, G., Bechini, L. 2011. An analysis of agricultural sustainability of cropping systems in arable and dairy farms in an intensively cultivated plain. *European Journal of Agronomy*, 34(2), 71-82. https://doi.org/10.1016/j.eja.2010.11.001

Garzon, I. (2006). *Reforming the common agricultural policy: history of a paradigm change*. Springer.

Gholizadeh, A., Žižala, D., Saberioon, M., Borůvka, L., 2018. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. *Remote Sensing of Environment*, 218, 89–103. https://doi.org/10.1016/j.rse.2018.09.015

Glenk, K., Shrestha, S., Topp, C. F., Sánchez, B., Iglesias, A., Dibari, C., Merante, P. 2017. A farm level approach to explore farm gross margin effects of soil organic carbon management. *Agricultural systems*, 151, 33-46. https://doi.org/10.1016/j.agsy.2016.11.002

Guo, L., Hao, H., Liu, Y., Ma, H., An, J., Sun, Q., Yang, Z., 2017. The Assessment of Soil Quality on the Arable Land in Yellow River Delta Combined with Remote Sensing Technology. *World Journal of Engineering and Technology*, 5, 18. https://doi.org/10.4236/wjet.2017.55B003

Hormann, K., Agathos, A., 2001. The point in polygon problem for arbitrary polygons. *Computational geometry*, 20(3), 131-144. https://doi.org/10.1016/S0925-7721(01)00012-8

ISTAT, 2021. *Principali statistiche geografiche sui comuni*. Online at: https://www.istat.it/it/archivio/156224. Last accessed 16/12/2021. In Italian

JRC, 2019a. Cross-Compliance. Available at: https://marswiki.jrc.ec.europa.eu/wikicap/index.php/Cross_Compliance

JRC, 2019b. Good Agricultural and Environmental Conditions (GAEC). Available at: https://marswiki.jrc.ec.europa.eu/wikicap/index.php/Good_Agricultural_and_Environmental_Conditions_%28GAEC%29

JRC, 2020. LUCAS: Land Use and Coverage Area frame Survey. Available at: https://esdac.jrc.ec.europa.eu/projects/lucas

Kästner, M., Kriebel, K.T., 2001. Alpine cloud climatology using long-term NOAA-AVHRR satellite data. *Theoretical and applied climatology*, 68(3), 175-195. https://doi.org/10.1007/s007040170044

Kelly, E., Latruffe, L., Desjeux, Y., Ryan, M., Uthes, S., Diazabakana, A., Dillon, E., Finn, J., 2018. Sustainability indicators for improved assessment of the effects of agricultural policy across the EU: Is FADN the answer?. *Ecological indicators*, 89, 903-911. https://doi.org/10.1016/j.ecolind.2017.12.053

Kerr, D. D., Ochsner, T. E. 2020. Soil organic carbon more strongly related to soil moisture than soil temperature in temperate grasslands. Soil Science Society of America Journal, 84(2), 587-596. https://doi.org/10.1002/saj2.20018

Kleijn, D., Sutherland, W.J., 2003. How effective are European agri-environment schemes in conserving and promoting biodiversity?. *Journal of applied ecology*, 40(6), 947-969. https://doi.org/10.1111/j.1365-2664.2003.00868.x

Kremmydas, D., Athanasiadis, I.N., Rozakis, S. 2018. A review of agent based modeling for agricultural policy evaluation. *Agricultural systems*, 164, 95-106. https://doi.org/10.1016/j.agsy.2018.03.010

Latawiec, A.E., Strassburg, B.B., Silva, D., Alves-Pinto, H.N., Feltran-Barbieri, R., Castro, A., Iribarrem, A., Cordeiro Rangel, M., Kalif, K.A.B., Gardner, T., Beduschi, F., 2017. Improving land management in Brazil: A perspective from producers. *Agriculture, ecosystems & environment*, 240, 276-286. https://doi.org/10.1016/j.agee.2017.01.043

Leteinturier, B., Herman, J.L., De Longueville, F., Quintin, L., Oger, R., 2006. Adaptation of a crop sequence indicator based on a land parcel management system. *Agriculture, Ecosystems & Environment*, 112(4), 324-334. https://doi.org/10.1016/j.agee.2005.07.011

Mary, S., 2013. Assessing the impacts of pillar 1 and 2 subsidies on TFP in French crop farms. *Journal of Agricultural Economics*, 64(1), 133-144. https://doi.org/10.1111/j.1477-9552.2012.00365.x

Micheletti, A., Aletti, G., Ferrandi, G., Bertoni, D., Cavicchioli, D., Pretolani, R., 2020. A weighted χ^2 test to detect the presence of a major change point in non-stationary Markov chains. *Statistical Methods & Applications*, 899–912. https://doi.org/10.1007/s10260-020-00510-0

Mulder, V.L., de Bruin, S., Schaepman, M.E., Mayr, T.R., 2011. The use of remote sensing in soil and terrain mapping — A review. Geoderma 162, 1–19. https://doi.org/10.1016/j.geoderma.2010.12.018

Nguyen, K.A., Liou, Y.A., Tran, H.P., Hoang, P.P., Nguyen, T.H., 2020. Soil salinity assessment by using near-infrared channel and Vegetation Soil Salinity Index derived from Landsat 8 OLI data: a case study in the Tra Vinh Province, Mekong Delta, Vietnam. *Progress in Earth and Planetary Science*, 7, 1. https://doi.org/10.1186/s40645-019-0311-0

Noormets, A., Ewers, B., Sun, G., Mackay, S., Zheng, D., McNulty, S. G., and Chen, J. 2006. Water and carbon cycles in heterogeneous landscapes: an ecosystem perspective. in Ecology of Hierarchical Landscapes: From Theory to Application, 89-123. ISBN 1-60021-047-3

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde, O., 2018. LUCAS Soil, the largest expandable soil dataset for Europe: a review. *European Journal of Soil Science*, 69(1), 140-153. https://doi.org/10.1111/ejss.12499

Paracchini, M.L., Bulgheroni, C., Borreani, G., Tabacco, E., Banterle, A., Bertoni, D.,Rossi, G., Parolo, G., Origgi, R., De Paola, C. 2015. A diagnostic system to assess sustainability at a farm level: The SOSTARE model. *Agricultural Systems*, 133, 35–53. http://dx.doi.org/10.1016/j.agsy.2014.10.004

Paut, R., Sabatier, R., Tchamitchian, M., 2020. Modelling crop diversification and association effects in agricultural systems. *Agriculture, Ecosystems & Environment*, 288, 106711. https://doi.org/10.1016/j.agee.2019.106711

Paz-Kagan, T., Shachak, M., Zaady, E., Karnieli, A., 2014. Evaluation of ecosystem responses to land-use change using soil quality and primary productivity in a semi-arid area, Israel. *Agriculture, ecosystems & environment*, 193, 9-24. https://doi.org/10.1016/j.agee.2014.04.024

Perego, A., Basile, A., Bonfante, A., De Mascellis, R., Terribile, F., Brenna, S., Acutis, M., 2012. Nitrate leaching under maize cropping systems in Po Valley (Italy). *Agriculture, Ecosystems & Environment,* 147, 57-65. https://doi.org/10.1016/j.agee.2011.06.014

Posthumus, H., Morris, J. 2010. Implications of CAP reform for land management and runoff control in England and Wales. *Land Use Policy*, 27(1), 42-50. https://doi.org/10.1016/j.landusepol.2008.06.002

Primdahl, J., Vesterager, J.P., Finn, J.A., Vlahos, G., Kristensen, L., Vejre, H., 2010. Current use of impact models for agri-environment schemes and potential for improvements of policy design and assessment. *Journal of environmental management*, 91(6), 1245-1254. https://doi.org/10.1016/j.jenvman.2009.12.012

Primi, R., Filibeck, G., Amici, A., Bückle, C., Cancellieri, L., Di Filippo, A., Gentile, C., Guglielmino, A., Latini, R., Mancini, L.D., Mensing, S.A., Rossi, C.M., Rossini, F., Scoppola, A., Sulli, C., Venanzi, R., Ronchi, B., Piovesan, G., 2016. From Landsat to leafhoppers: A multidisciplinary approach for sustainable stocking assessment and ecological monitoring in mountain grasslands. *Agriculture, Ecosystems & Environment*, 234, 118–133. https://doi.org/10.1016/j.agee.2016.04.028

Pufahl, A., Weiss, C.R., 2009. Evaluating the effects of farm programmes: results from propensity score matching. *European Review of Agricultural Economics*, 36(1), 79-101. https://doi.org/10.1093/erae/jbp001

Ronchi, S., Salata, S., Arcidiacono, A., Piroli, E. and Montanarella, L., 2019. Policy instruments for soil protection among the EU member states: A comparative analysis, *Land Use Policy*, 82, pp. 763-780. https://doi.org/10.1016/j.landusepol.2019.01.017

Rossi, R. (2018). CAP horizontal regulation. Financing, management and monitoring of the common agricultural policy for 2021-2027. Policy brief available at http://www.europarl.europa.eu/RegData/etudes/BRIE/2018/628302/EPRS_BRI%282018%29628302_EN.pdf last access: 19th February 2020

Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W. 1974. Monitoring vegetation systems in the Great Plains with ERTS. *NASA special publication, 351*(1974), 309.

Rousi, M., Sitokonstantinou, V., Meditskos, G., Papoutsis, I., Gialampoukidis, I., Koukos, A., Karathanassi, V., Drivas, T., Vrochidis, S., Kontoes, C., Kompatsiaris, I., 2020. Semantically enriched crop type classification and Linked Earth Observation Data to support the Common Agricultural Policy monitoring. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 529-552, 2021, https://doi.org/10.1109/JSTARS.2020.3038152

Rumora, L., Miler, M., Medak, D. (2021). Contemporary comparative assessment of atmospheric correction influence on radiometric indices between Sentinel-2A and Landsat 8 imagery. *Geocarto International*, 36(1), 13-27. https://doi.org/10.1080/10106049.2019.1590465

Sheffield, K., Morse-McNabb, E., 2015. Using satellite imagery to asses trends in soil and crop productivity across landscapes. *IOP Conference Series: Earth and Environmental Science*, 25, 012013. https://doi.org/10.1088/1755-1315/25/1/012013

Shoshany, M., Goldshleger, N., Chudnovsky, A., 2013. Monitoring of agricultural soil degradation by remote-sensing methods: a review. *International Journal of Remote Sensing*, 34, 6152–6181. https://doi.org/10.1080/01431161.2013.793872

Sitokonstantinou, V., Papoutsis, I., Kontoes, C., Lafarga Arnal, A., Armesto Andrés, A.P., Garraza Zurbano, J.A., 2018. Scalable parcel-based crop identification scheme using Sentinel-2 data time-series for the monitoring of the common agricultural policy. *Remote Sensing*, 10(6), 911. https://doi.org/10.3390/rs10060911

Tóth, G., Jones, A., Montanarella, L. (Eds.)., 2013. *LUCAS Topsoil Survey: methodology, data and results*. Luxembourg: Publications Office of the European Union. https://doi.org/10.2788/97922

USGS, 2018. *Landsat 8 Data Users Handbook.*

Venables, W.N., Ripley, B.D., 2002. *Modern Applied Statistics with S*, 4th ed, Statistics and Computing, Statistics,Computing Venables,W.N.:Statistics w.S-PLUS. Springer-Verlag, New York. https://doi.org/10.1007/978-0-387-21706-2

Viaggi, D., Gomez y Paloma, S., Mishra, A., Raggi, M., 2013. The role of the EU Common Agricultural Policy: assessing multiple effects in alternative policy scenarios *Land Use Policy*, 31, pp. 99-101. https://doi.org/10.1016/j.landusepol.2012.04.019

Walker, K.J., Critchley, C.N.R., Sherwood, A.J., Large, R., Nuttall, P., Hulmes, S., R. Rosed, Mountford, J.O., 2007. The conservation of arable plants on cereal field margins: an assessment of new agri-environment scheme options in England, UK. *Biological Conservation*, 136(2), 260-270. https://doi.org/10.1016/j.biocon.2006.11.026

Wang, L., Qu, J.J., 2007. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophysical Research Letters*, 34. https://doi.org/10.1029/2007GL031021

Yuzugullu, O., Lorenz, F., Fröhlich, P., and Liebisch, F. 2020. Understanding Fields by Remote Sensing: Soil Zoning and Property Mapping. Remote Sensing, 12(7), 1116. https://doi.org/10.3390/rs12071116

Zahm, F., Viaux, P., Vilain, L., Girardin, P., Mouchet, C., 2008. Assessing farm sustainability with the IDEA method–from the concept of agriculture sustainability to case studies on farms. *Sustainable development*, 16(4), 271-281. https://doi.org/10.1002/sd.380

Zhang, Y.; Guo, L.; Chen, Y.; Shi, T.; Luo, M.; Ju, Q.; Zhang, H.; Wang, S. 2019, Prediction of Soil Organic Carbon based on Landsat 8 Monthly NDVI Data for the Jianghan Plain in Hubei Province, China. Remote Sensing, 11, 1683. https://doi.org/10.3390/rs11141683

Annex 1 – Farmland uses 2014-2017 in the greening groups

Group 1 – Not eligible farms

Farmland use	2014 (before greening) Area (hectares) (a)	2014 (before greening) Area (%) (b)	Average 2015-2017 (after greening) Area (hectares) (c)	Average 2015-2017 (after greening) Area (%) (d)	Difference Area (%) (e) = (d) - (b)
Maize	8.146	25,3%	7.215	22,4%	-2,9%
Maize for silage	1.056	3,3%	1.265	3,9%	0,7%
Rotation ryegrass + maize for silage	639	2,0%	705	2,2%	0,2%
Wheat	3.580	11,1%	4.331	13,5%	2,3%
Barley	664	2,1%	825	2,6%	0,5%
Triticale	388	1,2%	441	1,4%	0,2%
Soybean	1.771	5,5%	1.692	5,3%	-0,2%
Pulses	48	0,1%	37	0,1%	0,0%
Horticulture	978	3,0%	1.126	3,5%	0,5%
Other arable crops	682	2,1%	553	1,7%	-0,4%
Ryegrass	34	0,1%	377	1,2%	1,1%
Grass herbages	518	1,6%	663	2,1%	0,5%
Legume herbages	6	0,0%	45	0,1%	0,1%
Mixed herbages	353	1,1%	477	1,5%	0,4%
Alfalfa	6.717	20,9%	6.145	19,1%	-1,8%
Other temporary grassland	4.995	15,5%	4.830	15,0%	-0,5%
Fallow land	533	1,7%	615	1,9%	0,3%
Non-eligible surfaces	1.048	3,3%	813	2,5%	-0,7%
Total Balanced Farmland	32.155	100,0%	32.155	100,0%	0,0%

Source: own elaboration on SIARL data

Group 2 – Eligible and compliant farms

Farmland use	2014 (before greening) Area (hectares) (a)	2014 (before greening) Area (%) (b)	Average 2015-2017 (after greening) Area (hectares) (c)	Average 2015-2017 (after greening) Area (%) (d)	Difference Area (%) (e) = (d) - (b)
Maize	21.740	29,1%	17.998	24,1%	-5,0%
Maize for silage	7.446	10,0%	8.205	11,0%	1,0%
Rotation ryegrass + maize for silage	4.172	5,6%	4.738	6,3%	0,8%
Wheat	8.964	12,0%	10.923	14,6%	2,6%
Barley	1.590	2,1%	2.010	2,7%	0,6%
Triticale	1.352	1,8%	1.014	1,4%	-0,5%
Soybean	6.237	8,3%	6.271	8,4%	0,0%
Pulses	53	0,1%	172	0,2%	0,2%
Horticulture	1.699	2,3%	1.998	2,7%	0,4%
Other arable crops	1.801	2,4%	1.377	1,8%	-0,6%
Ryegrass	146	0,2%	793	1,1%	0,9%
Grass herbages	2.317	3,1%	1.173	1,6%	-1,5%
Legume herbages	9	0,0%	87	0,1%	0,1%
Mixed herbages	435	0,6%	542	0,7%	0,1%
Alfalfa	11.950	16,0%	12.859	17,2%	1,2%
Other temporary grassland	3.711	5,0%	3.558	4,8%	-0,2%
Fallow land	401	0,5%	372	0,5%	0,0%
Non-eligible surfaces	789	1,1%	718	1,0%	-0,1%
Total Balanced Farmland	74.810	100,0%	74.810	100,0%	0,0%

Source: own elaboration on SIARL data

Group 3 – Eligible and not compliant farms

Farmland use	*2014 (before greening)*		*Average 2015-2017 (after greening)*		*Difference*
	Area (hectares) *(a)*	*Area (%)* *(b)*	*Area (hectares)* *(c)*	*Area (%)* *(d)*	*Area (%)* *(e) = (d) - (b)*
Maize	38.097	45,0%	29.241	34,5%	-10,5%
Maize for silage	13.403	15,8%	11.838	14,0%	-1,8%
Rotation ryegrass + maize for silage	6.376	7,5%	8.415	9,9%	2,4%
Wheat	7.832	9,2%	10.217	12,1%	2,8%
Barley	1.956	2,3%	3.025	3,6%	1,3%
Triticale	2.398	2,8%	1.971	2,3%	-0,5%
Soybean	1.405	1,7%	4.458	5,3%	3,6%
Pulses	31	0,0%	126	0,1%	0,1%
Horticulture	3.264	3,9%	3.423	4,0%	0,2%
Other arable crops	1.093	1,3%	984	1,2%	-0,1%
Ryegrass	103	0,1%	543	0,6%	0,5%
Grass herbages	694	0,8%	896	1,1%	0,2%
Legume herbages	3	0,0%	115	0,1%	0,1%
Mixed herbages	346	0,4%	326	0,4%	0,0%
Alfalfa	1.781	2,1%	3.482	4,1%	2,0%
Other temporary grassland	4.549	5,4%	4.239	5,0%	-0,4%
Fallow land	97	0,1%	872	1,0%	0,9%
Non-eligible surfaces	1.251	1,5%	508	0,6%	-0,9%
Total Balanced Farmland	84.680	100,0%	84.680	100,0%	0,0%

Source: own elaboration on SIARL data

- We integrate remote sensing into ex-post assessment of CAP effects
- We evaluate the effects of CAP greening on soil quality parameters
- Integration of administrative data with remote sensing offers full land coverage
- Membership in greening groups has moderate effects on soil quality
- The approach is reproducible in time and space

The authors declare that there is no conflict of interest all ethical practices have been followed in relation to the development, writing, and publication of the article.