From regional to parcel scale: A high-resolution map of cover crops across Europe combining satellite data with statistical surveys

Arthur Fendrich, Francis Matthews, Elise van Eynde, Marco Carozzi, Zheyuan Li, Raphael d'Andrimont, Emanuele Lugato, Philippe Martin, Philippe Ciais, Panos Panagos

Science of the Total Environment 873 (2023) 162300

Contents lists available at ScienceDirect

Science of the Total Environment

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

From regional to parcel scale: A high-resolution map of cover crops across Europe combining satellite data with statistical surveys

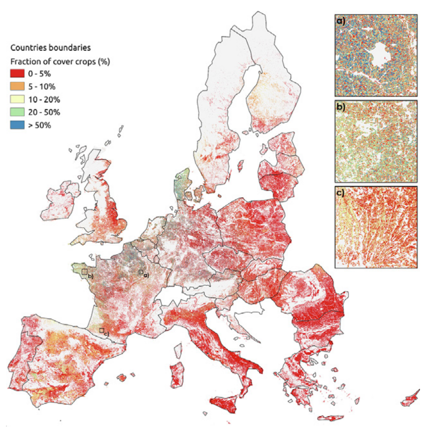
Arthur Nicolaus Fendrich[a,b,c,*], Francis Matthews[a,d], Elise Van Eynde[a], Marco Carozzi[c], Zheyuan Li[e,f], Raphael d'Andrimont[a], Emanuele Lugato[a], Philippe Martin[c], Philippe Ciais[b], Panos Panagos[a]

[a] European Commission, Joint Research Centre (JRC), Ispra 21027, Italy
[b] Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ-UPSACLAY, Gif sur Yvette 91190, France
[c] Université Paris-Saclay, INRAE, AgroParisTech, UMR SAD-APT, 91120, Palaiseau, France
[d] KU Leuven, Unit of Geography and Tourism, Celestijnenlaan 200e, Leuven 3001, Belgium
[e] School of Mathematics and Statistics, Henan University, Kaifeng 475001, China
[f] Department of Statistics and Actuarial Science, Simon Fraser University, University Dr W, 8888, Burnaby, BC V5A 1S6, Canada

HIGHLIGHTS

- Cover crops play a relevant role for conservation in the Common Agricultural Policy 2023–2027.
- Despite the potential of farmers' declarations, detailed data on cover crops are scarce.
- The disaggregation of survey data on cover crops using satellite information is proposed.
- Quantitative validation at the parcel level in France yielded an area under the curve of 0.74.
- Despite limitations, this work is the first effort to obtain a cover crop map at European scale.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Paulo Pereira

Keywords:
Cover crops
Common agricultural policy
Spatially-explicit
Disaggregation
European Union

ABSTRACT

The reformed Common Agricultural Policy of 2023–2027 aims to promote a more sustainable and fair agricultural system in the European Union. Among the proposed measures, the incentivized adoption of cover crops to cover the soil during winter provides numerous benefits such as improved soil structure and reduced nutrient leaching and erosion. Despite this recognized importance, the availability of spatial data on cover crops is scarce. The increasing availability of field parcel declarations in the European Union has not yet filled this data gap due to its insufficient information content, limited public availability and a lack of standardization at continental scale. At present, the best information available is regionally aggregated survey data, which although indicative, hinders the development of spatially accurate studies. In this work, we propose a statistical model relating Sentinel-1 data to the existence of cover crops at the 100-m spatial resolution over the entirety of the European Union and United Kingdom and estimate its parameters using the spatially aggregated survey data. To validate the method in a spatially-explicit way, predictions were compared against farmers' registered declarations in France, where the adoption of cover crops is widespread. The results indicate a good agreement between predictions and parcel-level data. When interpreted as a binary classifier, the model yielded an Area Under the Curve (AUC) of 0.74 for the whole country. When the country was divided into five regions for the evaluation of regional biases, the AUC values were 0.77, 0.75, 0.74, 0.70, and 0.65 for the

* Corresponding author at: European Commission, Joint Research Centre, Via E. Fermi 2749, Ispra 21027, Italy.
E-mail addresses: arthur.fendrich@ec.europa.eu arthur.fendrich@lsce.ipsl.fr (A.N. Fendrich).

http://dx.doi.org/10.1016/j.scitotenv.2023.162300
Received 2 November 2022; Received in revised form 12 January 2023; Accepted 13 February 2023
Available online 22 February 2023

North, Center, West, East, and South regions respectively. Despite limitations such as the lack of data for validation outside France, and the non-standardized nomenclature for cover crops among Member States, this work constitutes the first effort to obtain a relevant cover crop map at a European scale for researchers and practitioners.

1. Introduction

The Common Agricultural Policy (CAP), 2023–2027 introduces a series of reforms in key areas for the European Union (EU) Member States (MSs). These reforms aim to promote more intelligent, competitive, sustainable, and diversified agriculture and forestry, develop the socioeconomic structure of rural areas, and protect the environment. In particular, climate action, including carbon sequestration, is one of the EU's main priorities since it helps achieve the commitments made under the Paris Agreement. In this sense, conservation agriculture practices can help promote a shift of existing agricultural fields towards more sustainable systems. Such practices include minimizing soil disturbance, maintaining permanent ground cover, and adopting combined rotations (Hobbs et al., 2007; Palm et al., 2014). Among the existing options, the adoption of cover crops (CCs) constitutes one important example of a conservation measure to protect the soil surface against soil erosion. CCs are plants grown with the purpose of protecting the soil and improving its quality and health (Delgado et al., 2017). If CC biomass is incorporated into the soil, it positively impacts agronomic and environmental outcomes, such as soil carbon stocks (McClelland et al., 2021; McDaniel et al., 2014; Ruis and Blanco-Canqui, 2017). The adoption of CCs also has additional benefits, such as reducing nutrient leaching (Nyakatawa et al., 2001), improving soil structure through increased infiltration and water holding capacity (Nyakatawa et al., 2001; Panagos et al., 2015; Smith et al., 1987), and improving the biological quality of the soil (Muhammad et al., 2021; Kim et al., 2020). CCs thereby play a strategic role in soil conservation policies since they are one of the few components of the erosion process that can be directly mitigated through human interventions by farmers and policy-makers (Panagos et al., 2015). For their general contribution to soil health and climate change mitigation, CCs have been receiving increasing attention (Koudahe et al., 2022).

A key process justifying the implementation of CCs is soil erosion. Accelerated by human activities for more than 4000 years (Jenny et al., 2019), soil erosion is currently the most common form of land degradation in the world, affecting over 1 billion ha of the Earth's surface (Borrelli et al., 2020). The average erosion rates worldwide are estimated to be 2.8–2.9 t ha^{-1} yr^{-1} (Borrelli et al., 2017), around 3 to 20 times higher than the natural soil formation rates, of 0.15–1 t ha^{-1} yr^{-1} (FAO, 2015). For the future, model forecasts indicate that current erosion rates may further increase from 30 to 66 % worldwide (Borrelli et al., 2020), with local variations. In Europe, for example, the expected increase in soil erosion rates ranges from 13 to 22.5 % (Panagos et al., 2021). As an established tool to combat erosion (Borrelli et al., 2022), CC uptake has been shown to decrease erosion rates by 15 to 23 % (Nyakatawa et al., 2001; Verstraeten et al., 2006) and mitigate on- and off-site damages (Montgomery, 2007), such as degradation of arable areas, pollution and eutrophication of rivers and lakes, and contamination of aquatic and marine ecosystems (Poesen, 2017; Amundson et al., 2015).

With the modern-day availability of Integrated Administration and Control System (IACS) data in the EU, another potential source of CC information is farmers' declaration, such as the Land Parcel Identification System (LPIS) and the Geospatial Aid Application (GSAA). These large-scale spatial databases contain annual declarations made for CAP measures by EU farmers, and their elements correspond to the boundaries of agricultural parcels and their corresponding main crops. While seemingly promising, these can be managed and operated independently at the regional or country level, so reporting secondary crops or CCs is sometimes possible but not consistent across MSs. Therefore, despite the increasing importance given to CCs, the availability of data about their use and presence remains primarily limited to coarse-scale statistical surveys. Among the European Union MSs, the best information available about CCs can be found in the Farm Structure Survey (FSS), which surveyed the 27 European countries down to the NUTS3 (the third level of the Nomenclature of Territorial Units for Statistics classification) regional level every 3 or 4 years (European Commission, 2022d), and the Survey on Agricultural Production Methods carried out in 2010 (European Commission, 2022f) at farm scale. While such information can be helpful for applications permitting spatial generalization, it does not provide sufficient spatial detail to allow, for example, a precise evaluation of the local impacts of CCs on soil erosion and carbon content.

Some consequences of lacking detailed and georeferenced CC data can be found throughout continental-scale modeling exercises. For example, when modeling the C-factor of the Universal Soil Loss Equation, Panagos et al. (2015) used the CC area at NUTS2 level, the best available information at a European scale comprising 216 regions in the 27 MSs plus the United Kingdom (UK). In this case, the authors adopted the same value for all pixels within a NUTS2 region. Such an assumption is very strong and unlikely to represent field reality since it implicitly assumes that all detailed units (e.g., farms or parcels) behave identically. A similar assumption was also adopted in coupling erosion and carbon models to assess the combined effect of good agricultural and environmental practices on erosion and carbon budget at the national level (Borrelli et al., 2016). While the same assumption was not adopted by (Borrelli and Panagos (2020a), the lack of spatially detailed information about CCs was the limiting factor in their input datasets. In all these examples, a detailed CC map would have represented an improvement in the characterization of physical processes.

While several initiatives to gather CC data exist, the information is not always publicly available, and no harmonized dataset is available at the EU level. This latter issue is difficult because the CC eligibility for CAP subsidies differs in each MS/region because of policy, management, and climate reasons. At the EU scale, the continent-wide Land Use/Cover Area frame statistical Survey Soil (LUCAS) field survey (European Commission, 2022b; Orgiazzi et al., 2017) offers limited opportunity because it is predominantly made during the main cropping season and therefore cannot capture CC information. In addition to this, optical satellite imagery for the spatial monitoring of CC uptake often faces the challenge of a high incidence of clouds during rainy winter months in which CCs can be detected (Beriaux et al., 2021). The combination of these factors limits the availability of input data for automated computational techniques of CC detection with European-scalability. Even though the methodology for survey collections might change in the future, these reasons are probably why most recent computationally and data-intensive efforts using new technologies such as Copernicus Sentinel-1/2 satellites focus on mapping the main crop of the cropping season (d'Andrimont et al., 2021; Meroni et al., 2021).

Given the outlined scarcity of primary and auxiliary data, one possible alternative is using spatial disaggregation methods. Such a category of methods attempts to reconstruct the fine-resolution information from areal features (e.g. regional statistics polygons) to allow detailed spatial analysis (Comber and Zeng, 2019). With different approaches and assumptions, the application of disaggregation methods can be found in different disciplines, such as soil mapping (Møller et al., 2019), epidemiology (Utazi et al., 2018), disease mapping (Weiss et al., 2019), demography (Jia and Gaughan, 2016), among others.

This work focuses on the problem of disaggregating existing CC data in the European MSs and the UK from NUTS2 level to a finer spatial resolution using satellite data. The objectives are: a) to develop the first dataset of predicted CC occurrence at a high (i.e., 100-m) spatial resolution for Europe for 2016–2017; b) to develop the assumptions of a statistical model for the occurrence of CCs, which can be transposed to other study areas; c) to validate these newly produced maps quantitatively against parcel-scale observations; and d) to discuss some possible implications and applications of this new CC dataset. A statistical model for the occurrence of CCs is first built. The model uses 12-day median composites of remotely-sensed synthetic aperture radar (SAR) data from Sentinel-1. Its parameters are estimated iteratively in such a way that the aggregation of predicted values approximates the values reported by statistical surveys at the coarse regional level. Then, parcel-level data from France's Registre Parcellaire Graphique (RPG) (Institut National de l'Information Géographique et Forestière, 2022), one of the few systems in European countries where declarations of CCs are public, is used to validate model predictions at the finest possible (i.e., parcel) spatial scale. Such an analysis allows understanding the predictions' strengths and limitations. Finally, some policy-relevant aspects of CCs are discussed.

2. Methods

2.1. Study area

The study area includes all croplands of the current EU MSs plus the UK, which covers about 156 million ha and 67 million parcels. According to the Farm Field Survey (FSS), CC application has increased from 6.5 % of all agricultural lands in 2010 to 8.9 % in 2016 (Borrelli and Panagos, 2020b). The adoption of CCs is currently an underused farming practice (Kathage et al., 2022a) which is likely to increase in the EU in the future.

2.2. Input data

Multi-temporal Sentinel-1 data was used to monitor changes in the landscape surface condition through time. The Sentinel-1 SAR constellation revisits the EU territory with a minimum 6-day revisit period since 2016 (until the Sentinel-1B defect in December 2021), providing a dense temporal time series for phenological monitoring. Compared to optical sensors (e.g. Sentinel-2), its microwave backscatter retrieval is practically uninfluenced by atmospheric conditions. In the agricultural context, microwave backscattering is sensitive to crop canopy structure, which means it can detect plant growth at the parcel spatial resolution. For these reasons, Sentinel-1 data offers a consistent source of plant phenological data in winter months for mapping CCs in combination with other computational techniques. Spatio-temporally consistent time-series of Sentinel-1 data were generated as follows: i) analysis-ready Sentinel-1 data was accessed in Google Earth Engine (Gorelick et al., 2017) (i.e., COPERNICUS/S1_GRD), already pre-processed to account for thermal-noise removal, radiometric calibration, and terrain correction; ii) a temporal stack of 31 rasters of the VV and VH bands was created, for both ascending and descending orbits, giving 12-day composites from the 10th of August 2016 to the same date in 2017; iii) the cross-polarization ratio (CR) was calculated as CR = VV/VH, giving 31 spatio-temporal layers for the period of study. Finally, the data was resampled using the median statistic to a 100-m spatial resolution, the target resolution adopted for the new spatially explicit CC dataset for reasons of computational trade-offs vs spatial resolution (data size) in spatial disaggregation models.

The choice of the CR time series as an indicator of the existence of CCs came from recent evidence highlighting the correlation of this index with the more well-known Normalized Difference Vegetation Index in the context of crop phenology (Meroni et al., 2021), crop dynamics (Veloso et al., 2017) and vegetation dynamics (Ma et al., 2022; Vreugdenhil et al., 2020). The choice of the period of analysis (i.e., 10th of August 2016–2017) was mainly driven by the need to match the period of most recent the coarse CC data available in Europe. In this case, it corresponds to the 2016 dataset on CCs per NUTS2 region published by the European Commission, which estimates the total CC area in arable lands (European Commission, 2022c). The exact definition of the variable used (i.e., "cover or intermediate crops") reads:

"An area of arable land on which plants are sown specifically to reduce the loss of soil, nutrients and plant protection products during the winter or other periods when the land would otherwise be bare and susceptible to losses. The economic interest of these crops is low, and the main goal is soil and nutrient protection. (…) These crops should not be mistaken for normal winter crops or grassland" (European Commission, 2022e).

[(NPI333YA)]

In order to filter the location of arable lands, the CORINE Land Cover 2018 at a 100-m spatial resolution and with a minimum mapping unit of 25 ha was used (Copernicus, 2022). By doing so, predictions of CC existence were restriced to a spatial domain where agricultural activity was previously detected in an external dataset. From the complete CORINE database, the class arable land was created by selecting the classes "non-irrigated arable land", "permanently irrigated land", and "rice fields". Additionally, other agricultural classes were included[1] to avoid having an overly restrictive spatial domain. Even though the period of CORINE does not match the period of 2016–2017, it can be accepted as adequate under the assumption that no drastic changes occurred in European arable lands from 2016 to 2017, the year where CORINE Land Cover 2018 images were mainly taken (Büttner and Kosztra, 2022). Then, because the amount of arable land calculated using FSS (European Commission, 2022c) and CORINE are different, the CC area consistent with CORINE was defined by multiplying the original CC area from FSS by the ratio of CORINE to FSS arable land area (A), giving $CC_{Corine} = CC_{FSS} \cdot A_{Corine}/A_{FSS}$. Such an arable land ratio varies from 0.02 to 3.69, with an average of 1.02 (Supplementary Material, SM 1).

2.3. Disaggregation model

In order to disaggregate CC information from the (215) NUTS2 to the (156 million) pixels at 100-m resolution, the estimation method proposed by Fendrich et al. (2022) was used. As in any other regression modeling approach, the technique consists of first making assumptions about the relationships between explanatory and dependent variables at the fine scale (Fig. 1).

It was assumed that the fraction of CCs in a pixel varies in three dimensions, namely: space, time, and the observed Sentinel-1 CR. Each of these dimensions are assumed to vary (possibly smoothly) with nonlinear interactions between them. Such an assumption is equivalent to assuming that the interpretation of a given CR time series varies so that it might indicate the existence of CCs in one particular place and its neighboring areas but not in other more distant regions. In mathematical notation, this can be represented as:

$$y_i = g\left[\sum_t s(lat_i, long_i, t, CR_i(t))\right] + \varepsilon, \quad \varepsilon \sim \mathrm{N}(0, \sigma^2) \tag{1}$$

with y_i being the fraction from 0 to 100 % of CCs within pixel indexed i; lat_i, $long_i$, t, $CR_i(t)$ being latitude, longitude, time and the CR time-series with a timestep of 12 days, respectively. The function $g(x) = 1/(1 + exp(x))$ was chosen to be the link function, since y_i varies from 0 to 100 %. The term ε is the residual term, assumed to be normally distributed, and $s(\cdot)$ is the joint function to be estimated from the data. Penalized smoothers that can be represented using basis expansions and a penalty matrix to control function smoothness are a common choice for the one-dimensional smoothers inside $s(\cdot)$, and the interaction is often represented as tensor products (Wood, 2017). Under such a representation, Eq. (1) could be recognized as a nonlinear mixed model (Bates and Watts, 1988; Wolfinger, 1993).

In the representation of Eq. (1), the fraction of CC at the pixel level is assumed to be a random variable. Consequently, the sum of the CC fractions in the pixels that belong to a NUTS2 region creates another random variable

[1] Namely, "Vineyards", "Fruit trees and berry platforms", "Olive groves", "Annual crops associated with permanent crops", "Complex cultivation patterns", and "Land principally occupied by agriculture with significant areas of natural vegetation".

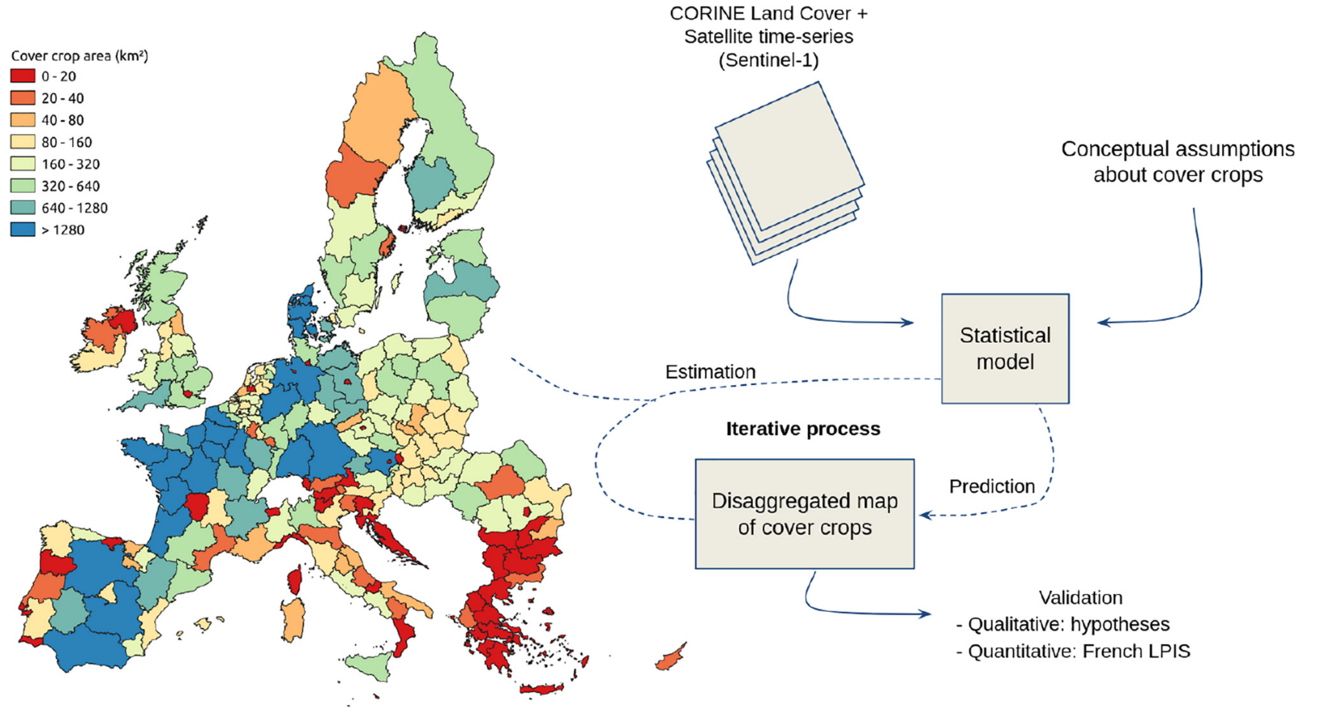


Fig. 1. Summary of the method used in the present work. Left: area of cover crops in 2016 visualized using the cover crop extent per region from the FSS (European Commission, 2022d).

representing the CC area at the NUTS2 level. It can be shown that, in this case, a second nonlinear mixed model can be derived for the NUTS2 level, and it preserves the original parameters necessary to construct the pixel-level smoothers of Eq. (1). Such a NUTS2-level model allows us to attempt to reconstruct the pixel-level information by performing regression analysis on the coarse data. During the parameter estimation phase, an optimization procedure handles the tradeoff between i) approximating the reconstructed to the observed values at the NUTS2 level; and ii) enforcing the mathematical assumptions for the function $s(\cdot)$ described above (i.e., that the one-dimensional smoothers are continuous functions, that the effect of the CR varies in space and time etc.).

Parameter estimation is done through a numerical optimization procedure, which maximizes the chance of observing the aggregated data given the assumptions made at the pixel level (i.e., maximum likelihood estimation). In this work, we slightly modified the original method of Fendrich et al. (2022) to use a quadratic instead of a first-order approximation to the log-likelihood in each iteration. Such a modification led to the natural interpretation of each iteration as a Newton-Raphson step of the original guess towards the maximum a posteriori estimates, similar to the scheme proposed by Rossell et al. (2021).

As disaggregation problems are fundamentally undetermined with infinite solutions, choosing proper assumptions is essential for narrowing the possibilities. Therefore, $s(\cdot)$ was assumed to be a tensor product, and several possibilities for the one-dimensional smoothers of Eq. (1) were tested, including thin-plate, cubic and P- splines, with different basis dimensions and penalty orders. The SM 4 shows the results for the four best models found during model selection according to the area under the curve (AUC) performance metric (see Section 2.4.2 for details about the validation dataset). The final alternative was chosen to be four P-splines (Wood, 2016), with basis dimensions 9 for the longitude and 8 for latitude (SM 3), time and CR. After the estimation of model parameters, the expected value of $s(\cdot)$ given the parameters was calculated and plotted to allow the visualization of the estimated one-dimensional smoothers. The predicted CC fraction and the corresponding uncertainty were calculated as the point estimate of the median and the bounds of the 90 % confidence intervals for the expected value of the response variable of Eq. (1) given the parameters, respectively.

2.4. Validation

2.4.1. Qualitative validation

Validating a disaggregation model is a challenging task. First and foremost, no fine-scale information is available for the whole study area. Otherwise, if such data were available for comparison against model predictions, it would be preferable not to use a disaggregation model but instead to use such information as an input to traditional mapping techniques. Given these considerations, the validation procedure was split into: i) a qualitative validation of the model's internal behavior; and ii) a quantitative validation of the external behavior (classification accuracy), consisting of a comparison against publicly available parcel data for France.

Developing a reasonable interpretation of the results obtained from a prediction is a vital step in regression modeling (Nisbet et al., 2009). Therefore, in the qualitative validation, the estimated smoother (Eq. (1)) was investigated to check its representation of logical aspects of CC phenology that should reasonably result in their identification. To do so, necessary but not sufficient conditions were first established for the model to be realistic. The first condition is for the model to assign a higher weight to the satellite observations made during cold months compared to hot months since that is when CCs can be observed in the field. The second condition is for the model to capture the different regional patterns of CC production properly based on the known a priori information. This secondary condition checked, for example, if the model could capture the regionalized temporal pattern of CCs in France, where sowing dates may vary from the end of July in the Moselle region (North-West France) until the beginning of November in the Pyrénées-Atlantiques (South-West France) (Journal Officiel de la République Française, 2018). In order to identify this information from Eq. (1), a first-order expansion of g(x) around $x = 0$ was used to calculate the approximate marginal contribution of the CR to the predicted fraction of CCs at the pixel level. Then, since the model of Eq. (1) is an interaction between several variables, only the marginal results were presented.

2.4.2. Quantitative validation

To complement the qualitative evaluation of the model, a quantitative validation was also developed to confront model predictions against comparable parcel-scale declarations made by farmers in the CAP context.

Such a validation started by searching for the most accurate classification of parcels between those with or without CCs in 2016. To do so, databases of European countries (Schneider et al., 2021) were inspected, and it was found that in most cases, farmers are only required to report their parcel's main crop. Nevertheless, in some countries, such as Portugal and France, farmers can, but are not required to, report multiple crops per year.

For the present work, the French RPG dataset was chosen, which gathers all annual declarations made by farmers to receive CAP subsidies (Levavasseur et al., 2016) and provides the cultivated crop or crop groups for each farmer at the parcel scale. This private and confidential information concerns the parcels belonging to farmers provided by the French Ministry of Environment. The RPG dataset was used to validate the predictions for three reasons. First, France is representative for being the largest MS, with a large share of its territory devoted to agricultural activity, and for covering several environmental conditions such as continental, Mediterranean and oceanic climates (Ols et al., 2020). Second, as shown in Fig. 1, France contains a large area of CCs, and the inclusion of CCs in the cropping system is, in some cases, mandatory to comply with the Nitrate directive (Official Journal of the European Union, 1991) and the CAP Greening (Kathage and Domínguez, 2019). Third the French Ministry of Agriculture and Food Sovereignty provides a list of the 44 crops that are considered to be CCs (Ministère de l'Agriculture et de l'Alimentation, 2022) to comply with the dedicated Ecological Focus Areas (EFAs) to safeguard and improve biodiversity, requested by the green direct payment of the CAP. Such a list facilitates farmers' declarations, which have to declare two of these crops in a mixture per parcel to be retained as EFAs.

However, the construction of the validation dataset faces some challenges. As the dataset is attached to compliance with regulations, motivations to meet the minimum requirements in the CAP can bias the area of CCs declared by farmers. For example, a farmer with more CCs than the minimum requirements may feel inclined to declare only the minimum, which would generate a bias towards underestimating CCs. Conversely, a farmer with half the minimum requirement could tend to overestimate the area of CCs in the declaration. Besides, the declaration of CCs is not mandatory in France due to other EFAs that can be present on the farm. Therefore, it is not possible to affirm with certainty whether a random parcel contains CCs or not. This problem was overcome by applying additional filters on the dataset to increase the confidence in the presence or absence of CCs on the parcels. Such a procedure aimed at obtaining a sample of parcels that could be used with greater certainty to validate the model at the scale of reference using the available data. The filters are described next.

First, three groups of parcels according to the different levels of uncertainty of CC existence were defined a priori: parcels without CCs in farms that did not declare CCs (PnFn, lower certainty); parcels without CCs in farms that declared CCs (PnFy, medium certainty); parcels with CCs in farms that declared CCs (PyFy, higher certainty). The sampling procedure for the PnFn group consisted of first filtering farms without CCs declared, and then randomly sampling parcels inside these farms. This group has the lowest certainty among the three because since no CCs are declared within the farms, the farmers may have chosen to declare other EFAs in their farms instead of CCs. In this case, some farms might have CCs on the field but not registered in the database.

For PnFy and PyFy, two filters were added at the farm level: i) only those with a sufficiently high ratio of surface as CCs on the farm (i.e., 20 %), compared to the minimum cover limit imposed by the CAP (5 %); and ii) only those with a large area (over 50 ha) to increase the chance of having a heterogeneous configuration of parcels. These filters assume that if farmers have declared well beyond the required minimum of 5 %, then they have declared the entire area of CCs in their parcels. However, since this assumption may not be true in some cases, more confidence can be assumed for the presence (PyFy) than the absence (PnFy) of CCs in this case, which justify their certainty assigned above. Finally, parcels whose polygons fall entirely inside the CORINE land cover classes described in Section 2.2 were randomly sampled for each of the three classes.

After these procedures, 19,390 parcels were used for validation: 8441 PnFn, 11,112 PnFy, and 12,256 PyFy, each with information about CC's existence and its main crops in 2016. Among the 19,553 sampled parcels without CCs (i.e., PnFn + PnFy), 5968 had a winter commercial crop as the main crop. Finally, because the size and shape of each land parcel are different, they were compared by calculating the weighted median value of all model predictions intersecting their boundaries. Boxplots were plotted to compare the distribution of predictions within the three classes of parcels described above, and receiver operating characteristic curves (i.e., ROC curves) were generated to assess the model's discrimination abilities.

For each ROC curve, the corresponding AUC was calculated. Their interpretation was made using the classification proposed by Bera et al. (2020). Weak: $0.5 \leq \text{AUC} < 0.6$; Moderate: $0.6 \leq \text{AUC} < 0.7$; Good: $0.7 \leq \text{AUC} < 0.8$; Very good: $0.8 \leq \text{AUC} < 0.9$; Perfect: $\text{AUC} \geq 0.9$. Analogous quantitative-qualitative relationships can be found in works such as Roy et al. (2020); Panahi et al. (2022), and others.

2.5. Implementation

All codes used were written in R 4.2.0 (R Core Team, 2022). The P-splines implementation used was that of the mgcv package (Wood, 2017), spatial data was manipulated using the terra package (Hijmans, 2022), and parallel computations were made with the snow package (Tierney et al., 2021). The row-wise Kronecker product and the log-likelihood were implemented with Rcpp and RcppArmadillo (Eddelbuettel and Balamuta, 2018; Eddelbuettel and Sanderson, 2014). All source codes and input datasets used, except for the parcels' boundaries used for validation, are publicly available in a dedicated GitHub repository: https://github.com/arthfen/CoverCrops/.

3. Results and discussion

3.1. A qualitative evaluation of the model response

For the marginal spatial effect, Fig. 2 shows the results for a winter month, December 2016. Six different locations are shown to highlight that each location has a distinct pattern for the effect of CR on the calculated fraction of CCs. For example, in Fig. 2, the model curves for South Italy and Bulgaria show that the low CC fractions (expected in these regions according to Fig. 1, left) are represented through a negative effect of the CC fraction for practically the full CR range. Conversely, for the other four locations where CCs tend to be more widespread, according to Fig. 1 (left), a different pattern can be observed. In these cases, Fig. 2 shows that an increase in the CC fraction is predicted for most of the CR range, and regional variability in the shape of the curve seems to be captured. As a reference for the CR values, Meroni et al. (2021) suggests a range from 0.1 to 0.2 in the trough (minimum) to 0.4–0.5 in the peak (maximum) of temporal series of crop fields in Europe (i.e., common wheat, rape, and maize), approximately the same range reported by Maurya et al. (2022) and Vavlas et al. (2020) in wheat fields in the North of India and the UK, respectively.

A distinct pattern in the marginal spatial effect can be seen in North-West France, where the effect on the CC fraction quickly peaks around 0.20 and decreases, and in Denmark, where the effect has a double peak, with the second starting around 0.20 and reaching its maximum around 0.30 (Fig. 2). In both cases, the fact that the model assigns the highest increase in CC fractions to relatively low CR values could mean that it is attempting to distinguish CCs from commercial winter crops, which could present a different signal due to fertilizer application. Such an interpretation is however uncertain since it could also represent other factors that affect the CR value, such as soil surface roughness from tillage operations (Vreugdenhil et al., 2020).

For the marginal temporal effect, the results of Fig. 3 are presented for Denmark (i.e., the blue line of Fig. 2). The Figure shows that, when weighing between available observations, the model attributes a relevant effect from the observations made in all months of the year, which can be seen by the large overall positive and negative effect on the CC fraction. For example, for low CR values ranging from 0.1 to 0.2, spring months

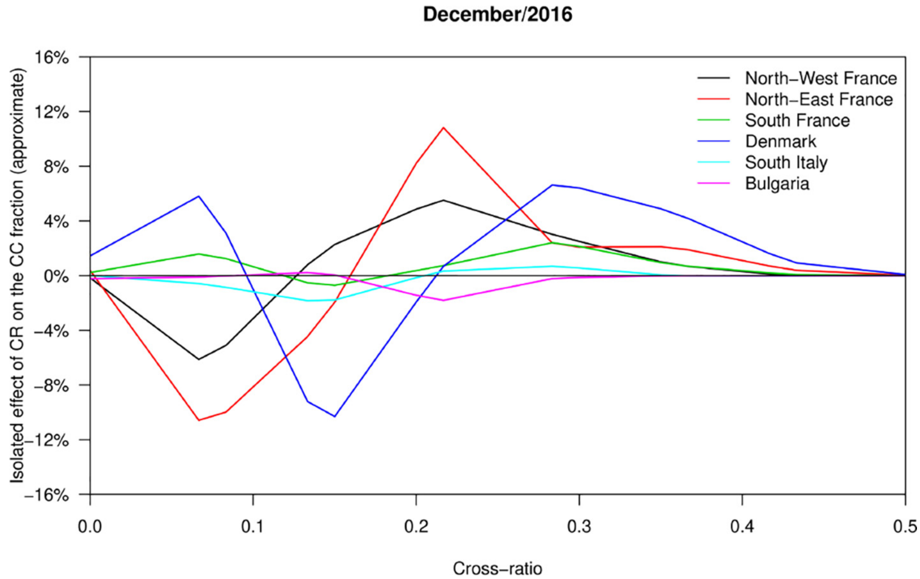


Fig. 2. Qualitative validation, marginal spatial effect: (approximate) isolated effect of the Cross-ratio (CR) on the predicted cover crop (CC) fraction as estimated by the model, for one of the 12-day median CR values in December/2016.

have a comparable effect to that of winter but in the opposite direction. Such a result resembles the observations made by Nowak et al. (2021), who showed that different spring-grown crops in France present different soil cover patterns during cold months. The hypothesis that cold months could have a higher influence than hot months seems to be valid only for CR values above 0.2. In practice, these results suggest that the approach to detect winter soil cover based on satellite data for winter months only proposed by Nowak et al. (2021) might be insufficient for CCs. It also suggests that a complete time series must be considered to separate CCs from commercial winter crops. In the latter case, the combination of complete time series and a flexible and probabilistic model could better identify what is a CC within a crop succession in a context-dependent way.

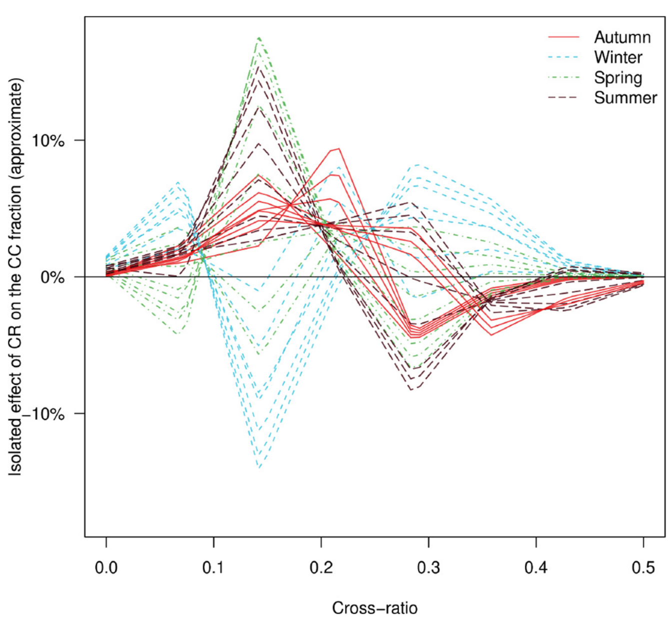


Fig. 3. Qualitative validation, marginal temporal effect: (approximate) isolated effect of the Cross-ratio (CR) on the predicted cover crop (CC) fraction as estimated by the model, for a random location in Denmark. Every line corresponds to one of the 12-day median CR values used for fitting the model.

3.2. A quantitative evaluation of model performance

Fig. 4 shows the calculation of weighted median predictions inside French parcels with accompanying declarations as described in Section 2.3, for two cases: all parcels included (i.e., $n \geq 1$ pixel, mean area of 5.2 ha), and only parcels intersecting more than 10 pixels included (i.e., $n \geq 10$ pixels, mean area of 9.6 ha). Such a distinction was made in order to evaluate the impact of sub-pixel confusion (a mix of landscape dynamics at the sub-pixel level) in the input data. For the case of all parcels included, the model could predict higher values for PyFy (right) than PnFn and PnFy (left, center), which indicates an ability to distinguish between parcels with and without CCs, respectively. The proximity between the distributions of PyFy and PnFy seems to indicate some confusion between parcels of farms known to have CCs. Such a result could indicate that the model performs better when the adoption of CCs is widespread at the farm level. However, such confusion seems to be minimized when only parcels intersecting more than 10 pixels are included in the analysis. In this case, PnFy approaches PnFn, with lower values predicted than PyFy.

Despite the results shown in Fig. 4, a question could be posed about the model's ability to distinguish between different types of winter cover. To evaluate this matter, Fig. 5 shows the analysis results that use crop information at the parcel level to separate winter commercial crops from CCs. Parcels with winter commercial crops were selected based on the sowing period of the main crop reported to the RPG, excluding CC species. It can be seen that a similar pattern as Fig. 4 is found, indicating a higher predicted fraction of CCs in parcels that truly contain CCs. Such a result is relevant, as it indicates that the assumption in Eq. (1) of a nonlinear effect of CR on the CC fraction is adequate to differentiate between the multiple signal patterns of soil cover in winter.

While Figs. 4 and 5 show the general pattern of model validation in France, regional patterns to understand the model's biases were also investigated. Fig. 6 shows the reclassification of the French NUTS2 areas into five contiguous regions (namely, South, East, North, West, and Center), and the model validation within these five areas. For each region, the ROC curve shows the true positive rate in the vertical axis against the false positive rate in the horizontal axis, summarizing the model performance when used as a binary classifier for the presence/absence of CCs in French parcels. The ROC curve is presented here, along with its corresponding AUC (Maimon and Rokach, 2010). In all regions and the whole of France, the AUC calculated is greater than 0.5, indicating that the model performs better than random chance. The results also show that, compared to the overall

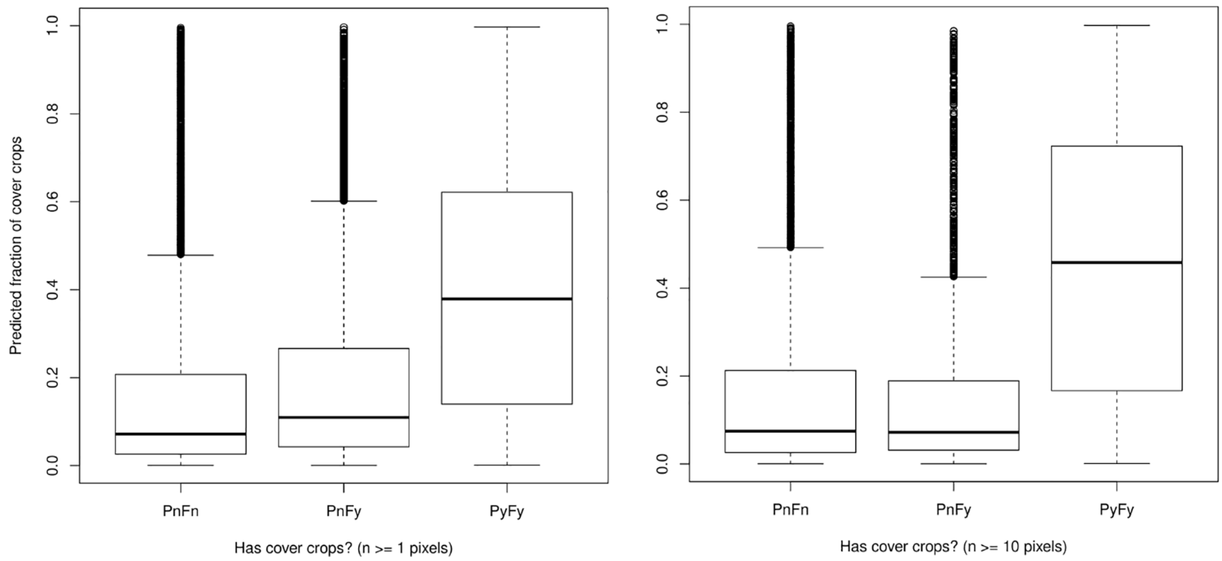


Fig. 4. Quantitative validation: distribution of the weighted median values calculated for French parcels in arable lands. Two cases: all parcels included (left), and only parcels intersecting more than 10 pixels included (right). PnFn are parcels without cover crops in farms that did not declare cover crops; PnFy are parcels without cover crops in farms that declared cover crops; PyFy are parcels with cover crops in farms that declared cover crops.

performance of AUC = 0.74, the West region of France yields an identical AUC, the North and Center regions yield a higher AUC (0.77 and 0.75, respectively), and the South and East regions a lower AUC (0.70 and 0.65, respectively). According to the classification presented in Section 2.4.2, this means that the overall classification and 4 out of 5 regions presented a good performance, while the East region had a moderate performance. The variable pattern across the country also shows that regional classification biases exist in the model and that more errors tend to happen in the South and East regions of France.

For the European-wide model (Fig. 6, bottom left), a threshold value was calculated as being the point in the ROC curve at which the Euclidean distance to the theoretical optimum is minimized (i.e., the top-left corner). The confusion matrix generated for this threshold value (Table 1) shows that when parcels contained CCs in the RPG, the model predictions were correct in 68.1 % of the cases. When the parcels did not contain CCs, the model was correct 74.6 % and 67.5 % for PnFn and PnFy, respectively. It also shows that, when the model predicted the inexistence of CCs, it was wrong in 22.1 % of the cases, and therefore correct in 77.9 % of them. When the model predicted the existence of CCs, it was correct in 59.2 % of the cases and thus wrong in 40.8 % of them. In this case, this model seems to assign more false positives to PnFy (25.6 %) than to PnFn (15.2 %), which can be possibly explained by some farmers not declaring all their CCs due to other EFAs in the farm (see Section 2.4). While these results suggest a limited power of the model in differentiating CC presence, it is necessary to highlight that: i) they implicitly assume a single threshold value for the whole of France, which can be suboptimal according to the heterogeneous results presented in Fig. 6, and ii) another threshold to improve model's precision could be chosen. In this context, the use of continuous values (e.g., Fig. 7) might be appropriate depending on the application intended.

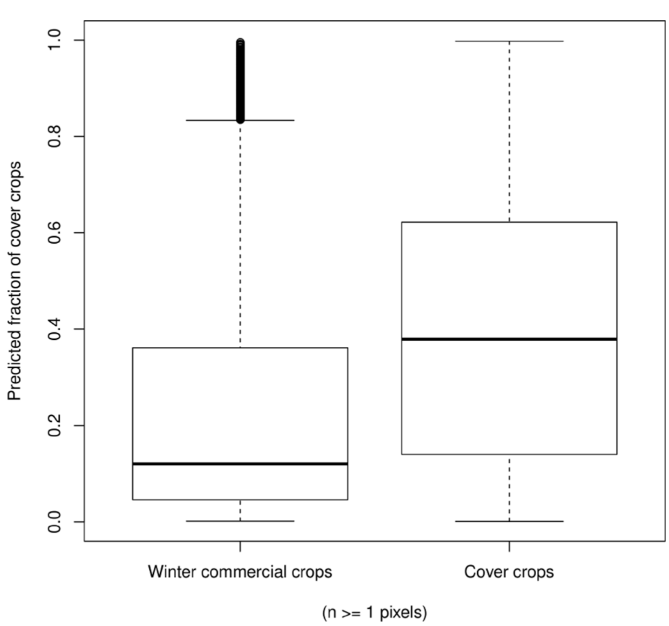


Fig. 5. Quantitative validation: distribution of the weighted median values calculated for French parcels with winter commercial crops and cover crops. The left boxplot merges all parcels with winter commercial crops within the groups PnFn (parcels without cover crops in farms that did not declare cover crops) and PnFy (parcels without cover crops in farms that declared cover crops).

Fig. 7 shows the model predictions for the arable land in the whole of Europe. The continental level map shows that the fraction predicted in most pixels with arable land tends to be zero and that CCs occur in concentrated regions but with local variations (Fig. 7, zoom). The marginal distributions show that CC fractions tend to be higher in the regions corresponding to Northern France, Germany, and Denmark, following Fig. 1. The standard deviation of model predictions, as well as the 0.05 and 0.95 quantiles are provided as Supplementary Materials (SM 6 and SM 6). These uncertainty margins were generated using the uncertainty around the estimated smoothers.

3.3. Future policy opportunities offered by spatially explicit cover crop predictions

The map produced in this work is the first European product at a 100-m spatial resolution predicting the fraction of CCs for the winter season of 2016–2017. Despite the limitations, with a lack of validation outside France and the existence of mixed satellite signals (Section 3.4), our dataset fills a gap and can be tested in several future applications ranging from fine-resolution input data for analysis of soil loss by water erosion, organic carbon sequestration, nutrient application, and to assist the implementation of soil conservation policies at a regional or national scale. Given that the FSS survey data for 2023/24 is still under collection at the time of writing, the current work also provides a reference methodology to be used for the spatial disaggregation of CCs and other policy-relevant information (e.g.

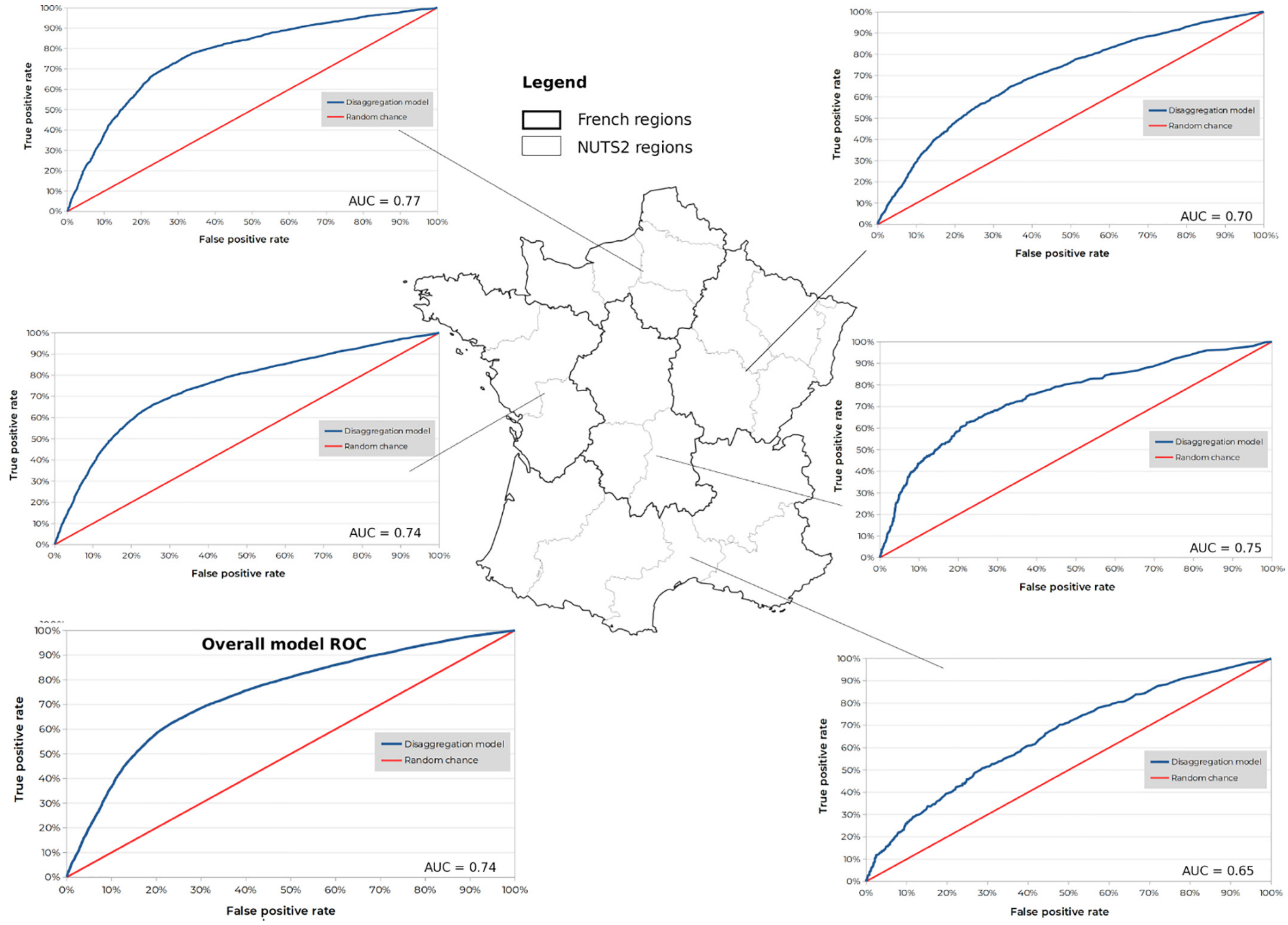


Fig. 6. Quantitative validation: regional variability of model predictions across France.

reduced tillage). Such a temporal analysis would enable the possibility of evaluating trends in CC adoption in recent years.

Two main policies are currently responsible for driving European farmers' adoption of CCs; the Nitrates Directive, and the CAP. Such policies, especially the Nitrates Directive, are now the strongest determinants of the adoption rates and intensities of CCs by farmers (Kathage et al., 2022b). Such factors overcome agronomy or environmental motivations, although this scenario might change in the future as CCs become economically incentivized, for example, for energy production (Launay et al., 2022). The spatial pattern of the vulnerable zones under the Nitrate Directive in France resembles the areas with high fractions of CCs in the map produced by this study (Fig. 7) (JRC, 2022; Tzilivakis et al., 2021). Furthermore, covering the soil in the most sensitive periods, a practice that includes adopting CCs, constitutes one of the GAECs introduced in the upcoming CAP 2023–2027 to protect soils in rainy susceptible seasons (e.g winter). In the proposal, the European Commission introduced various scenarios to protect soils using soil conservation measures such as the minimum green cover. The adoption of CCs with a cover rate of 75 % in one of the policy scenarios led to a reduction of soil erosion in arable lands and permanent crops by 15 and 30 %, respectively (European Commission, 2018; Panagos et al., 2021). Therefore, the new CAP 2023–2027 is likely to increase the area of agricultural land with CCs in the EU.

One of the Good Agricultural and Environmental Conditions (GAECs) of the new CAP prescribes that MSs must require farmers to apply crop rotation (European Commission, 2022g), a practice that also composes the list that can be supported by eco-schemes (European Commission, 2021), and a summary of the first strategic plans shows that 16 MSs include the practice of crop rotation with a secondary or CCs during the winter season to protect soils (European Commission, 2022g). In this sense, while the currently available CC data at the NUTS2 level collected by the FSS is useful to monitor the overall adoption rates and intensities of CCs, these coarse data are less suitable for evaluating the effect agricultural practices by farmers on the environmental and climatic objectives that are set by European policies (Matthews et al., 2023). By contrast, the use of techniques to develop refined spatial datasets as produced in this study is highly valuable to assess the influence of European policies on the on-ground-farmers decisions and the subsequent impact on climate and the environment. Since the new CAP emphasizes performance and results, assessments with detailed data like the one produced in this study will become more important.

Table 1
Quantitative validation: confusion matrix for the overall model in France. The results are based on the assumption of a single threshold for the whole country.

Situation	Model prediction: contains CCs?	
	No	Yes
PnFn	6296 (74.6 %)[a] (35.6 %)[b]	2145 (25.4 %)[a] (15.2 %)[b]
PnFy	7498 (67.5 %)[a] (42.4 %)[b]	3614 (32.5 %)[a] (25.6 %)[b]
PyFy	3907 (31.9 %)[a] (22.1 %)[b]	8349 (68.1 %)[a] (59.2 %)[b]

[a] Refers to the 'row' percentage.
[b] Refers to the 'column' percentage.

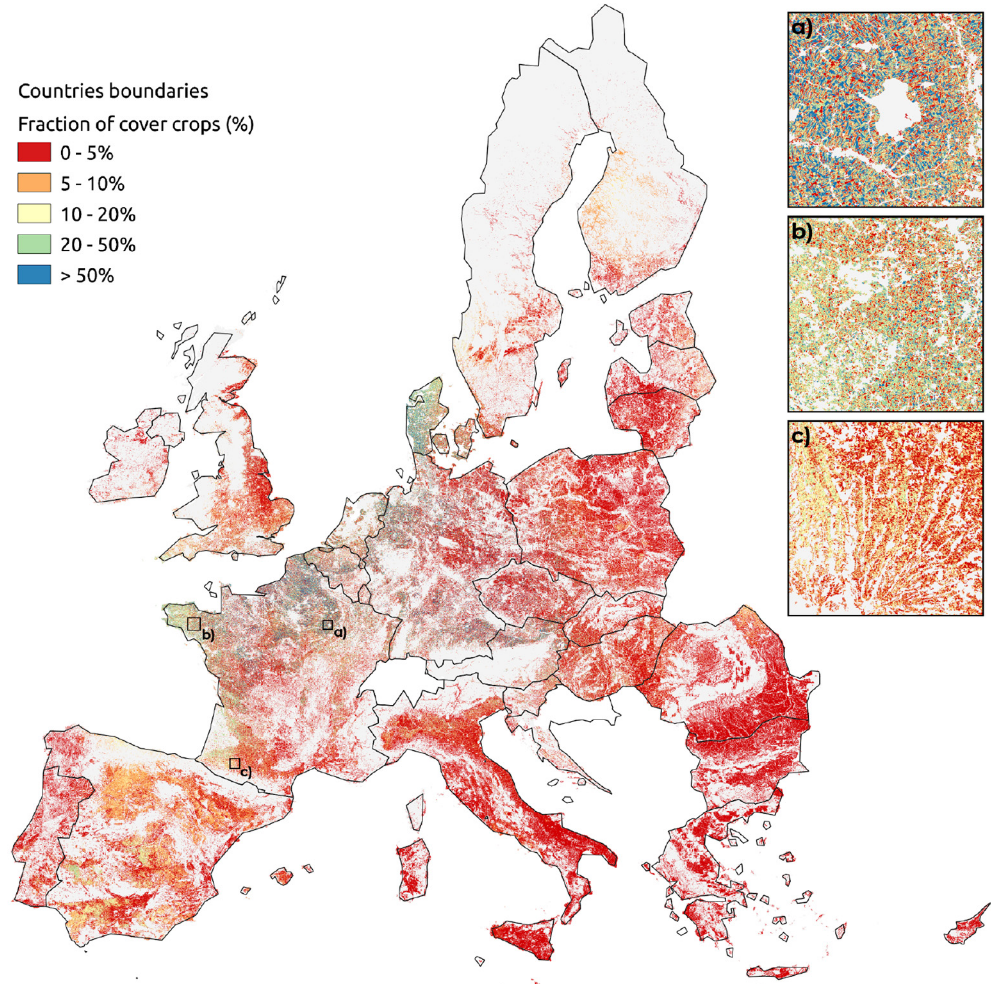


Fig. 7. Left: model predictions of the occurrence of cover crops (CCs) in Europe. Right: Three zooms: Predictions on the East, West and South of France (a, b and c, respectively).

3.4. Model limitations and future improvements

While the present results represent an advance for CC mapping, several limitations of the current approach can be listed and possibly developed in further works. The disaggregation model is strongly dependent on the FSS CC data at the NUTS2 level. In the last decades and once every ten years, the FSS has been a part of the Census of Agriculture, which guarantees high reliability of the values provided. However, the values available for 2016 and used in the current work correspond to sample estimates (European Commission, 2022a). This means that not only is the sample size used relatively small, but also the data itself contains uncertainties. Although the disaggregation model mitigates the problem by not enforcing the equality constraint but only approximating (Fendrich et al., 2022) (SM 2), more reliable data on CCs could potentially lead to more precise results. While a difference exists between the total CC area in our predictions and the FSS values rescaled by CORINE arable land (i.e., 13.5 Mha vs. 9.3 Mha, respectively), removing pixels with a low CC share (~ 10%) can be used as a post-processing step to approximate the two quantities. In the near future, area monitoring systems have the potential to provide more information about CCs at the EU level (Official Journal of the European Union, 2021).

Here, the spatial resolution of 100-m was adopted, approximately ten times coarser than the original resolution of the Sentinel-1 data. In practice, this means that the pixels in the present work can potentially have mixed signals in areas of high heterogeneity, which adds an extra layer of uncertainty where non-unique cultivations are found in close proximity. As shown in Fig. 4, this effect directly impacts the predictions for smaller parcels but tends to be minimized for larger units. Even though the model could be run at 10-m resolution, this would bring new computational and practical challenges since the number of pixels to predict would increase 100-fold. Such a high-resolution application would also require highly accurate land cover masks to be used as spatial input data. While the CORINE croplands dataset, as used in this study, is widespread in academic studies, it does not capture the fine-scale variability at the parcel level. Replacing it with parcel-level data in the future can be a direction for further work.

The additional benefit of a standardized validation dataset at a continental scale would be very relevant, since management practices present a high variability across the EU (Panagos et al., 2015). Due to the current lack of information, it was necessary to assume in this work that if the model presents reasonable results for France, then it is appropriate for other regions. France is a particularly good choice of MS due to data availability, the current status of CC implementation, and its high diversity of climate regions. However, the fact that the performance of the results in other regions is still unknown must be improved in further work. Therefore, we highlight the need for increased data availability and an accompanying definition of what does or does not belong to the CCs group among EU countries. In this sense, it is necessary to reinforce that the results obtained in the present work are conditional on the broad definition presented in Section 2.2. As an outcome of this work, we justify the need for a move towards a consistent terminology for CCs in the EU in order to facilitate the synergistic use of datasets and improve future prediction exercises.

4. Conclusions

In this work, a statistical disaggregation model is proposed to derive CC information at 100-m resolution from aggregated statistics reported in the Farm Field Survey at the NUTS2 level in Europe (i.e., 215 regions with CC information). The transference from the coarse (regional aggregations) to the fine (pixel) level was made through a statistical model constructed based on assumptions relating CC phenology in arable lands to a full annual time series of the cross ratio (CR) from Sentinel-1. To quantify the model accuracy, the best available data sampled from spatially-explicit farmers' declarations in the French RPG was used to validate the model at the field parcel scale. The models' interior behavior was shown to be coherent with reality, modifying its sensitivity to the Sentinel-1 CR time series according to the region and period of the year under consideration. In regions with known CC implementation, the model properly identified the importance of considering a complete time series to generate context-dependent model predictions. This multi-temporal approach was important in the model's successful distinction of different types of winter cover, in which lower fractions of CCs were predicted in places with winter commercial crops. In general, the results showed that the model was able to successfully predict higher CC fractions in areas where they are planted in fields. Overall, when interpreted as a binary classifier for the whole of France, the model yielded an Area Under the Curve (AUC) of 0.74. On a regional basis, the AUC values were 0.77, 0.75, 0.74, 0.70, and 0.65 for the North, Center, West, East, and South regions, respectively, showing geographical variation in the model accuracy. Despite discussed limitations, this derived data layer can provide an important and updateable information source for researchers and practitioners requiring a spatially explicit knowledge of CC implementation.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.162300, and the datasets at 100-m spatial resolution are available in the European Soil Data Centre 2.0 (ESDAC) (JRC, 2023; Panagos et al., 2022).

CRediT authorship contribution statement

Arthur Nicolaus Fendrich: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft. **Francis Matthews:** Conceptualization, Methodology, Investigation, Data curation, Writing – original draft. **Elise Van Eynde:** Conceptualization, Methodology, Investigation, Data curation, Writing – original draft. **Marco Carozzi:** Conceptualization, Investigation, Resources, Data curation, Writing – original draft. **Zheyuan Li:** Methodology, Software, Formal analysis, Investigation, Writing – original draft. **Raphael d'Andrimont:** Conceptualization, Methodology, Investigation, Writing – original draft. **Emanuele Lugato:** Conceptualization, Investigation, Writing – original draft. **Philippe Martin:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – original draft, Supervision. **Philippe Ciais:** Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Supervision. **Panos Panagos:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – original draft, Supervision.

Data availability

The datasets of predicted CC fraction, standard deviation, and the 5 % and 95 % quantiles at 100-m spatial resolution are available in the European Soil Data Centre 2.0 (ESDAC) (JRC, 2023; Panagos et al., 2022).

Declaration of competing interest

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

Acknowledgments

This work was realized with the collaboration of the European Commission Joint Research Centre under the Collaborative Doctoral Partnership Agreements No. 35403 and 35332, and partially funded by the project.

References

Amundson, R., Berhe, A.A., Hopmans, J.W., Olson, C., Sztein, A.E., Sparks, D.L., 2015. Soil and human security in the 21st century. Science 348. https://doi.org/10.1126/science.1261071 URL: 10.1126/science.1261071.

Bates, D.M., Watts, D.G. (Eds.), 1988. Nonlinear Regression Analysis And Its Applications. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470316757 URL: doi:10.1002/9780470316757.

Bera, B., Saha, S., Bhattacharjee, S., 2020. Forest cover dynamics (1998 to 2019) and prediction of deforestation probability using binary logistic regression (BLR) model of Silabati watershed, India. TreesFor.People 2, 100034. https://doi.org/10.1016/j.tfp.2020.100034 URL: 10.1016/j.tfp.2020.100034.

Beriaux, E., Jago, A., Lucau-Danila, C., Planchon, V., Defourny, P., 2021. Sentinel-1 time series for crop identification in the framework of the future CAP monitoring. Remote Sens. 13, 2785. https://doi.org/10.3390/rs13142785 URL: 10.3390/rs13142785.

Borrelli, P., Panagos, P., 2020a. 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 URL: 10.1016/j.landusepol.2020.104467.

Borrelli, P., Panagos, P., 2020b. 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 URL: 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 URL: 10.1016/j.landusepol.2015.09.033.

Borrelli, P., Robinson, D.A., Fleischer, L.R., Lugato, E., Ballabio, C., Alewell, C., Meusburger, K., Modugno, S., Schütt, B., Ferro, V., Bagarello, V., Oost, K.V., Montanarella, L., Panagos, P., 2017. An assessment of the global impact of 21st century land use change on soil erosion. Nat. Commun. 8. https://doi.org/10.1038/s41467-017-02142-7 URL: doi:10.1038/s41467-017-02142-7.

Borrelli, P., Robinson, D.A., Panagos, P., Lugato, E., Yang, J.E., Alewell, C., Wuepper, D., Montanarella, L., Ballabio, C., 2020. Land use and climate change impacts on global soil erosion by water (2015-2070). Proc. Natl. Acad. Sci. 117, 21994–22001. https://doi.org/10.1073/pnas.2001403117 URL: doi:10.1073/pnas.2001403117.

Borrelli, P., Ballabio, C., Yang, J.E., Robinson, D.A., Panagos, P., 2022. GloSEM: high-resolution global estimates of present and future soil displacement in croplands by water erosion. Sci.Data 9. https://doi.org/10.1038/s41597-022-01489-x URL: 10.1038/s41597-022-01489-x.

Büttner, G., Kosztra, B., 2022. Clc2018 technical guidelines. https://land.copernicus.eu/user-corner/technical-library/clc2018technicalguidelines_final.pdf Accessed: 2022-10-06.

Comber, A., Zeng, W., 2019. Spatial interpolation using areal features: a review of methods and opportunities using new forms of data with coded illustrations. <sb:contribution><sb:title>Geogr. </sb:title></sb:contribution><sb:host><sb:issue><sb:series><sb:title>Compass</sb:title></sb:series></sb:issue></sb:host> 13. https://doi.org/10.1111/gec3.12465 URL: 10.1111/gec3.12465.

Copernicus, 2022. Clc 2018. https://land.copernicus.eu/pan-european/corine-land-cover/clc2018 Accessed: 2022-09-20.

d'Andrimont, R., Verheghen, A., Lemoine, G., Kempeneers, P., Meroni, M., van der Velde, M., 2021. From parcel to continental scale – a first European crop type map based on sentinel-1 and LUCAS Copernicus in-situ observations. Remote Sens. Environ. 266, 112708. https://doi.org/10.1016/j.rse.2021.112708 URL: 10.1016/j.rse.2021.112708.

Delgado, J.A., Reeves, D.W., Follett, R.F., 2017. Cover crops. Encyclopedia of Soil Science, Third edition CRC Press, pp. 484–487 https://doi.org/10.1081/e-ess3-120054071 URL: 10.1081/e-ess3-120054071.

Eddelbuettel, D., Balamuta, J.J., 2018. Extending R with C++: a brief introduction to Rcpp. Am. Stat. 72, 28–36. https://doi.org/10.1080/00031305.2017.1375990.

Eddelbuettel, D., Sanderson, C., 2014. Rcpparmadillo: accelerating R with high-performance C++ linear algebra. Comput. Stat. Data Anal. 71, 1054–1063. https://doi.org/10.1016/j.csda.2013.02.005.

European Commission, 2022a. Eurostat: farm structure (ef). https://ec.europa.eu/eurostat/cache/metadata/EN/ef_esqrs_de.htm Accessed: 2022-09-22.

European Commission, 2022b. Eurostat: Lucas - land use and land cover survey. https://web.archive.org/web/20220901034430/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=LUCAS_-_Land_use_and_land_cover_survey Accessed: 2022-10-10.

European Commission, 2022c. Eurostat: soil cover by nuts2 regions. https://ec.europa.eu/eurostat/databrowser/view/ef_mp_soil/default/table?lang=en Accessed: 2022-09-20.

European Commission, 2022. Glossary: Farm structure survey (fss). https://web.archive.org/web/20220901054908/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Farm_structure_survey_(FSS) Accessed: 2022-10-06.

European Commission, 2022e. Glossary: soil cover. https://web.archive.org/web/20220119025400/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Soil_cover Accessed: 2022-09-19.

European Commission, 2022. Glossary: Survey on agricultural production methods (sapm). https://web.archive.org/web/20220901012231/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Survey_on_agricultural_production_methods_(SAPM) Accessed: 2022-10-10.

European Commission, 2022g. Proposed cap strategic plans and commission observations: summary overview for 27 member states. https://agriculture.ec.europa.eu/system/files/2022-07/csp-overview-28-plans-overview-june-2022_en.pdf Accessed: 2022-09-19.
Commission, 2018. Commission staff working document swd(2018) 301 final. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018SC0301.
European Commission, 2021. List of potential agricultural practices that eco-schemes could support. https://agriculture.ec.europa.eu/system/files/2021-01/factsheet-agri-practices-under-ecoscheme_en_0.pdf.
FAO, 2015. Status of the World's Soil Resources: Main Report. FAO: ITPS OCLC: 945442780.
Fendrich, A.N., Neto, E.S.H., Moreira, L.E.M., Neto, D.D., 2022. A scalable method for the estimation of spatial disaggregation models. Comput.Geosci. 166, 105161. https://doi.org/10.1016/j.cageo.2022.105161 URL: 10.1016/j.cageo.2022.105161.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031 URL: 10.1016/j.rse.2017.06.031.
Hijmans, R.J., 2022. terra: Spatial Data Analysis. URL https://CRAN.R-project.org/package=terra r package version 1.6-17.
Hobbs, P.R., Sayre, K., Gupta, R., 2007. The role of conservation agriculture in sustainable agriculture. Philos. Trans. R. Soc., B 363, 543–555. https://doi.org/10.1098/rstb.2007.2169 URL: doi:10.1098/rstb.2007.2169.
Institut National de l'Information Géographique et Forestière, 2022. Registre parcellaire graphique. https://geoservices.ign.fr/rpg Accessed: 2022-10-06.
Jenny, J.P., Koirala, S., Gregory-Eaves, I., Francus, P., Niemann, C., Ahrens, B., Brovkin, V., Baud, A., Ojala, A.E.K., Normandeau, A., Zolitschka, B., Carvalhais, N., 2019. Human and climate global-scale imprint on sediment transfer during the holocene. Proc. Natl. Acad. Sci. 116, 22972–22976. https://doi.org/10.1073/pnas.1908179116 URL: 10.1073/pnas.1908179116.
Jia, P., Gaughan, A.E., 2016. Dasymetric modeling: a hybrid approach using land cover and tax parcel data for mapping population in Alachua County, Florida. Appl. Geogr. 66, 100–108. https://doi.org/10.1016/j.apgeog.2015.11.006 URL: 10.1016/j.apgeog.2015.11.006.
Journal Officiel de la République Française, 2018. Arrêté du 9 avril 2018 modifiant l'arrêté du 12 novembre 2015 fixant certaines dispositions relatives au paiement pour les pratiques agricoles bénéfiques pour le climat et l'environnement dit « paiement vert » prévu par la politique agricole commune. https://www.legifrance.gouv.fr/download/pdf?id=-TK_H2XcPnMCxk14S-Q3vZtam0AAwx52TSzzQW0k1-4= Accessed: 2022-10-04.
JRC, 2022. Nitrated directive - vulnerable zones reporting 7. https://water.jrc.ec.europa.eu/portal/apps/webappviewer/index.html?id=b33a220c1b284583851e93a245da02ef Accessed: 2022-10-21.
JRC, 2023. European soil data centre. https://esdac.jrc.ec.europa.eu/ Accessed: 2023-01-04.
Kathage, J., Domínguez, I.Pérez, 2019. Adoption of cover crops for climate mitigation in the Eu. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC116730/ccc_report.pdf Accessed: 2022-10-13.
Kathage, J., Smit, B., Janssens, B., Haagsma, W., Adrados, J.L., 2022a. How much is policy driving the adoption of cover crops? Evidence from four EU regions. Land Use Policy 116, 106016. https://doi.org/10.1016/j.landusepol.2022.106016 URL: 10.1016/j.landusepol.2022.106016.
Kathage, J., Smit, B., Janssens, B., Haagsma, W., Adrados, J.L., 2022b. How much is policy driving the adoption of cover crops? Evidence from four EU regions. Land Use Policy 116, 106016. https://doi.org/10.1016/j.landusepol.2022.106016 URL: 10.1016/j.landusepol.2022.106016.
Kim, N., Zabaloy, M.C., Guan, K., Villamil, M.B., 2020. Do cover crops benefit soil microbiome? A meta-analysis of current research. Soil Biol. Biochem. 142, 107701. https://doi.org/10.1016/j.soilbio.2019.107701 URL: 10.1016/j.soilbio.2019.107701.
Koudahe, K., Allen, S.C., Djaman, K., 2022. Critical review of the impact of cover crops on soil properties. Int. Soil Water Conserv. Res. 10, 343–354. https://doi.org/10.1016/j.iswcr.2022.03.003 URL: 10.1016/j.iswcr.2022.03.003.
Launay, C., Houot, S., Frédéric, S., Girault, R., Levavasseur, F., Marsac, S., Constantin, J., 2022. Incorporating energy cover crops for biogas production into agricultural systems: benefits and environmental impacts. A review. Agron. Sustain. Dev. 42. https://doi.org/10.1007/s13593-022-00790-8 URL: 10.1007/s13593-022-00790-8.
Levavasseur, F., Martin, P., Bouty, C., Barbottin, A., Bretagnolle, V., Thérond, O., Scheurer, O., Piskiewicz, N., 2016. RPG explorer: a new tool to ease the analysis of agricultural landscape dynamics with the land parcel identification system. Comput. Electron. Agric. 127, 541–552. https://doi.org/10.1016/j.compag.2016.07.015 URL: 10.1016/j.compag.2016.07.015.
Ma, C., Johansen, K., McCabe, M.F., 2022. Monitoring irrigation events and crop dynamics using sentinel-1 and sentinel-2 time series. Remote Sens. 14, 1205. https://doi.org/10.3390/rs14051205 URL: 10.3390/rs14051205.
Maimon, O., Rokach, L. (Eds.), 2010. Data Mining And Knowledge Discovery Handbook. Springer US https://doi.org/10.1007/978-0-387-09823-4 Accessed: 2022-10-04.
Matthews, F., Verstraeten, G., Borrelli, P., Panagos, P., 2023. A field parcel-oriented approach to evaluate the crop cover-management factor and time-distributed erosion risk in Europe. Int. Soil Water Conserv. Res. https://doi.org/10.1016/j.iswcr.2022.09.005 URL: doi:10.1016/j.iswcr.2022.09.005.
Maurya, A.K., Bhargava, N., Singh, D., 2022. Efficient selection of SAR features using ML based algorithms for accurate FVC estimation. Adv. Space Res. 70, 1795–1809. https://doi.org/10.1016/j.asr.2022.06.039 URL: doi:10.1016/j.asr.2022.06.039.
McClelland, S.C., Paustian, K., Schipanski, M.E., 2021. Management of cover crops in temperate climates influences soil organic carbon stocks: a meta-analysis. Ecol. Applic. 31. https://doi.org/10.1002/eap.2278 URL: doi:10.1002/eap.2278.
McDaniel, M.D., Tiemann, L.K., Grandy, A.S., 2014. Does agricultural crop diversity enhance soil microbial biomass and organic matter dynamics? A meta-analysis. Ecol. Appl. 24, 560–570. https://doi.org/10.1890/13-0616.1 URL: 10.1890/13-0616.1.
Meroni, M., d'Andrimont, R., Vrieling, A., Fasbender, D., Lemoine, G., Rembold, F., Seguini, L., Verhegghen, A., 2021. Comparing land surface phenology of major european crops as derived from SAR and multispectral data of sentinel-1 and -2. Remote Sens. Environ. 253, 112232. https://doi.org/10.1016/j.rse.2020.112232 URL: doi:10.1016/j.rse.2020.112232.
Cultures et précisions. liste des cultures à utiliser pour renseigner le descriptif des parcelles. https://www.telepac.agriculture.gouv.fr/telepac/pdf/tas/2020/Dossier-PAC-2020_notice_cultures-precisions.pdf.
Møller, A.B., Malone, B., Odgers, N.P., Beucher, A., Iversen, B.V., Greve, M.H., Minasny, B., 2019. Improved disaggregation of conventional soil maps. Geoderma 341, 148–160. https://doi.org/10.1016/j.geoderma.2019.01.038 URL: 10.1016/j.geoderma.2019.01.038.
Montgomery, D.R., 2007. Soil erosion and agricultural sustainability. Proc. Natl. Acad. Sci. 104, 13268–13272. https://doi.org/10.1073/pnas.0611508104 URL: doi:10.1073/pnas.0611508104.
Muhammad, I., Wang, J., Sainju, U.M., Zhang, S., Zhao, F., Khan, A., 2021. Cover cropping enhances soil microbial biomass and affects microbial community structure: a meta-analysis. Geoderma 381, 114696. https://doi.org/10.1016/j.geoderma.2020.114696 URL: 10.1016/j.geoderma.2020.114696.
Nisbet, R., Elder, J.F., Miner, G., 2009. Handbook of Statistical Analysis And Data Mining Applications. Academic Press/Elsevier, Amsterdam; Boston OCLC: ocn316327105.
Nowak, B., Marliac, G., Michaud, A., 2021. Estimation of winter soil cover by vegetation before spring-sown crops for mainland France using multispectral satellite imagery. Environ. Res. Lett. 16, 064024. https://doi.org/10.1088/1748-9326/ac007c URL: 10.1088/1748-9326/ac007c.
Nyakatawa, E., Reddy, K., Lemunyon, J., 2001. Predicting soil erosion in conservation tillage cotton production systems using the revised universal soil loss equation (RUSLE). Soil Tillage Res. 57, 213–224. https://doi.org/10.1016/s0167-1987(00)00178-1 URL: 10.1016/s0167-1987(00)00178-1.
Official Journal of the European Union, 1991. Council directive 91/676/eec of 12 december 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31991L0676 from=EN. Accessed: 2022-09-23.
Official Journal of the European Union, 2021. Regulation (eu) 2021/2116 of the European parliament and of the council of 2 December 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2116 OCLC: ocn316327105.
Ols, C., Hervé, J.C., Bontemps, J.D., 2020. Recent growth trends of conifers across western Europe are controlled by thermal and water constraints and favored by forest heterogeneity. Sci. Total Environ. 742, 140453. https://doi.org/10.1016/j.scitotenv.2020.140453 URL: 10.1016/j.scitotenv.2020.140453.
Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde, O., 2017. LUCAS soil, the largest expandable soil dataset for Europe: a review. Eur. J. Soil Sci. 69, 140–153. https://doi.org/10.1111/ejss.12499 URL: 10.1111/ejss.12499.
Palm, C., Blanco-Canqui, H., DeClerck, F., Gatere, L., Grace, P., 2014. Conservation agriculture and ecosystem services: an overview. Agric. Ecosyst. Environ. 187, 87–105. https://doi.org/10.1016/j.agee.2013.10.010 URL: 10.1016/j.agee.2013.10.010.
Panagos, P., Borrelli, P., Meusburger, K., Alewell, C., Lugato, E., Montanarella, L., 2015. Estimating the soil erosion cover-management factor at the European scale. Land Use Policy 48, 38–50. https://doi.org/10.1016/j.landusepol.2015.05.021 URL: 10.1016/j.landusepol.2015.05.021.
Panagos, P., Ballabio, C., Himics, M., Scarpa, S., Matthews, F., Bogonos, M., Poesen, J., Borrelli, P., 2021. Projections of soil loss by water erosion in Europe by 2050. Environ. Sci. Pol. 124, 380–392. https://doi.org/10.1016/j.envsci.2021.07.012 URL: 10.1016/j.envsci.2021.07.012.
Panagos, P., Liedekerke, M.V., Borrelli, P., Köninger, J., Ballabio, C., Orgiazzi, A., Lugato, E., Liakos, L., Hervas, J., Jones, A., Montanarella, L., 2022. European Soil Data Centre 2.0: soil data and knowledge in support of the Eu policies. Eur. J. Soil Sci. 73. https://doi.org/10.1111/ejss.13315 URL: 10.1111/ejss.13315.
Panahi, M., Khosravi, K., Golkarian, A., Roostaei, M., Barzegar, R., Omidvar, E., Rezaie, F., Saco, P.M., Sharifi, A., Jun, C., Bateni, S.M., Lee, C.W., Lee, S., 2022. A country-wide assessment of Irans land subsidence susceptibility using satellite-based InSAR and machine learning. Geocarto Int., 1–23 https://doi.org/10.1080/10106049.2022.2086631 URL: 10.1080/10106049.2022.2086631.
Poesen, J., 2017. Soil erosion in the Anthropocene: research needs. Earth Surf. Process. Landf. 43, 64–84. https://doi.org/10.1002/esp.4250 URL: 10.1002/esp.4250.
R Core Team, 2022. R: A Language And Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria URL: https://www.R-project.org/.
Rossell, D., Abril, O., Bhattacharya, A., 2021. Approximate Laplace approximations for scalable model selection. J. R. Stat. Soc. Ser. B Stat Methodol. 83, 853–879.
Roy, P., Chakrabortty, R., Chowdhuri, I., Malik, S., Das, B., Pal, S.C., 2020. Development of different machine learning ensemble classifier for gully erosion susceptibility in Gandheswari watershed of West Bengal, India. Machine Learning for Intelligent Decision Science. Springer, Singapore, pp. 1–26 https://doi.org/10.1007/978-981-15-3689-2_1 URL: doi:10.1007/978-981-15-3689-2_1.
Ruis, S.J., Blanco-Canqui, H., 2017. Cover crops could offset crop residue removal effects on soil carbon and other properties: a review. Agron. J. 109, 1785–1805. https://doi.org/10.2134/agronj2016.12.0735 URL: 10.2134/agronj2016.12.0735.
Schneider, M., Broszeit, A., Körner, M., 2021. Eurocrops: A Pan-European Dataset for Time Series Crop Type Classification. https://doi.org/10.2760/125905 arXiv:arXiv:2106.08151.
Smith, M.S., Frye, W.W., Varco, J.J., 1987. Legume winter cover crops. Advances in Soil Science. Springer, New York, pp. 95–139 https://doi.org/10.1007/978-1-4612-4790-6_3 URL: 10.1007/978-1-4612-4790-6_3.
Tierney, L., Rossini, A.J., Li, N., Sevcikova, H., 2021. snow: Simple Network of Workstations. https://CRAN.R-project.org/package=snow Accessed: 2022-10-04.
Tzilivakis, J., Warner, D.J., Green, A., Lewis, K.A., 2021. A broad-scale spatial analysis of the environmental benefits of fertiliser closed periods implemented under the Nitrates

Directive in Europe. J. Environ. Manage. 299, 113674. https://doi.org/10.1016/j.jenvman.2021.113674.
Utazi, C., Thorley, J., Alegana, V., Ferrari, M., Nilsen, K., Takahashi, S., Metcalf, C., Lessler, J., Tatem, A., 2018. A spatial regression model for the disaggregation of areal unit based data to high-resolution grids with application to vaccination coverage mapping. Stat. Methods Med. Res. 28, 3226–3241. https://doi.org/10.1177/0962280218797362 URL: 10.1177/0962280218797362.
Vavlas, N.C., Waine, T.W., Meersmans, J., Burgess, P.J., Fontanelli, G., Richter, G.M., 2020. Deriving wheat crop productivity indicators using sentinel-1 time series. Remote Sens. 12, 2385. https://doi.org/10.3390/rs12152385 URL: 10.3390/rs12152385.
Veloso, A., Mermoz, S., Bouvet, A., Toan, T.L., Planells, M., Dejoux, J.F., Ceschia, E., 2017. Understanding the temporal behavior of crops using sentinel-1 and sentinel-2-like data for agricultural applications. Remote Sens. Environ. 199, 415–426. https://doi.org/10.1016/j.rse.2017.07.015 URL: 10.1016/j.rse.2017.07.015.
Verstraeten, G., Oost, K., Rompaey, A., Poesen, J., Govers, G., 2006. Evaluating an integrated approach to catchment management to reduce soil loss and sediment pollution through modelling. Soil Use Manag. 18, 386–394. https://doi.org/10.1111/j.1475-2743.2002.tb00257.x URL: 10.1111/j.1475-2743.2002.tb00257.x.
Vreugdenhil, M., Navacchi, C., Bauer-Marschallinger, B., Hahn, S., Steele-Dunne, S., Pfeil, I., Dorigo, W., Wagner, W., 2020. Sentinel-1 cross ratio and vegetation optical depth: a comparison over Europe. Remote Sens. 12, 3404. https://doi.org/10.3390/rs12203404 URL: 10.3390/rs12203404.
Weiss, D.J., Lucas, T.C.D., Nguyen, M., Nandi, A.K., Bisanzio, D., Battle, K.E., Cameron, E., Twohig, K.A., Pfeffer, D.A., Rozier, J.A., Gibson, H.S., Rao, P.C., Casey, D., Bertozzi-Villa, A., Collins, E.L., Dalrymple, U., Gray, N., Harris, J.R., Howes, R.E., Kang, S.Y., Keddie, S.H., May, D., Rumisha, S., Thorn, M.P., Barber, R., Fullman, N., Huynh, C.K., Kulikoff, X., Kutz, M.J., Lopez, A.D., Mokdad, A.H., Naghavi, M., Nguyen, G., Shackelford, K.A., Vos, T., Wang, H., Smith, D.L., Lim, S.S., Murray, C.J.L., Bhatt, S., Hay, S.I., Gething, P.W., 2019. Mapping the global prevalence, incidence, and mortality of plasmodium falciparum, 2000–17: a spatial and temporal modelling study. Lancet 394, 322–331. https://doi.org/10.1016/s0140-6736(19)31097-9 URL: doi:10.1016/j.rse.2020.112232.
Wolfinger, R., 1993. Laplace's approximation for nonlinear mixed models. Biometrika 80, 791–795. https://doi.org/10.1093/biomet/80.4.791 URL: 10.1093/biomet/80.4.791.
Wood, S.N., 2016. P-splines with derivative based penalties and tensor product smoothing of unevenly distributed data. Stat. Comput. 27, 985–989. https://doi.org/10.1007/s11222-016-9666-x URL: 10.1007/s11222-016-9666-x.
Wood, S., 2017. Generalized Additive Models: An Introduction With R. 2nd ed. Chapman and Hall/CRC.