



# Time series of Landsat-based bimonthly and annual spectral indices for continental Europe for 2000–2022

Xuemeng Tian<sup>1,2</sup>, Davide Consoli<sup>1</sup>, Martijn Witjes<sup>1</sup>, Florian Schneider<sup>3</sup>, Leandro Parente<sup>1</sup>,  
Murat Şahin<sup>1</sup>, Yu-Feng Ho<sup>1</sup>, Robert Minařík<sup>1</sup>, and Tomislav Hengl<sup>1</sup>

<sup>1</sup>OpenGeoHub, Doorwerth, the Netherlands

<sup>2</sup>Laboratory of Geo-Information Science and Remote Sensing,  
Wageningen University & Research, Wageningen, the Netherlands

<sup>3</sup>Thünen Institute of Climate-Smart Agriculture, Braunschweig, Germany

**Correspondence:** Xuemeng Tian (xuemeng.tian@opengeohub.org)

Received: 28 June 2024 – Discussion started: 12 September 2024

Revised: 28 November 2024 – Accepted: 29 December 2024 – Published: 26 February 2025

**Abstract.** The production and evaluation of the analysis-ready and cloud-optimized (ARCO) data cube for continental Europe (including Ukraine, the UK, and Türkiye), derived from the Landsat analysis-ready dataset version 2 (ARD V2) produced by Global Land Analysis and Discovery (GLAD) team and covering the period from 2000 to 2022, is described. The data cube consists of 17 TB of data at a 30 m resolution and includes bi-monthly, annual, and long-term spectral indices on various thematic topics, including surface reflectance bands, normalized difference vegetation index (NDVI), soil adjusted vegetation index (SAVI), fraction of absorbed photosynthetically active radiation (FAPAR), normalized difference snow index (NDSI), normalized difference water index (NDWI), normalized difference tillage index (NDTI), minimum normalized difference tillage index (minNDTI), bare soil fraction (BSF), number of seasons (NOS), and crop duration ratio (CDR). The data cube was developed with the intention to provide a comprehensive feature space for environmental modeling and mapping. The quality of the produced time series was assessed by (1) assessing the accuracy of gap-filled bimonthly Landsat data with artificially created gaps; (2) visual examination for artifacts and inconsistencies; (3) plausibility checks with ground survey data; and (4) predictive modeling tests, examples with soil organic carbon (SOC) and land cover (LC) classification. The time series reconstruction demonstrates high accuracy, with a root mean squared error (RMSE) smaller than 0.05, and  $R^2$  higher than 0.6, across all bands. The visual examination indicates that the product is complete and consistent, except for winter periods in northern latitudes and high-altitude areas, where high cloud and snow density introduce significant gaps and hence many artifacts remain. The plausibility check further shows that the indices logically and statistically capture the processes. The BSF index showed a strong negative correlation ( $-0.73$ ) with crop coverage data, while the minNDTI index had a moderate positive correlation ( $0.57$ ) with the Eurostat tillage practice survey data. The detailed temporal resolution and long-term characteristics provided by different tiers of predictors in this data cube proved to be important for both soil organic carbon regression and LC classification experiments based on 60 723 LUCAS observations: long-term characteristics (tier 4) were particularly valuable for predictive mapping of SOC and LC, coming out on top of variable importance assessment. Crop-specific indices (NOS and CDR) provided limited value for the tested applications, possibly due to noise or insufficient quantification methods. The data cube is made available at <https://doi.org/10.5281/zenodo.10776891> (Tian et al., 2024) under a CC-BY license and will be continuously updated.

## 1 Introduction

Earth observation (EO) data are increasingly recognized for its critical role in environmental sciences (Kansakar and Hosain, 2016; Salcedo-Sanz et al., 2020). Compared to traditional field surveys, they offer high-resolution, large-scale, and recurrent spatial environmental data at relatively low cost. These capabilities are important for comprehensive environmental studies, ongoing monitoring, and enables effective management (Chatenoux et al., 2021; Giuliani et al., 2021). EO data not only offer valuable insights into the Earth's surface through various spectral bands but also can highlight specific land surface features through spectral indices derived by processing them. In addition, satellite data serve as an essential input for various models that study physical processes, forecast future scenarios, and inform policy- and decision-makers (Salcedo-Sanz et al., 2020).

As EO technology develops, more opportunities emerge. The development of a wider range of spectral indices allows for a more detailed analysis using different combinations of satellite bands (Montero et al., 2023). Improvements in spatial and temporal resolutions enable more detailed observations. In addition, a longer record of satellite imagery facilitates long-term environmental studies. These advances help us better understand the environmental dynamics and enhance natural resource management strategies. However, they also introduce new challenges. One of the key challenges to fully unlock the potential of EO data for environmental applications, as highlighted by numerous researchers (Killough, 2018; Wagemann et al., 2021; Giuliani et al., 2021; Chatenoux et al., 2021; Montero et al., 2023), is the gap between the demand for detailed EO data and the limited processing capabilities of most users. As the volume of EO data and the complexity of spectral indices increase, processing typically requires specialized expertise and costly computational resources both locally and in the cloud. For example, while platforms such as MODIS, Landsat, and Sentinel openly provide valuable satellite data, these data often require preprocessing to remove poor-quality pixels, such as those affected by cloud cover and geometric distortions (Wulder et al., 2022; Radeloff et al., 2024). While many preprocessed layers are available, they tend to be scattered across various data portals and often focus on limited themes. Accessing these ready-to-use datasets usually demands domain-specific knowledge to even be aware of their existence. In addition, non-standardized data formats further complicate the ability of users to integrate these datasets for specific applications (Lokers et al., 2016; Wagemann et al., 2021). Thus, there is an essential need for solutions that ensure easy access to open environmental data. Furthermore, these solutions should also facilitate the easy integration of these data, enabling its combined use to effectively address critical environmental and economic challenges (Giuliani et al., 2017).

Numerous platforms have emerged in the last decade that aim to make the management, processing, and analysis of big EO data operational, including Google Earth Engine (GEE, Gorelick et al., 2017), Sentinel Hub (SH, <https://www.sentinel-hub.com/>, last access: 27 June 2024), Open Data Cube (ODC, <https://www.opendatacube.org/>, last access: 27 June 2024), Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu/>, last access: 27 June 2024), and openEO cloud (<https://openeo.cloud/>, last access: 27 June 2024). Each platform has its strengths and weaknesses: GEE offers extensive datasets and robust processing but is limited by its closed nature, restricting external contributions and governance. Sentinel Hub faces similar limitations. openEO provides flexibility with user-defined functions but lacks guaranteed compatibility across different backend systems. Among these, the Open Data Cube (ODC) stands out for its open-source nature, flexibility, and strong community support, making it a leading option (Gomes et al., 2020). As a result, many analysis-ready data (ARD) cubes have been developed using ODC principles. ARD refers to a multidimensional time series stack of spatially aligned pixels that are ready for analysis, eliminating the need for additional data harmonization (Giuliani et al., 2017, 2021). These efforts reduce technical complexity and have been shown to be effective in delivering information efficiently, as demonstrated by successful implementations in various domains, such as the vegetation dynamics (Obuchowicz et al., 2023), snow cover (Poussin et al., 2023), and drying trend (Poussin et al., 2021).

EcoDataCube.eu, developed by Witjes et al. (2023), is the first product to cover the entire EU with a sufficient temporal range to support a long-term analysis of land degradation and land use change dynamics. This work extends the concept of ARD to ARCO (analysis-ready and cloud-optimized) data cubes (Stern et al., 2022; Iosifescu Enescu et al., 2021), which helps to optimize cloud-based data management and processing. Cloud optimization allows for on-the-fly access, reduces latency through partial and parallel reads, and efficient metadata handling. Cloud-Optimized GeoTIFF (COG, <https://www.cogeo.org/>, last access: 27 June 2024), adopted by EcoDataCube.eu, is a good example of such a format. COG files are structured to facilitate network access through hypertext transfer protocol (HTTP) range requests, ensuring compatibility with cloud object storage systems. This design supports integration with high-level analysis libraries and distributed frameworks. Additionally, EcoDataCube.eu also minimizes data gaps caused by clouds, which is a major obstacle to unlocking the potential of EO data (Baumann, 2024). This is achieved through time-reconstruction algorithms to ensure optimal cloud-free conditions (Zhou et al., 2016; Consoli et al., 2024). This approach not only prepares the data for spatio-temporal analysis but also proves beneficial for modeling as models typically require complete input data to function effectively.

EcoDataCube.eu provides valuable insights through its quarterly temporal resolution and raw bands, making it a significant resource for Earth observation applications. To further reduce the burden on end users and increase the use and impact of EO data, we developed the ARCO data cube for the EU by refining the temporal resolution, processing reflectance bands to derive spectral indices, and updating the data layers until 2022. Our Landsat-based spectral index data cube spans from 2000 to 2022, covering the pan-European area, including Ukraine, the UK, and Türkiye, with data at a detailed 30 m resolution. Using the Global Land Analysis and Discovery (GLAD) team Landsat analysis-ready dataset version 2 (ARD V2) presented by Potapov et al. (2020) as input, we adopt the state-of-the-art approach of Consoli et al. (2024) to generate four tiers of environmental predictors:

1. bimonthly aggregated cloud-optimized bands,
2. bimonthly spectral indices derived from bands,
3. annual indices derived by analyzing the bimonthly time series of indices,
4. long-term indices reflecting features across 2 decades.

The foundational tier of the data cube consists of bimonthly cloud-optimized bands, derived using time reconstruction methods to address the gaps left by cloudy pixels (Consoli et al., 2024). Building on this, we calculate bimonthly spectral indices as the second tier of predictors. This selection includes the most widely used indices, covering key aspects such as vegetation, crops, soil, and water (Montero et al., 2023): the normalized difference vegetation index (NDVI) is used to evaluate vegetation health and biomass. Complementing NDVI, the soil adjusted vegetation index (SAVI) has been shown to be more effective in areas with sparse vegetation by minimizing the impact of soil brightness on vegetation sensing (Huete, 1988; Rhyma et al., 2020; Reddy et al., 2022). Additionally, the fraction of absorbed photosynthetically active radiation (FAPAR) provides a more direct measurement of plant productivity (Myeni and Williams, 1994). The normalized difference water index (NDWI), also known as the normalized difference moisture index (NDMI), provides insights into water dynamics and climatic characteristics (Gao, 1996). The normalized difference snow index (NDSI) helps identify snowy areas (Salomonson and Appel, 2006). The normalized difference tillage index (NDTI), also known as normalized burn ratio 2 (NBR2), shows potential in tillage detection, post-fire recovery studies, and soil sealing identification (Zheng et al., 2012; Daughtry et al., 2010; Eskandari et al., 2016; Beeson et al., 2020; Sonmez and Slater, 2016; Storey et al., 2016; Etehad Osgouei et al., 2019; Xiang et al., 2022).

Through time series analysis, we also derive several annual indices that capture temporal patterns within the year as our third tier of predictors. This includes the number of seasons (NOS) and crop duration ratio (CDR) from NDVI

time series, which shed light on the intensity of cropland use (Siebert et al., 2010; Li et al., 2014; Patel and Oza, 2014; Estel et al., 2016). The bare soil fraction (BSF) measures the duration of soil exposure, which is related to soil health (Demattê et al., 2020; Mzid et al., 2021; Sharma et al., 2018; Turmel et al., 2015). Although NDTI is positively correlated with the crop residue cover ratio (CRC), using it to distinguish between conventional, conservative, and zero-tillage practices can be challenging without knowing the specific timing of tillage or planting (Zheng et al., 2012, 2013). Therefore, its annual derivation, minimum NDTI (minNDTI), is adopted as a proxy for tillage due to its ease of application (Zheng et al., 2012). We also calculate annual percentile aggregations (25th, 50th, and 75th) of NDVI and NDWI to quantify annual vegetation and water conditions. Finally, from these annual time series, we develop long-term indices that reveal decadal characteristics as the tier-4 predictors.

Given the comprehensive and extensive nature of this data cube, it is impractical to validate every derived spectral index against ground-truth data. Plausibility checks were conducted selectively, focusing on third-tier predictors where feasible and necessary and where relevant ground data were available. These checks involved matching the data with the ground survey statistics to highlight the advantages and limitations of the data. Our predictor data cube will be compared with EcoDataCube.eu, which has been utilized for land cover (LC) classification by Witjes et al. (2023). By benchmarking against EcoDataCube.eu, our aim is to highlight the improvements and advantages offered by our spectral index data cube, specifically in terms of accuracy, robustness, and usability for land cover classification. To demonstrate the versatility and utility of the data cube for various environmental modeling purposes, we also present a case study focused on building regression modeling for soil organic carbon (SOC) using this data cube. Finally, we discuss the recommended uses, limitations, and future development of this data cube and compare it with other similar projects and initiatives. All the layers we produced are available under the open data license CC-BY and can be accessed and visualized via the following link: <https://stac.ecodatacube.eu> (last access: 27 June 2024) as Cloud-Optimized GeoTIFFs (Witjes et al., 2023). To enable users to easily assess the suitability of datasets for specific applications, metadata are embedded within the layer name in a structured and standardized format, aligning with data-fitness-for-use assessment methodologies (Yang et al., 2013; Pôças et al., 2014; Lokers et al., 2016; Wentz and Shimizu, 2018; Lacagnina et al., 2022). A copy of the data, including bimonthly reflectance bands, spectral indices, annual aggregations, and long-term analysis, is also available on Zenodo (Tian et al., 2024a, <https://doi.org/10.5281/zenodo.10776891>); the complete code used to derive all indices is available via <https://doi.org/10.5281/zenodo.12907281> (Tian et al., 2024b).

## 2 Material and method

### 2.1 Landsat ARD V2

The data cube presented in this study is based on the Landsat analysis-ready dataset version 2 (ARD V2), developed by the Global Land Analysis and Discovery (GLAD) team at the University of Maryland (Potapov et al., 2020). Landsat ARD V2, the second version of Landsat ARD, consists of 16 d tiled composites with 23 images per year from 1997 to 2022, totaling 598 images. It includes seven reflectance bands (blue, green, red, near-infrared, short-wave infrared 1, short-wave infrared 2, and thermal) and a detailed quality flag that classifies each pixel as land, water, cloud, cloud shadow, topographic shadow, hill shade, snow, haze, cloud proximity, shadow proximity, other shadows, or buffered proximity of the previously mentioned classes. These quality flags enable the identification of poor-quality pixels, including those affected by clouds, cloud shadows, haze, or other shadowy conditions. The presence of these gaps requires significant preprocessing before the data can be used for direct modeling and analysis. As one of the few globally consistent archives for historical time series of normalized surface reflectances derived from various Landsat satellite collections, Landsat ARD V2 offers long-term availability (since 1997) and detailed spatial resolution (30 m). For this study, we cropped the global Landsat ARD V2 layers to a pan-European extent, including Ukraine, the UK, and Türkiye. The GLAD ARD tile system is adopted to support data organization, parallel processing, and spatial inference when necessary.

### 2.2 TSIRF framework and scikit-map

As a comprehensive framework for processing EO time series, the Time Series Iteration-free Reconstruction Framework (TSIRF) enables diverse time series processing techniques by simply adjusting the convolution kernel (Consoli et al., 2024). We adopted TSIRF for temporal aggregation and gap filling to impute data gaps created after removing clouds using the ARD 2 quality assessment mask. This framework is implemented through the Python package scikit-map (<https://github.com/openlandmap/scikit-map>, last access: 27 June 2024), which optimizes speed by leveraging data structures for parallel computing. In addition to TSIRF, scikit-map also offers straightforward capabilities for band operations, trend calculations, temporal statistics, parallel raster reading and writing, and processing. These functionalities enable efficient processing, analysis, and visualization of large multidimensional raster datasets. In this work, all data processing and map production were performed using scikit-map Python library.

### 2.3 Spectral index predictor data cube

The preparation of the spectral biophysical index data cube involves developing four tiers of predictors from Landsat ARD V2 step by step. These tiers differ in temporal resolution and processing complexity, as outlined in Fig. 1 and detailed in the following subsections. Section 2.3.1 introduces the production of tier-1 bimonthly, gap-free Landsat reflectance bands. Tier-2 predictors are bimonthly time series for six biophysical indices (Sect. 2.3.2). From tier-2 indices, time series analyses are conducted to derive annual predictors (Sect. 2.3.3). Finally, further time series analyses are applied to tier-3 predictors to extract long-term temporal features representing persistent states (Sect. 2.3.4).

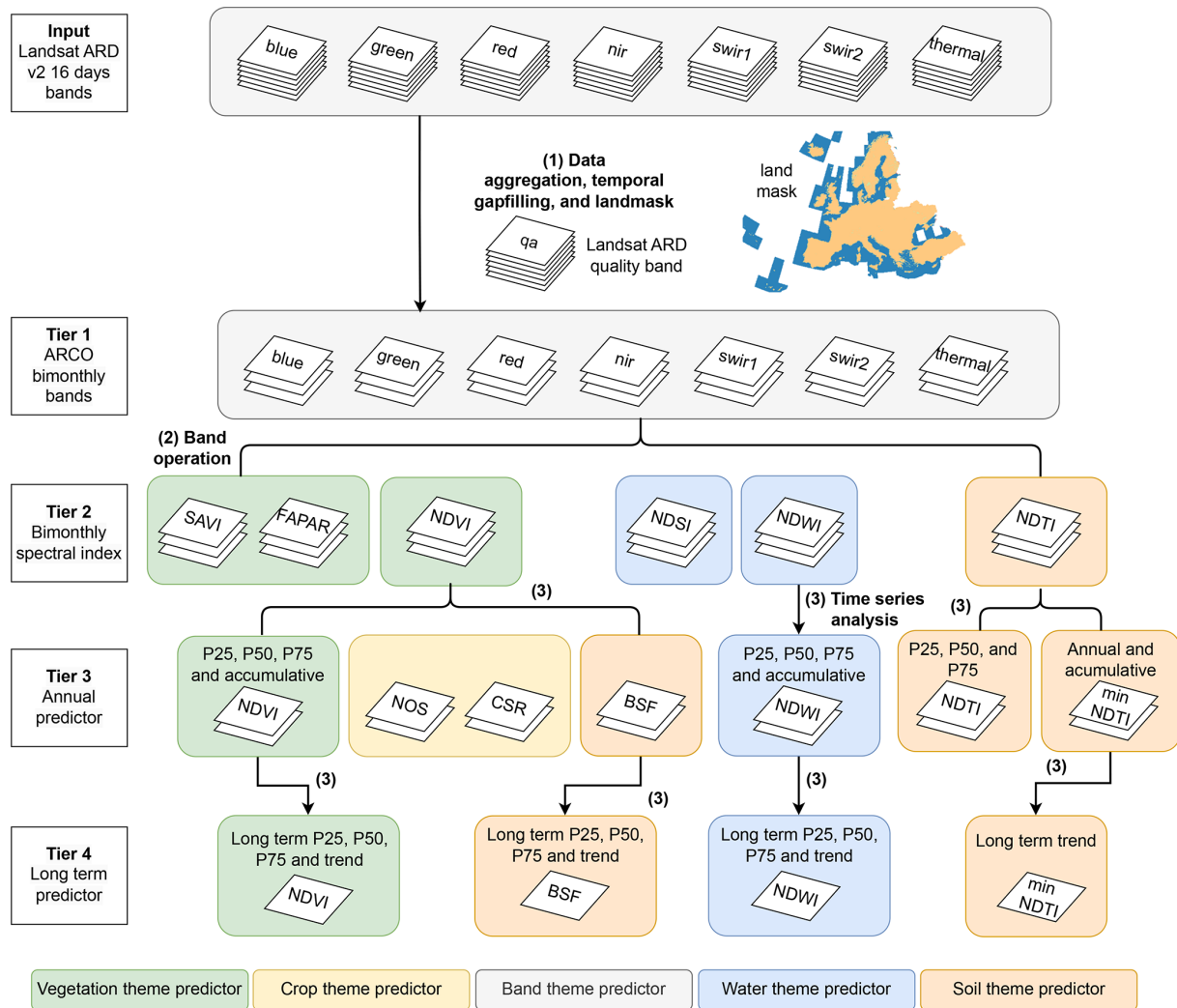
#### 2.3.1 First tier of predictors: bimonthly Landsat bands

The first-tier predictors are ARCO land surface reflectance bands, gap-filled using Landsat ARD V2 data. Pixels with insufficient quality were identified using the quality band and masked, creating gaps in the time series. These gaps were initially reduced by calculating a weighted average from several scenes (typically six to seven) within and adjacent to the 2-month period of the original 16 d interval time series, resulting in a bimonthly product (one image every 2 months). Despite this aggregation, significant gaps remained, as shown in Fig. 2, which were subsequently filled using the seasonally weighted average generalization (SWAG) method developed by Consoli et al. (2024). SWAG assigns weights for aggregation based on the clear-sky fraction of the corresponding tiles. For gap filling, SWAG respects seasonality and causality by reconstructing each pixel's time series using weighted averages, prioritizing images from the same period in previous years and relying only on past data.

This tier-1 product is part of the global bimonthly aggregation framework established by Consoli et al. (2024) and represents the European subset of that effort. Within this work, the data have been refined for regional applications by applying the pan-European land mask (<https://doi.org/10.5281/zenodo.8171860>, Tian et al., 2023) to exclude waterbodies and assembling it into continental mosaics stored as COGs, as shown in step 1 of Fig. 1. These ARCO Landsat surface reflectance layers that are gap-filled bimonthly provide the foundational material for producing spectral indices and their derivatives.

#### 2.3.2 Second tier of predictors: bimonthly spectral index

The second tier of predictors is calculated per pixel using ARCO surface reflectance bands, as shown in step 2 of Fig. 1. Direct satellite-derived indices make use of the unique spectral signatures of different objects on the surface of the land to detect their presence within a pixel. Direct satellite-derived indices include the normalized difference vegetation index (NDVI), soil adjusted vegetation index (SAVI),



**Figure 1.** General workflow for processing Landsat-based spectral index predictors. They formed four different tiers based on their level of processing and different temporal resolutions throughout the workflow. The predictors could also be categorized into five thematic groups, each framed by different colors in the figure, including vegetation, water, band-specific properties, crops, and soil characteristics. Step 3 in the workflow, time series analysis, incorporates three different temporal operations: temporal aggregation, which extracts statistical measures over specific periods to represent those intervals; trend analysis, which evaluates the directional changes in a predictor over time; and cumulative analysis, which sum up the corresponding annual predictor values from a starting baseline year (i.e., 2000) to the specified index year. Land mask source: <https://doi.org/10.5281/zenodo.8171860> (Tian et al., 2023).

fraction of absorbed photosynthetically active radiation (FAPAR), normalized difference tillage index (DTI), normalized difference water index (NDWI), and normalized difference snow index (NDSI). The calculation formula we adopted is described in Table 1.

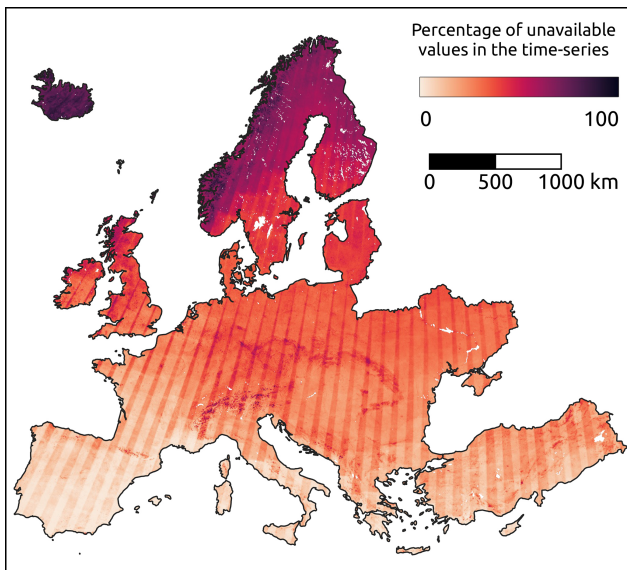
### 2.3.3 Third tier of predictors: annual aggregated index

In step 3, the aggregated annual indices are derived by extracting key temporal features or statistics from the bimonthly index series for each year. The following indices are then provided at a 30 m spatial resolution on an annual basis:

- *Annual P25, P50, and P75 aggregation of NDVI, NDWI, and NDTI.* These are calculated by identifying the values at the 25th, 50th, and 75th percentile of the sorted bimonthly NDVI and NDWI values for each pixel within a year.
- *Minimum normalized tillage index (minNDTI).* This index is determined as the minimum value of the six NDTI values over a year (Zheng et al., 2012).
- *Bare soil fraction (BSF).* This index is calculated by dividing the number of pixels classified as bare surface within a year’s time series (identified by the criterion of

**Table 1.** Overview of the third tier of predictors.

Index	Band operation	Reference
NDVI	$\frac{\text{nir}-\text{red}}{\text{nir}+\text{red}}$	Tucker (1979)
NDTI	$\frac{\text{swir1}-\text{swir2}}{\text{swir1}+\text{swir2}}$	Van Deventer et al. (1997)
NDWI	$\frac{\text{nir}-\text{swir1}}{\text{nir}+\text{swir1}}$	Gao (1996)
NDSI	$\frac{\text{green}-\text{swir1}}{\text{green}+\text{swir1}}$	Salomonson and Appel (2006)
SAVI	$\frac{(\text{nir}-\text{red})}{(\text{nir}+\text{red}+0.5)} \times 1.5$	Huete (1988)
FAPAR	$\frac{(\text{ndvi}-0.03) \times (0.95-0.001)}{0.96-0.03} + 0.001$	Robinson et al. (2018)



**Figure 2.** Per-pixel count of available values in the bimonthly aggregated Landsat time series from 1997 to 2022. Darker areas are more affected by presence of data gaps in the time series. In addition to cloud presence and snow cover, it is possible to notice patterns determined by overlapping scenes in the original Landsat raw images. Adapted from Consoli et al. (2024).

NDVI values below 0.35) by the total number of pixels analyzed in that year (Castaldi et al., 2019).

- *Number of season (NOS)*. This index indicates the frequency of cropping cycles within a year, calculated by identifying peaks in the NDVI time series (Li et al., 2014; Liu et al., 2020). Time steps with NDVI values greater than 0.5 are first flagged as candidate peaks, possibly corresponding to an actual cropping cycle. A prominence filter of 0.25 ensures that each candidate peak is at least 0.25 higher than its surrounding troughs, addressing consistently high NDVI values in areas such as forests. Adjacent peaks with intervals shorter than 2 months are merged, keeping the one with higher NDVI values, to ensure an accurate representation of distinct cropping cycles.

- *Crop duration ratio (CDR)*. It quantifies the length of active cropping periods, calculated as the ratio of time during which a pixel is in an active cropping state, defined as when the vegetation signals reach at least half the amplitude of the phenological curve's peak values (White et al., 1997). The peak values are determined by the average NDVI values of the peaks identified during the calculation of NOS.
- *Accumulated NDVI P50, NDWI P50, NDVI of the months of July and August, and minNDTI*. The cumulative indices are calculated by summing their values over time from the year 2000. They are expected to reflect the cumulative effects of water content, vegetation health, and tillage practices on each pixel.

#### 2.3.4 Fourth tier of predictors: long-term temporal feature index

In step (4), the indices that quantify long-term temporal features over 2000–2022 are calculated. The long-term change trends are derived from annual aggregates to represent the changes from 2000 to 2022. For this, we adopted the Theil–Sen estimator to fit the trend for 2000–2022. This is a statistical, non-parametric approach that is resistant to outliers as it calculates the median slope among all point pairs (Theil, 1950; Sen, 1968). This results in trend maps for bare soil fraction (BSF), NDVI P50, NDWI P50, and minNDTI. Additionally, the 25th, 50th, and 75th percentiles of NDVI, NDWI, and BSF are derived for the years 2000–2022 to provide a general overview of their distribution and general state over the this period.

#### 2.4 Plausibility check

Due to the limited availability of land survey data that align temporally, spatially, and thematically with our data cube, we were unable to perform traditional accuracy validation on most index layers, which requires extensive ground-truth datasets for direct comparison. Instead, we conducted a plausibility check to assess whether the data align logically and statistically with established land surface processes. Apart from the tier-1 product, we focus our quality assessment on

the third tier of predictors – annual aggregated indices such as minNDTI, BSF, NOS, and CDR – since they are more complex and less established than simpler spectral indices like NDVI and NDWI.

#### 2.4.1 Quality assessment of tier-1 product

The quality of the gap-filled time series for bimonthly Landsat surface reflectance bands was evaluated using 2746 time series randomly sampled from Europe. Details are provided in the supplemental computational notebook available in the “Code availability” section. To simulate real-world conditions, 10 % artificial gaps were introduced into the data. Performance metrics were used to assess the accuracy of the TSIRF model, including  $R^2$ , which measures the proportion of variance explained by the model; root mean squared error (RMSE), quantifying the average magnitude of prediction errors; and concordance correlation coefficient (CCC; Lawrence and Lin, 1989), evaluating the agreement between observed and predicted values. These are reported to indicate the accuracy of TSIRF in reconstructing the tier-1 land surface reflectance bands.

#### 2.4.2 Index statistics by crop type: BSF, NOS, and CDR

To examine the effectiveness of crop-specific parameters – BSF, NOS, and CDR – in reflecting the crop patterns; we used the harmonized LUCAS (Land Use and Cover Area frame Survey) data, which detail crop type information at specific locations and times (d’Andrimont et al., 2020). Three NUTS2 regions are selected to capture a diverse array of European agricultural practices and climates: Picardy region (FRE2), Piedmont province (ITC1), and Malopolska province (PL21). Picardy is chosen for its temperate maritime climate and highly mechanized large-scale farms. Piedmont is included to represent its unique damp rice paddies, which illustrate specialized crop cultivation. Malopolska offers insights into agriculture under a more continental climate, in contrast with other regions. We identified points belonging to four predominantly grown crops in the selected NUTS2 regions – FRE2, ITC1, and PL21 from the most recent LUCAS 2018 dataset. These points were then overlaid onto the 2018 index maps for BSF, NOS, and CDR, followed by the calculation of the statistics for crop-specific indices among the different crop types. By comparing the statistics of these indices across regions and crops, we conducted an evaluation of the indices’ sensitivity and accuracy in capturing the distinctions in cropping patterns and reflecting the specific agricultural dynamics.

#### 2.4.3 Correlation with land survey data: BSF

We compared BSF maps against two land survey datasets, with approximately 150 records of crop cover duration from 2007 to 2016 on cropland (Edlinger et al., 2022, 2023).

The hypothesis is that the proportion of time that crops cover the soil for is inversely related to the presence of bare soil during the same time period. By implementing a correlation analysis between BSF values derived from satellite data and ground-based measurements of crop cover duration, our objective was to assess the reliability of the BSF in capturing the extent of bare soil across agricultural landscapes.

#### 2.4.4 Correlation with regional statistics: minNDTI

To validate the minNDTI’s effectiveness in reflecting tillage practices across the EU, we compared it with Eurostat’s tillage area statistics ([https://doi.org/10.2908/EF\\_MP\\_PRAC](https://doi.org/10.2908/EF_MP_PRAC), Eurostat, 2024). Eurostat’s data originate from farm structure surveys conducted in 28 EU countries and are available in a vector shapefile format covering 319 NUTS2 regions and detailing arable areas under different tillage practices. The breakdown of arable land area based on tillage practices is as follows:

$$\begin{aligned} \sum(\text{TIL\_CV} + \text{TIL\_CSERV} + \text{TIL\_ZERO} + \text{ARAXTIL}) \\ = \text{ARA}, \end{aligned} \quad (1)$$

where TIL\_CV refers to areas under conventional tillage, TIL\_CSERV refers to areas under conservation tillage, TIL\_ZERO refers to areas under zero tillage, ARA refers to total arable land areas, and ARAXTIL refers to areas land excluding from tillage survey; all areas are recorded in hectares. Notably, ARAXTIL represents land that is excluded from tillage survey due to not being sown/cultivated during the reference year, which includes areas with multi-annual plants, such as temporary grassland, leguminous plants, and industrial crops (hops or aromatic plants). These are recorded in hectares and together make up the total arable land area (ARA).

To spatially match the Eurostat data with our predictor layers, the data preparation process involved four steps:

1. *Tillage share calculation.* For valid NUTS2 data records with all four types of tillage practices, calculate the share of each tillage practice within each NUTS2 region.
2. *Crop masking.* Remove non-arable pixels from the minNDTI maps using the EU crop mask developed by d’Andrimont et al. (2021).
3. *Spatial overlay.* Overlay the Eurostat survey data with crop-masked minNDTI map layers of matching years.
4. *Average minNDTI calculation.* Calculate the average minNDTI value within each NUTS2 region.

The relationship between each tillage practice and minNDTI is assessed separately using ordinary least-squares (OLS) regression models, correlating the shares of each tillage practice with the average minNDTI values within

**Table 2.** Typical CRC value range for each tillage practice type.

Tillage practice type	CRC range (%)
Conventional tillage	0–30
Conservation tillage	30–70
Zero tillage	70–100
Excluded tillage practices	0–100

each NUTS2 region. To capture the collective impact of all tillage practices, we introduced the concept of weighted crop residue cover (WCRC). This metric aggregates the influence of various tillage practices, considering minNDTI's correlation with crop residue cover (CRC):

$$\text{WCRC} = \sum_{\text{tillage types}} (\text{typical CRC value} \times \text{area share}), \quad (2)$$

where WCRC integrates typical CRC values for each tillage type, weighted by their respective area shares within each NUTS2 region.

We fitted typical CRC values for each type of tillage practice from their respective ranges, as outlined in Table 2 (Magdoff and Van Es, 2000; Zheng et al., 2012). This optimization process identified the values within these ranges that maximize the correlation between WCRC and minNDTI. For excluded tillage practices, we assume a broad range of 0 % to 100 % to cover all possibilities, reflecting the uncertainty in its typical CRC values. The correlation analysis assesses the relationship between the estimated WCRC and the average minNDTI of the NUTS2 regions.

## 2.5 Comparison to EcoDataCube in LC classification

The Landsat-based predictor data cube is intended to be an improvement of the version in EcoDataCube (Witjes et al., 2023). The switch to a bimonthly temporal aggregation from a seasonal one with three percentiles leads to a smaller number of variables (6 instead of 12) and a higher level of temporal resolution. In order to quantify the difference between these aggregation techniques, we compare the performance of sets of random forest classifiers trained on both sets of Landsat data to predict seven types of land cover: cropland, grassland, bare land, forest, artificial land, shrubland, and wetlands. We exclude the EcoDataCube non-Landsat data (digital terrain model, DTM, and Sentinel-2) and indices in this work to guarantee an objective comparison. Details about the training data are discussed in the next section, while an overview of the experiment workflow is presented in Fig. 3.

## 2.6 Modeling experiments

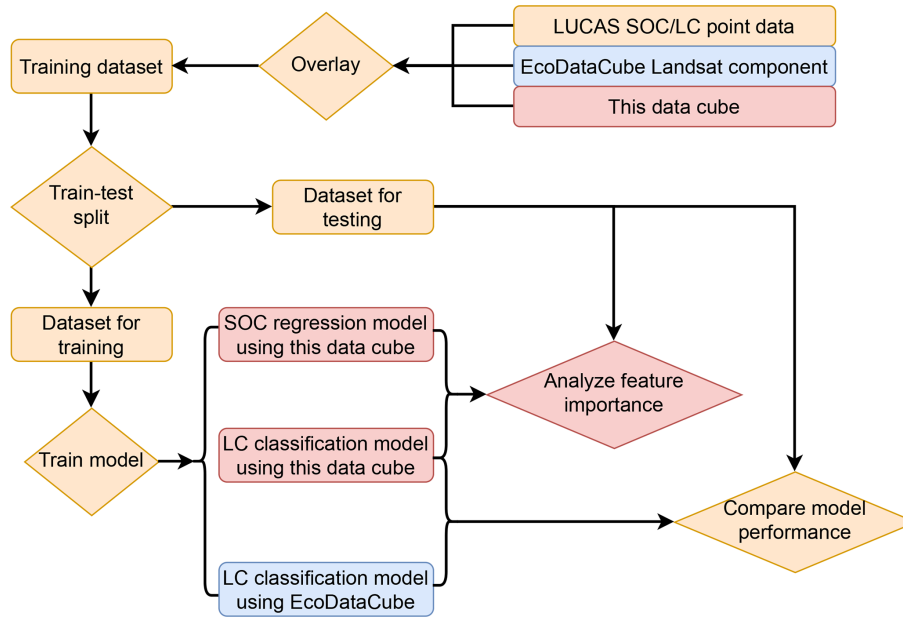
To demonstrate the utility of the indices in the data cube, we performed two modeling experiments: soil organic carbon (SOC) regression and LC classification. The general

workflow is depicted in Fig. 3. The data points used for SOC regression and land cover classification come from the LUCAS land cover and top soil surveys available. Only surveyed point data (i.e., identified by unique point id and sampling year) with both land cover and topsoil SOC data were selected for modeling experiments. These point data were overlaid with EcoDataCube Landsat data, along with all Landsat reflectance and biophysical index layers generated in this study, to create the training dataset for modeling experiments. Using stratified sampling, we select about 10 % of test data points based on a 30 km grid tiling system to ensure that the test data accurately represent the spatial distribution throughout the study area. This results in 52 306 training data points and 8417 test data points, which is, in total 60 723 points.

The models' performance for both tasks is evaluated by training them with different combinations of predictors, organized into two main categories: themes and tiers. The first division categorizes the predictors by themes: reflectance bands, vegetation, water, soil, and crop (see Table 3). The second division organizes predictors by tiers, which represent different levels of temporal aggregation and processing, as detailed in Sect. 2.3. These tiers, as described above, include bimonthly bands, bimonthly indices, annual indices, and long-term indices. By systematically testing models with these combinations, we can understand how indices developed for different purposes and how short-term variations, seasonal dynamics, and long-term temporal features contribute to the accuracy and robustness of predictive models. Both SOC regression and land cover classification models were trained using different groups from both divisions. We examined the feature space to identify the most influential predictors and compared their importance in the regression and classification tasks. This analysis aimed to highlight the flexibility and comprehensive nature of our data cube. Apart from comparing different feature combinations, we also provide the top 20 most important features for the model trained on all predictors and all point data. This provides a focused analysis that identifies which specific predictors are the most influential across the entire dataset when no subsets are used.

For the SOC regression experiment, we used random forest (RF) as an estimator to predict the SOC content throughout the EU. The target variable, SOC, was obtained from the LUCAS soil survey. We evaluated the performance of the regression model using different combinations of predictors to identify which set provided the most accurate predictions. Performance metrics used to assess model accuracy include  $R^2$ , RMSE, and CCC. We used a random forest regression model with the following parameters: a maximum depth of 20 to control complexity, sqrt for maximum features to improve performance, a minimum of two samples per leaf to prevent overfitting, a minimum of five samples per split for stable node splitting, and 800 estimators to ensure robust predictions. These parameters were selected through empirical testing and cross-validation, implemented using the scikit-





**Figure 3.** Workflow for modeling experiments, illustrating the comparison between EcoDataCube and the proposed data cube in the LC classification experiment as well as the feature importance analysis conducted for the LC classification model and the SOC regression model using the proposed data cube. Components specific to EcoDataCube are marked in blue, those unique to the data cube produced in this work are marked in red, and shared general components are marked in yellow.

**Table 3.** Theme-based division of predictors.

Theme	Predictors
Reflectance bands	Bimonthly Landsat bands Annual P25, P50, and P75 of Landsat bands
Vegetation	Bimonthly NDVI, SAVI, FAPAR Annual NDVI P25, P50, and P75; accumulated NDVI P50 and NDVI of the bimonth of July and August Long-term P25, P50, P75, and trend of NDVI
Water	Bimonthly NDWI, NDSI Annual NDWI P25, P50, and P75 Long-term P25, P50, P75, and trend OF NDWI
Soil	Bimonthly NDTI Annual minNDTI, NDTI P25, P50, and P75, and BSF Long-term P25, P50, P75, and trend of BSF, and long-term trend of minNDTI
Crop	Annual NOS and CDR

learn library in Python. The metrics will also be calculated in the log1p space.

For the LC classification experiment, we employed a random forest classifier to predict LUCAS LC classes. Similarly to the SOC regression, we tested the model’s performance with different combinations of predictors to determine the most effective feature sets. Performance metrics used to evaluate the classification model include precision, assessing the accuracy of positive identifications; recall, evaluating the model’s ability to capture all relevant cases; and  $F_1$  score

(Taha and Hanbury, 2015), providing a balance between precision and recall for a comprehensive model assessment.

### 3 Result

This section begins with an accuracy assessment of time series reconstruction using the TSIRF method to generate tier-1 predictors. This is followed by a qualitative visual examination of spatial patterns in three zoomed-in areas, each representing distinct dominant land cover types and unique land dynamics over the past decades. Additionally, long-term

**Table 4.** Performance metrics for the reconstruction of the tier-1 product.

Metric	Blue	Green	Red	NIR	SWIR1	SWIR2	Thermal
RMSE	0.04	0.04	0.04	0.05	0.03	0.02	0.01
RMSE/ $\mu$	0.86	0.51	0.56	0.20	0.15	0.22	0.01
RMSE/ $\sigma$	0.56	0.56	0.58	0.60	0.41	0.40	0.32
$R^2$	0.67	0.67	0.67	0.62	0.83	0.84	0.90
CCC	0.83	0.82	0.81	0.79	0.91	0.91	0.95

trends of BSF, NDVI, and NDWI are analyzed using annual time series data from 2000 to 2022 at three LUCAS points within these areas. These analyses provide both spatial and temporal insights into land cover changes and agricultural practices, demonstrating how these indices capture and reflect patterns and trends over time. Where relevant data are available, quantitative assessments are also incorporated to further validate our approach, as elaborated in Sect. 2.4.

### 3.1 Accuracy of reconstructed tier-1 predictors

Table 4 presents performance metrics for the reconstructed Landsat surface reflectance bands across seven tier-1 spectral bands. Overall, the reconstruction demonstrates high accuracy, with low RMSE, high  $R^2$ , and CCC across all bands. The thermal band consistently shows the best performance, while the NIR band exhibits the lowest performance, with slight variations observed among the other bands. In a similar global-scale experiment, Consoli et al. (2024) reported generally higher reconstruction performance. For a more comprehensive perspective on global time series reconstruction, we refer readers to their work, which focuses exclusively on this topic.

### 3.2 Visual examination of tier-3 predictors

Figure 4 presents zoomed-in examples of tier-3 predictors: NDVI P50, NDWI P50, BSF, NOS, and CDR. They provide an overview of vegetation health, moisture levels, bare soil exposure, number of growing seasons, and crop duration. These visual representations complement the statistical analysis by highlighting spatial patterns that may not be evident through quantitative methods alone, providing an intuitive understanding of the spatial patterns of the indices. There are three presented regions with distinct features (Fig. 4): (A) an agricultural area in northern France, characterized by well-organized grid-like field patches of various crops; (B) a rice cultivation region that featured less regular field texture in northern Italy; and (C) the former Szczakowa sand mine in southern Poland (Pietrzykowski and Krzaklewski, 2007; Pietrzykowski, 2008).

NDVI and NDWI delineate agricultural patches with a 30 m spatial resolution (Fig. 4, region A). Forested areas show high values of NDVI and NDWI, while urban areas (scattered towns in Fig. 4, regions A and B) and ex-

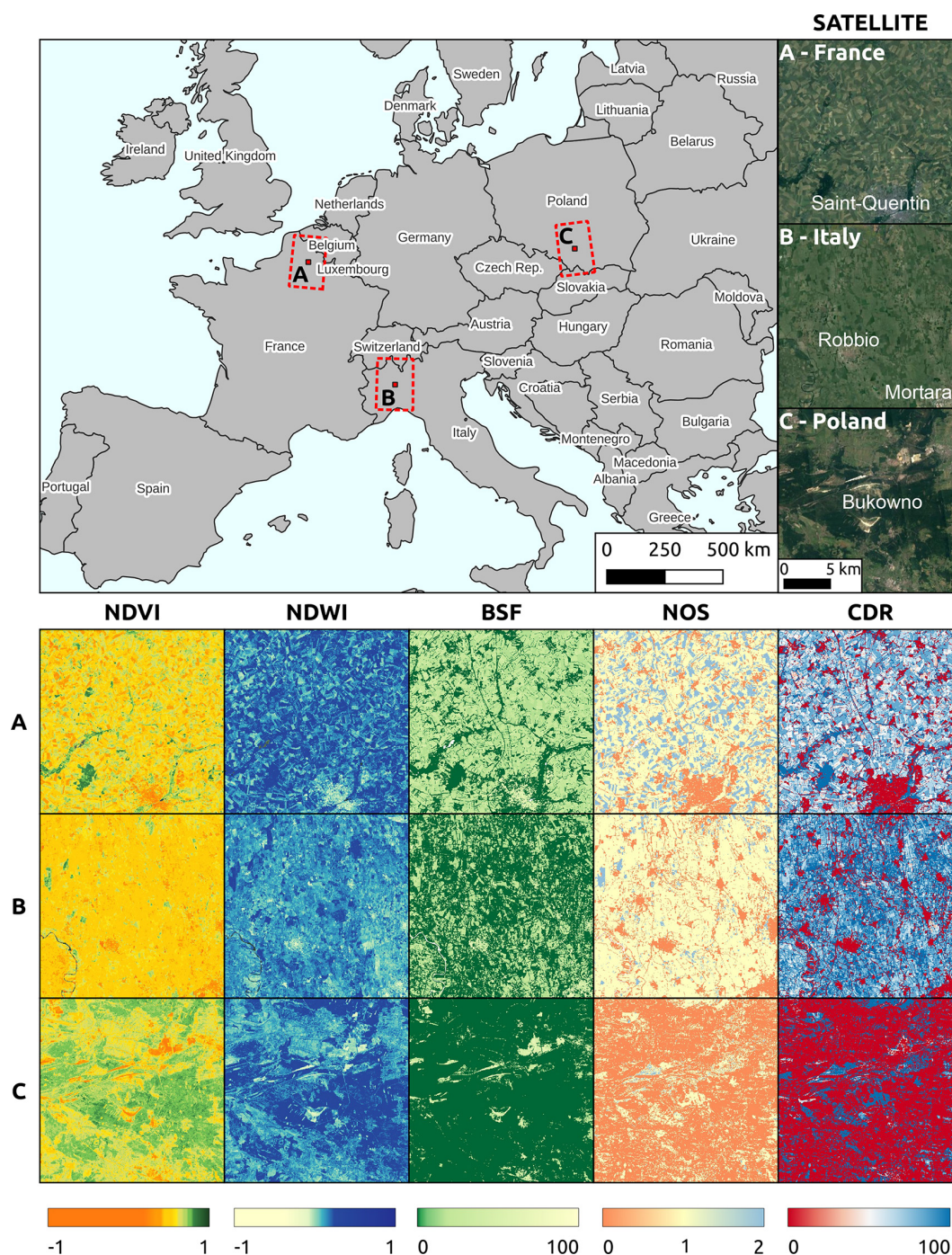
posed land from mining activities (Fig. 4, region C) exhibit low values. High NDWI values in region B, attributed to flooded paddy fields, do not always coincide with a high NDVI (Ranghetti et al., 2016). For example, there are several patches in Fig. 4 – B exhibiting high NDWI but moderate NDVI values. This observation could be attributed to the paddy fields in region B, which are flooded in a controlled way to meet the water demand of the rice plants (Ranghetti et al., 2016).

Although derived from the NDVI annual time series, the BSF shows different patterns. Urban and exposed soil areas have high BSF values, while woodlands have low values. Despite lower NDVI P50 values, region B has lower BSF values due to stable NDVI levels above a threshold, indicating less bare soil exposure throughout the year. This highlights that BSF reveals different aspects of the landscape compared to NDVI P50, making it a valuable complement to assess exposure to bare soil and vegetation dynamics. The crop-specific NOS and CDR provide insights into growing seasons and their duration. CDR highlights the duration for which crops occupy a pixel within a year and generally inversely relates to BSF. Most croplands in regions A and B have a NOS value of 1, indicating a single annual growing season, while urban and forested areas have a NOS value of 0 (Fig. 4, region C). Some areas show NOS values of 2, indicating possible practices such as winter cropping or double-cropping systems.

### 3.3 Trend analysis with land cover dynamics

Four trend maps were produced to analyze land surface trends between 2000 and 2022: NDVI annual median, NDWI annual median, BSF, and minNDTI (Fig. 5). The NDVI trends reveal a positive trend in most parts of Europe. However, specific regions, including the Alps, northern Europe, northern Scotland, Iceland, and Scandinavia, exhibit clear negative NDVI trends. These areas, characterized by high latitudes or altitudes, often experience persistent snow cover and frequent cloudiness. The NDWI trend map for these regions shows a positive trend, indicating an increase in vegetation water content. In other parts of Europe, the NDWI trends display a mix of positive and negative changes, lacking a definitive overall trend. The decreasing trend in NDVI in regions of high altitude and high latitude is also reflected in the BSF trend map (Fig. 6). In Spain and Türkiye, strong positive and negative trends are scattered, with the positive trend being more dominant. Generally, our results show an increase in minNDTI from 2000 to 2022 across Europe. However, there are exceptions, such as the negative trends observed in regions like Türkiye.

In region A, northern France, all BSF, NDVI, and NDWI trends show scattered increases and decreases, likely influenced by well-planned crop patches and their varying management practices (Fig. 7). Most areas exhibit positive NDVI and NDWI trends, except for urban areas such as Saint-Quentin, where both vegetation and water content are declin-

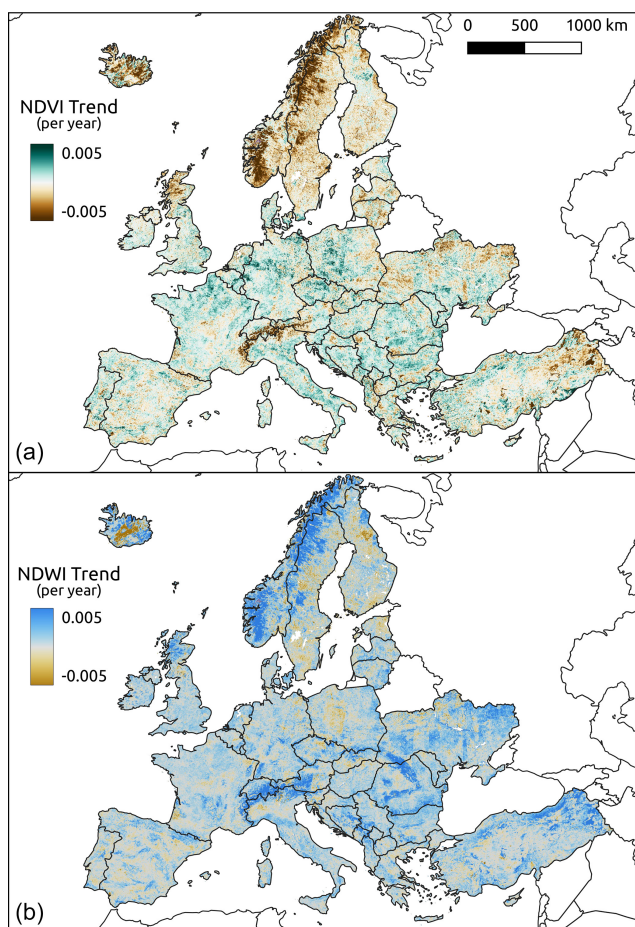


**Figure 4.** Zoomed-in examples of NDVI and NDWI median in the year 2018 in (A) northern France, with the city of Saint-Quentin in the southeast corner; (B) northern Italy, in the western Po Valley, east of Vercelli; and (C) southern Poland, within the Przemsza River basin, featuring the Szcakowa sand mine. Each image tile has a side length of 20 km. Satellite imagery source: © Google Earth (<https://earth.google.com/web/>, last access: 27 June 2024).

ing. The time series also shows significant variation, influenced by crop types and management practices.

In region B, northern Italy, the NDVI trends exhibit variability, but the overall trend is not as positive as in region A (Fig. 8). BSF generally remains stable, while some areas ex-

hibit an increasing trend, indicating greater soil exposure. The dominant decline in NDWI highlights a drying trend, likely due to the increasing adoption of dry seeding practices in rice cultivation (Ranghetti et al., 2016). The time series for



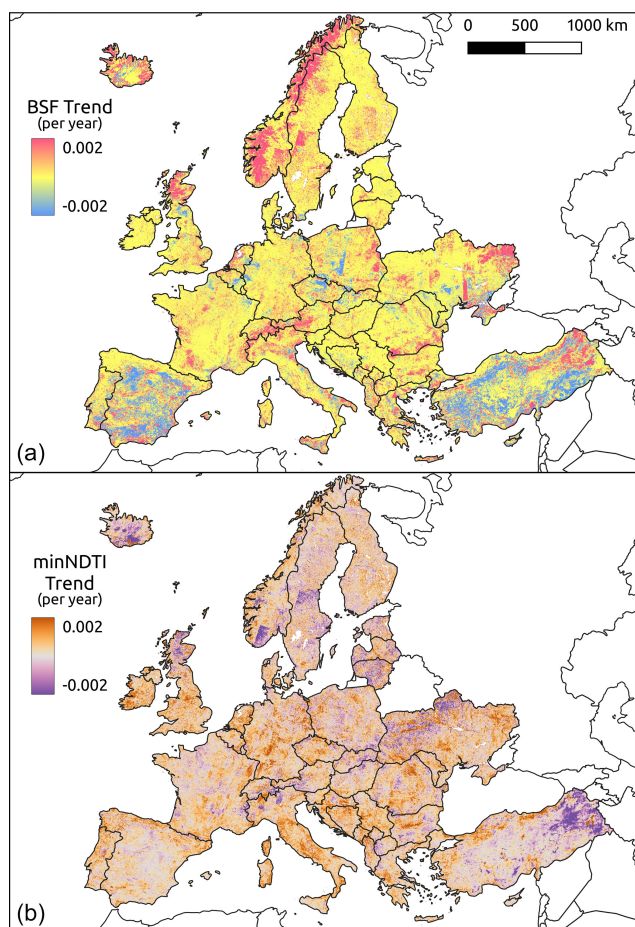
**Figure 5.** Long-term changing trend between 2000 and 2022 of NDVI (a) and NDWI (b) as the annual rate of change in index values.

the three indices has remained relatively stable over 20 years, reflecting consistent rice cultivation practices.

Region C in Poland, which has transitioned from a sand mine to a reforested area (Pietrzykowski and Krzaklewski, 2007; Pietrzykowski, 2008), shows the most significant positive trends in BSF, NDVI, and NDWI (Fig. 9). This indicates a steady increase in vegetation cover and water content, which is also reflected in the time series data. Although it was a coniferous forest from 2009 to 2018, the increases in NDVI and NDWI, coupled with the decrease in BSF, are likely due to the continuous growth of trees. Figure 10 shows satellite images from 2000, 2010, and 2020, each accompanied by the corresponding BSF images for comparison, clearly illustrating the progression of reforestation during the last 2 decades.

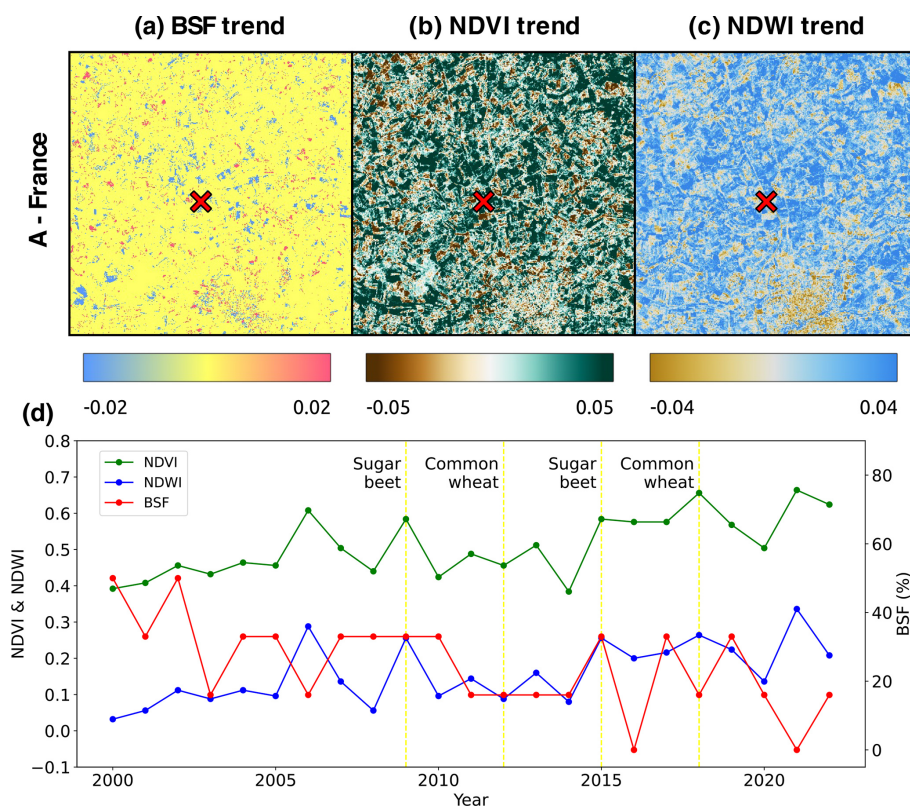
### 3.4 Crop specific indices statistics across European regions

Figure 11 shows the average values of BSF, NOS, and CDR for maize, sugar beet, and rice, which are common crops in



**Figure 6.** Long-term changing trend between 2000 and 2022 of BSF (a) and minNDTI (b) as the annual rate of change in index values.

the NUTS2 regions PL21, FRE2, and ITC1, respectively. The NDVI time series for 2018, from which these indices are derived, are shown in the bottom panel. Sugar beet in the FRE2 region has the highest average NOS value due to its distinct NDVI time series fluctuations, indicating pronounced seasonal variations. The NDVI time series for sugar beet shows a distinct peak during the summer months (July and August), a deep valley in March and April, and a less pronounced peak in winter. In comparison, the NDVI time series for maize in the PL21 region and rice in the ITC1 region display more uniform patterns, with smaller differences between peak and valley NDVI values, leading to lower NOS values. Sugar beet in the FRE2 region also exhibits the highest BSF values due to a long and deep valley period in the NDVI time series. In contrast, rice in the ITC1 region shows the lowest BSF values, reflecting its stable vegetation coverage throughout the year. Consequently, rice in the ITC1 region also has the highest CDR value. Although BSF and CDR often show a negative correlation, this is not always the case. For example, sugar beet in the FRE2 region has a higher BSF compared



**Figure 7.** The top panel (a–c) provides a zoomed-in view of NDVI, NDWI, and BSF trends for the Saint-Quentin region, north France (region A, identical in spatial location and extent to region A in Fig. 4), which has a consistent agricultural use with varied crop types across years. Images were generated by the authors using data produced in this study. Red crosses in the top panel mark the LUCAS point (ID: 38363000; lat 49.91122°, long 3.237571°), detailed in the bottom panel, panel (d). Panel (d) shows the time series of NDVI, NDWI, and BSF from 2000 to 2022 at this LUCAS point, which has been revisited four times in the years 2009, 2012, 2015, and 2018, with the surveyed land cover indicated beside vertical yellow lines.

to maize in the PL21 region but also has a higher CDR. This highlights that, although related, BSF and CDR can differ due to their different focus: BSF measures the duration of bare soil, whereas CDR emphasizes the active growing period.

Figure 12 shifts the focus from different crop types across regions to the same crop type – common wheat – grown in the FRE2, PL21, and ITC1 regions. The average values of BSF, NOS, and CDR for common wheat reveal notable regional variations despite the NDVI time series overlap. Common wheat in the FRE2 region displays significantly higher NOS values compared to the other two regions due to a secondary growth peak in winter. It also has the highest CDR value, indicating the most active growing cycles. In contrast, common wheat in the ITC1 region has the lowest BSF.

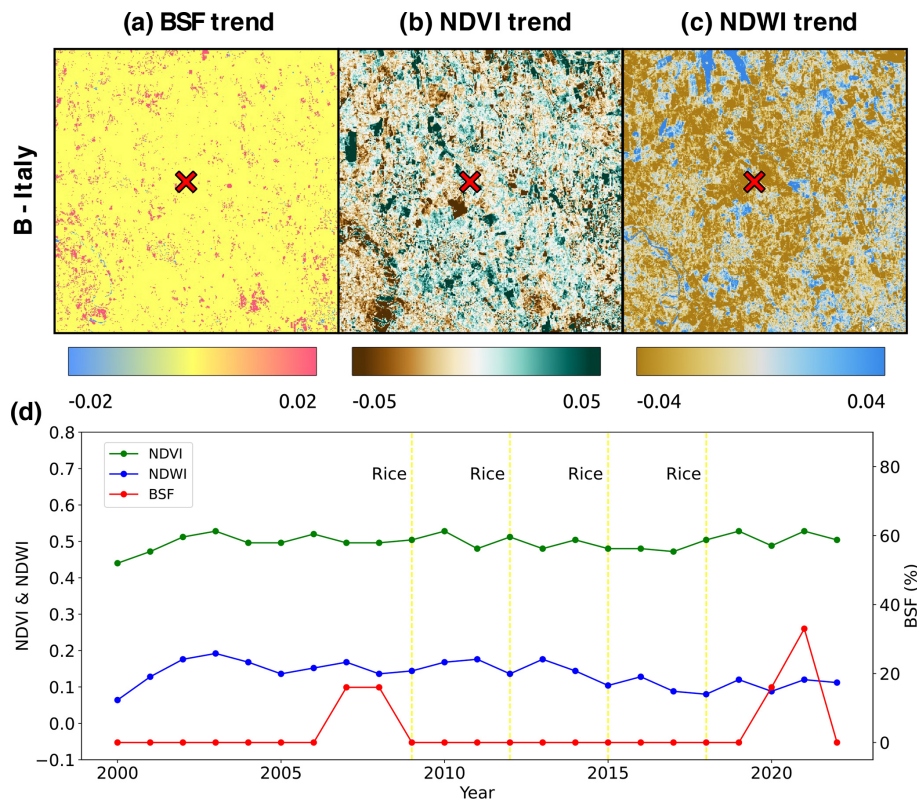
### 3.5 Assessing BSF in cropland

An inverse correlation, as illustrated in Fig. 13, is evident between the recorded crop coverage data and the BSF values from 2007 to 2016 in the cropland. Given the complexities involved, including the varying temporal resolutions of ob-

servations, the effects of crop residues, and the emergence of natural vegetation and weeds, a perfect one-to-one correlation between these variables is not expected. This observed correlation, in line with our hypothesis, confirms the validity of the BSF data.

### 3.6 Compare minNDTI with Eurostat at NUTS2 level

Figure 14 shows the Pearson correlation between the shares of each individual type of tillage practice and their impact on the minNDTI values. Although the correlation coefficients for all the tillage practices suggest only a moderate correlation between the mean minNDTI values and the respective shares of each tillage practice, comparing the regression coefficients provides valuable information about their relative impacts. The sensitivity of the average minNDTI to each tillage practice is represented by the regression slope coefficient obtained from the OLS model fitting. Conventional tillage practices, characterized by soil inversion and minimal soil cover post-harvest, exhibit the most negative influence on average minNDTI in NUTS2 regions. This is followed by conservation tillage practices. In contrast, zero-



**Figure 8.** The top panel (a–c) provides a zoomed-in view of NDVI, NDWI, and BSF trends for rural area besides city Vercelli in north Italy, with dominant rice cultivation (region B, identical in spatial location and extent to region B in Fig. 4). Images were generated by the authors using data produced in this study. Red crosses in the top panel mark the LUCAS point (ID = 42122479; lat 45.33338°, long 8.611634°), detailed in the bottom panel, panel (d). Panel (d) shows the time series of NDVI, NDWI, and BSF from 2000 to 2022 at this LUCAS point, which has been revisited four times in the years 2009, 2012, 2015, and 2018, with the surveyed land cover indicated beside vertical yellow lines.

tillage practices, which minimize soil disturbance, show the most positive effect on the average minNDTI. The area excluded from the tillage practices analysis, denoted as araxtil, consists mainly of multi-annual plants, resulting in a positive impact on the average minNDTI as indicated by the sensitivity analysis.

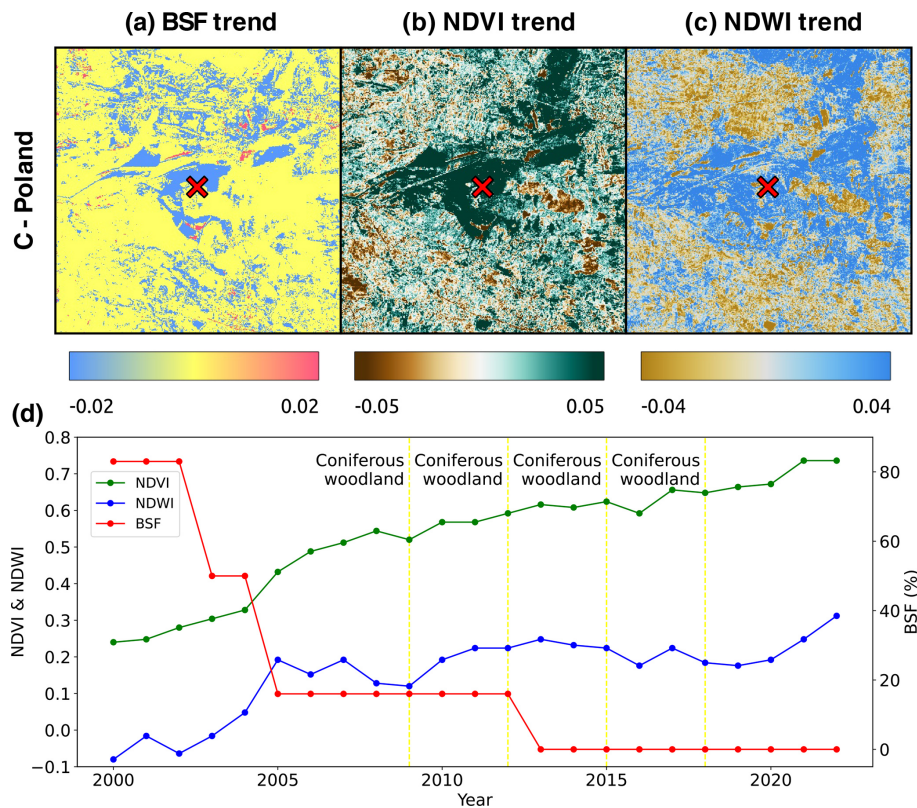
The correlation analysis between the WCRC and the average minNDTI of NUTS2 regions indicates that the cumulative impact of all tillage practices on the average minNDTI is more significant than any individual practice (Fig. 15). The optimal CRC for the tillage practices excluded from the analysis is relatively high at 75.08 %, which is consistent with Eurostat’s statement that the majority of land excluded from tillage practices comprises multi-annual plants.

To further illustrate the spatial distribution and temporal trends of minNDTI, we used the minNDTI map, masked by the crop mask of d’Andrimont et al. (2021), to generate maps showing the average minNDTI for 2016 across the croplands of each NUTS2 region. Furthermore, we created a map that illustrates the average minNDTI trend between 2000 and 2022 per NUTS2 region (Fig. 16). In 2016, NUTS2 regions in Scandinavia and Ireland exhibited high

minNDTI values, in contrast to regions in Spain and eastern Europe, which generally displayed lower minNDTI values on average. The trend map for minNDTI reveals that NUTS2 regions around the Baltic Sea are experiencing a relatively strong negative trend, indicating a possible increase in the intensity of tillage. In contrast, areas in eastern Britain, the Netherlands, and Austria exhibit a moderate positive trend. In general, the EU presents a negative trend in minNDTI. However, caution must be exercised when interpreting minNDTI data as the correlation between minNDTI values and tillage practices is not robust enough to draw definitive conclusions without additional information.

### 3.7 LC classification results compared to EcoDataCube.eu

Both LC models achieved a weighted  $F_1$  score of 0.77, indicating similar performance. However, the model trained on this Landsat-based predictor data cube (i.e., six bimonthly and three annual Landsat features per band) achieved a macro  $F_1$  score of 0.47 versus a macro  $F_1$  score of 0.45 for the model trained on Landsat bands from EcoDataCube.eu



**Figure 9.** The top panel (a–c) provides a zoomed-in view of NDVI, NDWI, and BSF trends for the former Szczakowa sand mine, south Poland (region C, identical in spatial location and extent to region C in Fig. 4). Images were generated by the authors using data produced in this study. Red crosses in the top panel (a–c) mark the LUCAS point, detailed in the bottom panel (ID = 49923058; lat 50.24478°, long 19.435963°). Panel (d) shows the time series of NDVI, NDWI, and BSF from 2000 to 2022 at this LUCAS point, which has been revisited four times in the years 2009, 2012, 2015, and 2018, with the surveyed land cover indicated beside vertical yellow lines.

(i.e., four quarterly Landsat data with three percentiles per quarter). This is mainly due to the higher classification accuracy for the relatively rare class of wetlands (202 points; see Fig. 17).

### 3.8 Feature importance analysis of SOC regression and LC classification

#### 3.8.1 Examination across tiers

From Table 5, we see the performance of the SOC regression model evaluated using different combinations of predictors categorized by tiers. When models are trained solely on one tier of predictors, tier-4 predictors (long-term characteristics) outperform the other tiers across all three metrics. In contrast, tier-2 predictors (bimonthly spectral indices) show the least optimal CCC and  $R^2$ , while tier-1 predictors (bimonthly reflectance bands) have the least optimal RMSE. For models trained on combinations of two tiers of predictors, the combination of tier-1 and tier-4 predictors performs the best across all three metrics, while the combination of tier-2 and tier-4 predictors performs the least well. When examining combinations of three tiers of predictors, most metrics are quite

close, but the combination of tier 1, tier 2, and tier 4 still stands out slightly. When comparing based on the number of tier sets combined, the general observation is that the more predictors combined, the better the model performed, with one exception: the model trained only on tier-4 predictors, despite being the smallest feature set, surprisingly achieved the best performance across all combinations.

An examination of the model performance across different tiers, as shown in Fig. 18, reveals that including tier-1 and tier-2 features generally has a negative effect. In contrast, incorporating tier-3 and tier-4 features positively impacts performance. Notably, tier-4 features show to have the most positive impact on model performance, consistent with the findings from Table 5.

Unlike the SOC regression experiment, the LC classification experiment indicates that combining features from tier 1, tier 2, and tier 4 results in the highest accuracy (see Table 6). Although tier 4 is the best single-tier feature set for models in terms of  $F_1$  score, precision, and recall, it does not outperform all possible feature combinations. Models utilizing tier-3 predictors (annual spectral bands and indices) show the least optimal performance. For models using two-tier combi-

**Table 5.** Metrics for SOC regression using different combinations of predictor based on tiers. The predictor combination yielding the best performance is highlighted in bold.

Number of tiers combined	Tiers	Feature number	CCC	RMSE	$R^2$
1	tier 1	42	0.635	0.651	0.489
	tier 2	36	0.623	0.660	0.476
	tier 3	38	0.637	0.652	0.487
	tier 4	13	<b>0.682</b>	<b>0.613</b>	<b>0.547</b>
2	tier 1 + tier 2	78	0.647	0.642	0.503
	tier 1 + tier 3	80	0.648	0.642	0.503
	tier 1 + tier 4	55	0.661	0.630	0.521
	tier 2 + tier 3	74	0.648	0.643	0.501
	tier 2 + tier 4	49	0.644	0.643	0.502
	tier 3 + tier 4	51	0.659	0.634	0.516
3	tier 1 + tier 2 + tier 3	116	0.652	0.639	0.507
	tier 1 + tier 2 + tier 4	91	0.664	0.629	0.523
	tier 1 + tier 3 + tier 4	93	0.663	0.630	0.522
	tier 2 + tier 3 + tier 4	87	0.662	0.631	0.520
4 (all predictors)	tier 1 + tier 2 + tier 3 + tier 4	129	0.665	0.629	0.523

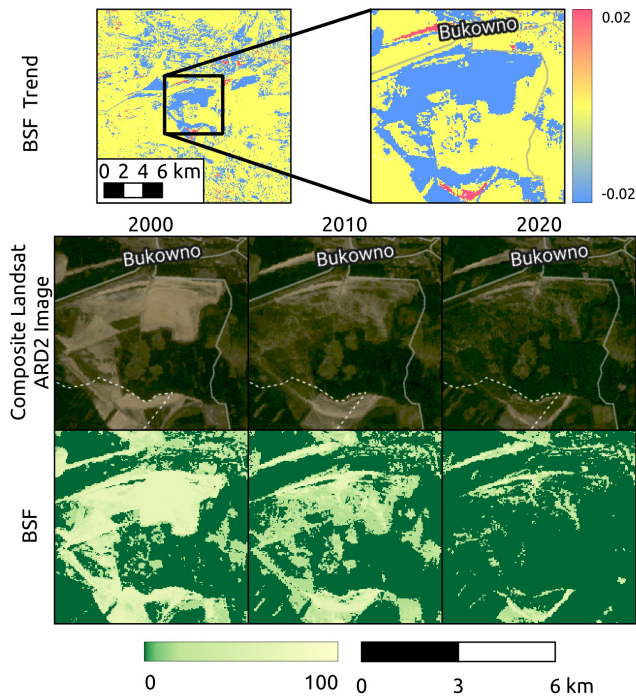
**Table 6.** Metrics for LC classification using different combinations of predictors based on tiers. The predictor combination yielding the best performance is highlighted in bold.

Number of tiers combined	Tiers	Feature number	$F_1$ score	Precision (UA)	Recall (PA)
1	tier 1	42	0.769	0.776	0.784
	tier 2	36	0.776	0.781	0.789
	tier 3	38	0.761	0.762	0.773
	tier 4	13	0.785	0.782	0.791
2	tier 1 + tier 2	78	0.782	0.787	0.794
	tier 1 + tier 3	80	0.774	0.777	0.787
	tier 1 + tier 4	55	0.785	0.792	0.799
	tier 2 + tier 3	74	0.777	0.781	0.790
	tier 2 + tier 4	49	0.785	0.791	0.797
	tier 3 + tier 4	51	0.775	0.777	0.787
3	tier 1 + tier 2 + tier 3	116	0.779	0.783	0.792
	tier 1 + tier 2 + tier 4	91	<b>0.790</b>	<b>0.796</b>	<b>0.802</b>
	tier 1 + tier 3 + tier 4	93	0.782	0.785	0.795
	tier 2 + tier 3 + tier 4	87	0.784	0.788	0.796
4 (all predictors)	tier 1 + tier 2 + tier 3 + tier 4	129	0.786	0.789	0.798

nations, the pairing of tier 1 and tier 4 yields the best results across all three metrics, followed by the combination of tier 2 and tier 4. When examining models that incorporate three tiers, the combination of tier 1, tier 2, and tier 4 stands out, achieving the highest  $F_1$  score, precision, and recall among all possible feature set combinations, even slightly outperforming the feature set with all features.

Figure 19 shows that including tier 1 (bimonthly reflectance bands), tier 2 (bimonthly indices), and tier 4 (long-term characteristics) improves model performance. Similarly to the SOC regression results, models incorporating long-term trends (tier 4) exhibit the best performance metrics on average, while models including annual features tend to decrease model performance.

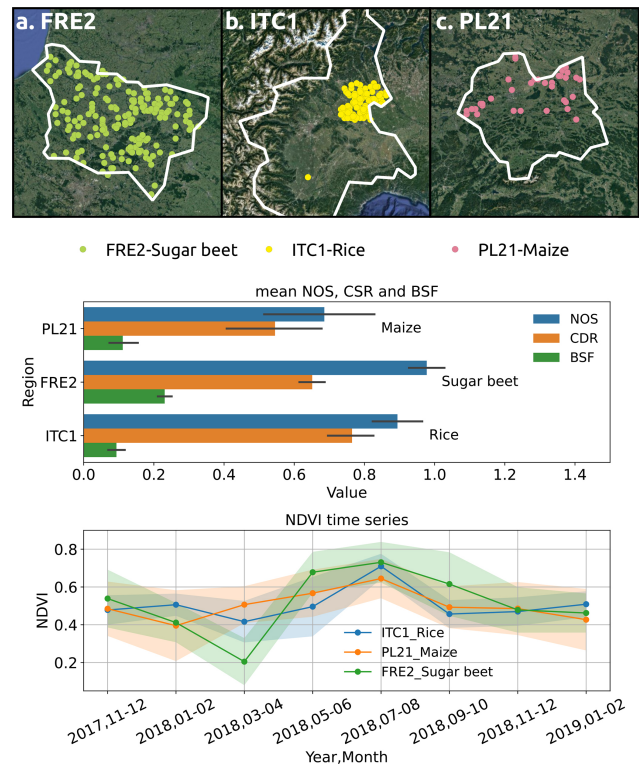




**Figure 10.** Time series comparison showing zoomed-in satellite and BSF images of reforested post-sand mine areas near Bukowno for the years 2000, 2010, and 2020. In the top left are the BSF trend images, where the black frame shows the zoomed-in area of which the satellite and BSF images are shown. The top-left image is the same as in Fig. 9. BSF images generated by the authors using data produced in this study. The composite Landsat ARD2 image was generated by the authors using data from Potapov et al. (2020), as described in Sect. 2.1.

### 3.8.2 Examination across themes

Table 7 presents the performance of the model for different combinations of predictors based on themes. When examining models trained solely on one predictor theme, reflectance band predictors perform the best across all metrics, while crop predictors perform the worst. It is important to note that crop predictors only include two variables: NOS and CDR. For models trained with combinations from two themes, those involving band predictors generally perform much better than others, especially when combined with water and vegetation predictors. In contrast, the inclusion of crop predictors appears to degrade the performance of the model. When three predictor themes are combined, the best performance is found with the combination of band, soil, and water predictors, closely followed by combinations of band, soil, and vegetation, as well as band, water, and vegetation. Models lacking band predictors but including crop predictors perform worse, which is also true for models trained on four themes. Adding crop predictors to models that already include all other predictors does not significantly change the performance of the model. When comparing model perfor-



**Figure 11.** Illustration of the locations of LUCAS 2018 survey points of selected crop types (a–c); statistics of various indices calculated from these LUCAS 2018 points across crops in the selected regions (middle panel); and average NDVI dynamics for the year 2018, including adjacent values from neighboring years (bottom panel). The LUCAS 2018 points data are from three NUTS2 regions: (a) the central part of the Picardy region (FRE2) in north France, (b) the western part of the Piedmont province (ITC1) in Italy, and (c) the Malopolska Province (PL21) in southern Poland. These three regions are close to or partially overlap with the areas shown in Fig. 4; thus, they are denoted by lowercase letters instead of capital ones to show the difference and similarity. Satellite imagery source: © Google Earth (<https://earth.google.com/web/>, last access: 27 June 2024).

mance across all possible theme-based combinations, it is evident that crop predictors do not add value to the model and can even harm performance when insufficient other predictors are available. In contrast, soil, vegetation, and water predictors all enhance the model, and the combination of these three themes provides the most significant improvement in performance.

Figure 20 shows the  $R^2$  comparison of using or not using features from a certain theme as input for the SOC regression model. Including more thematic groups of features almost always brings in added value, in particular in the case of reflectance bands, but not for crop indices. The general performance patterns of the LC classification models, trained on different feature sets, resemble those of the SOC regression models (see Table 8). For LC classification models trained

**Table 7.** Metrics for SOC regression using different combinations of predictors based on themes. The predictor combination yielding the best performance is highlighted in bold.

Number of themes combined	Feature combination	Feature number	CCC	RMSE	$R^2$
1	band	63	0.635	0.652	0.488
	crop	2	0.061	0.898	0.028
	soil	17	0.582	0.690	0.426
	vegetation	27	0.594	0.680	0.443
	water	20	0.591	0.681	0.441
2	band + crop	65	0.636	0.651	0.489
	band + soil	80	0.649	0.642	0.504
	band + vegetation	90	0.651	0.639	0.508
	band + water	83	0.651	0.639	0.508
	crop + soil	19	0.581	0.691	0.425
	crop + vegetation	29	0.593	0.680	0.442
	crop + water	22	0.590	0.682	0.440
	soil + vegetation	44	0.631	0.655	0.483
	soil + water	37	0.626	0.658	0.479
	vegetation + water	47	0.618	0.661	0.474
3	band + crop + soil	82	0.650	0.641	0.505
	band + crop + vegetation	92	0.650	0.640	0.507
	band + crop + water	85	0.651	0.639	0.507
	band + soil + vegetation	107	0.659	0.634	0.516
	band + soil + water	100	0.660	0.634	0.516
	band + vegetation + water	110	0.658	0.634	0.516
	crop + soil + vegetation	46	0.630	0.656	0.482
	crop + soil + water	39	0.625	0.659	0.477
	crop + vegetation + water	49	0.621	0.659	0.476
soil + vegetation + water	64	0.646	0.643	0.502	
4	band + crop + soil + vegetation	109	0.660	0.634	0.516
	band + crop + soil + water	102	0.660	0.633	0.517
	band + crop + vegetation + water	112	0.657	0.634	0.516
	band + soil + vegetation + water	127	<b>0.666</b>	<b>0.628</b>	<b>0.524</b>
	crop + soil + vegetation + water	66	0.645	0.643	0.502
5 (all predictors)	band + crop + soil + vegetation + water	129	<b>0.666</b>	<b>0.628</b>	<b>0.524</b>

on a single theme feature set, crop theme features result in the worst metrics, while reflectance bands yield the best performance. Similarly to SOC regression, two-theme combinations that include band theme features generally perform better. For three-theme combinations, the combination of band, soil, and water themes is the most effective. For models trained on combinations of four themes, the best performance is observed when crop features are excluded. In particular, the model trained on the entire feature space achieved the highest accuracy metrics, with a minimal difference to the model trained on all features except crop indices, specifically using the combination of band, soil, vegetation, and water themes (0.786331 versus 0.786330).

On average, models that incorporate Landsat bands achieved the highest  $F_1$  scores, while those including crop indices had the lowest, as shown in Fig. 21. Including addi-

tional thematic predictors, in addition to crop indices, generally improved model performance.

### 3.8.3 Comparison of feature space between SOC regression and LC classification

Comparing the results of the tier-based and theme-based performance for both SOC regression and LC classification, the improvement in model performance from combining more tiers is less significant than the enhancement seen when increasing the number of combined themes. The contribution of features from different tiers varies between SOC regression and LC classification: including tier-3 and tier-4 features improves the SOC regression model, while tier 3 is the only feature tier that decreases model performance for LC classification. In both experiments, tier 4 demonstrates its impor-

**Table 8.** Metrics for LC classification using different combinations of predictors based on themes. The predictor combination yielding the best performance is highlighted in bold.

Number of themes combined	Feature combination	Feature number	$F_1$ score	Precision	Recall
1	band	63	0.769	0.774	0.784
	crop	2	0.288	0.308	0.458
	soil	17	0.683	0.685	0.700
	vegetation	27	0.741	0.744	0.756
	water	20	0.724	0.727	0.739
2	band + crop	65	0.771	0.776	0.785
	band + soil	80	0.778	0.783	0.792
	band + vegetation	90	0.775	0.780	0.789
	band + water	83	0.777	0.781	0.790
	crop + soil	19	0.683	0.685	0.700
	crop + vegetation	29	0.741	0.745	0.756
	crop + water	22	0.725	0.730	0.741
	soil + vegetation	44	0.762	0.764	0.775
	soil + water	37	0.749	0.753	0.762
	vegetation + water	47	0.769	0.773	0.781
3	band + crop + soil	82	0.778	0.784	0.792
	band + crop + vegetation	92	0.774	0.780	0.788
	band + crop + water	85	0.776	0.781	0.789
	band + soil + vegetation	107	0.781	0.785	0.794
	band + soil + water	100	0.784	0.788	0.797
	band + vegetation + water	110	0.780	0.785	0.793
	crop + soil + vegetation	46	0.760	0.763	0.773
	crop + soil + water	39	0.747	0.751	0.761
	crop + vegetation + water	49	0.769	0.774	0.782
	soil + vegetation + water	64	0.779	0.782	0.791
4	band + crop + soil + vegetation	109	0.781	0.785	0.794
	band + crop + soil + water	102	0.784	0.789	0.797
	band + crop + vegetation + water	112	0.780	0.785	0.793
	band + soil + vegetation + water	127	<b>0.786</b>	<b>0.790</b>	<b>0.799</b>
	crop + soil + vegetation + water	66	0.778	0.782	0.791
5 (all predictors)	band + crop + soil + vegetation + water	129	<b>0.786</b>	<b>0.790</b>	<b>0.799</b>

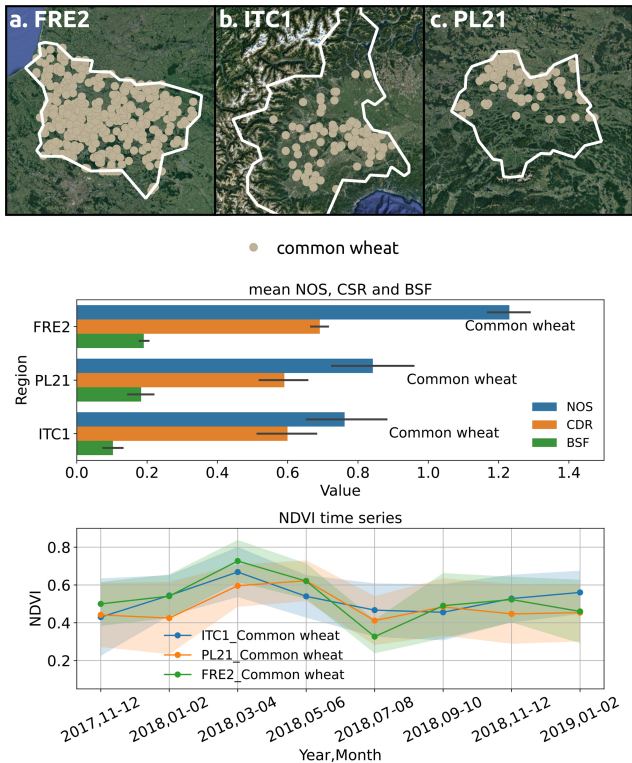
tance. Regarding the feature examination across themes, both experiments show similar results: reflectance bands provide a solid foundation for the models, and different thematic features add valuable information. However, crop indices contribute very little and can even degrade model performance when other variables are insufficient.

## 4 Discussion

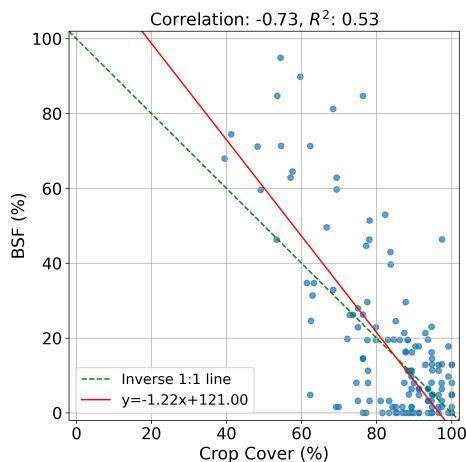
### 4.1 Comparison to other open EO datasets

Our ARCO data cube, with almost 17 TB of data for Europe at a 30 m spatial resolution from 2000 to 2022, represents a significant improvement over existing pan-EU EO datasets in terms of completeness, consistency, and comprehensiveness. Through an automated workflow, we ensured that all data

layers adhere to uniform standards in naming conventions, projection systems, spatial extents, and resolutions. This consistency allows users to easily combine layers over space and time, enabling detailed studies using various indices. Our multidimensional index data cube, optimized as ARCO mosaics to be cloud-free and gap-free, significantly reduces preprocessing efforts compared to using raw data from open data archives such as Sentinel (<https://dataspace.copernicus.eu/explore-data/data-collections/sentinel-data>, last access: 27 June 2024), MODIS (<https://modis.gsfc.nasa.gov/data/dataproduct/>, last access: 27 June 2024), or Landsat (<https://www.usgs.gov/products/data>, last access: 27 June 2024), which are freely provided by agencies like USGS, NASA, and the European Commission's Copernicus and Sentinel programs.



**Figure 12.** Illustration of the locations of LUCAS 2018 survey points of common wheat (a–c); statistics of various indices calculated from these LUCAS 2018 points of common wheat across regions (middle panel); and average NDVI dynamics for the year 2018, including adjacent values from neighboring years (bottom panel). The LUCAS 2018 point data are from the same three NUTS2 regions as in Fig. 11. Satellite imagery source: © Google Earth (<https://earth.google.com/web/>, last access: 27 June 2024).

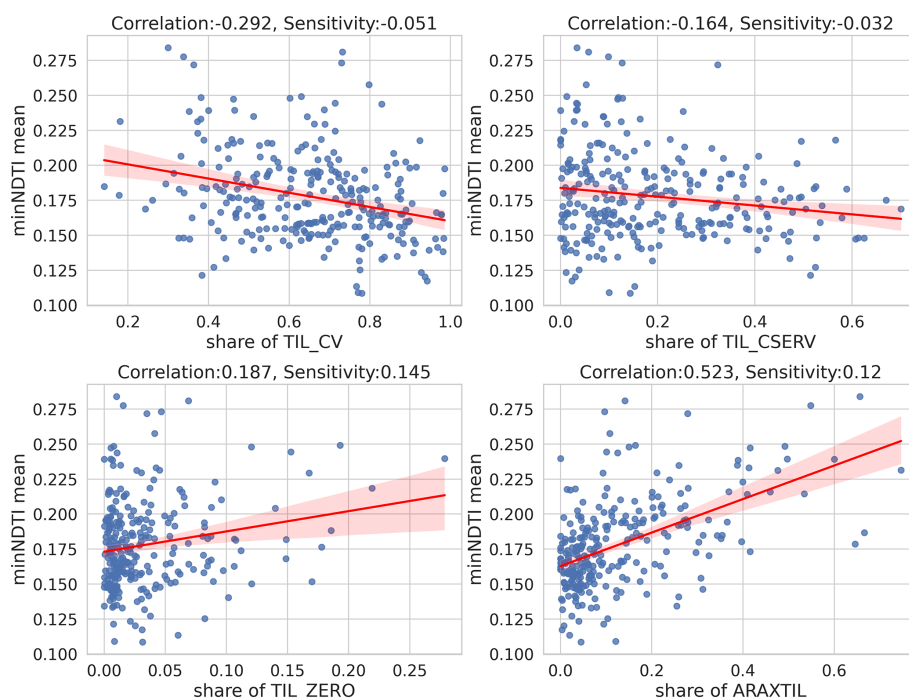


**Figure 13.** Relationship between crop cover proportion and BSF from 2007 to 2016, depicted with an inverse correlation measured by Pearson’s product moment correlation coefficient and coefficient of determination,  $R^2$ .

The concept of ARD has been applied in projects to manage multidimensional data cubes from remote sensing images for specific countries and regions, such as the Australian Geoscience Data Cube (Lewis et al., 2017), Swiss Data Cube (Giuliani et al., 2017), and Africa Regional Data Cube (Kilough, 2019). Although Witjes et al. (2023) offers the first pan-European data cube, it is limited to a quarterly temporal resolution and only cloud-optimized reflectance bands. Our data cube extends this work by refining the temporal resolution and developing four tiers of predictors. In addition to bimonthly indices that capture detailed environmental dynamics, annual aggregations enable yearly comparisons, and long-term characteristics represent general states and trends of vegetation, water, and crop management across Europe. This advancement benefits from two advantages of the Landsat program, the world’s longest continuously running EO mission, and detailed 30 m spatial resolution (Wulder et al., 2022). These indices complement ground survey data, which often lack temporal continuity and are limited in duration due to the significant time and effort required.

Although some datasets provide continuous time series of spectral indices, they are often scattered and inconsistent in format, resolution, and coverage (Wagemann et al., 2021). For instance, most reconstruction efforts focus on NDVI (Zhou et al., 2016), while indices such as NOS, CDR, and BSF are less available (Estel et al., 2015, 2016; Ettehad Osgouei et al., 2019; Dematté et al., 2020). Our data cube’s multidimensional nature provides comprehensive insights into various aspects. The BSF index, for example, helps with crop management by mitigating soil exposure risks and revealing larger land use changes, such as deforestation (Dematté et al., 2020). Combined indices provide a holistic view, e.g., with NOS and CDR, along with BSF, detailing crop cycles and identifying intensively cultivated areas (Estel et al., 2016). Coupled with climate and human footprint data, these indices facilitate studies of the interactions between the climate, human activities, and environments (Brown et al., 2012; Seifert and Lobell, 2015).

Although our Landsat index data cube encompasses a wide range of indices, it is less specialized than studies designed to quantify specific processes. For example, BSF, proposed in previous work (Dematté et al., 2020; Mzid et al., 2021), is mapped here for the first time annually across continental Europe as a complete, gap-free time series and has proven useful in modeling experiments. However, it is not a direct representation of soil surface. In contrast, many studies mapping bare soil focus on identifying the “barest” pixel to create a bare soil composite (BSC) (Diek et al., 2017; Rogge et al., 2018; Safanelli et al., 2020; Dematté et al., 2020; Mzid et al., 2021; Rizzo et al., 2023). These studies explored more sophisticated methods to ensure the algorithm accurately identifies truly bare pixels while effectively filtering out non-soil signals, such as vegetation. Typically, this involves a combination of indices and, in some cases, land use masks (Rogge et al., 2018; Diek et al., 2017; Safanelli et al., 2020;

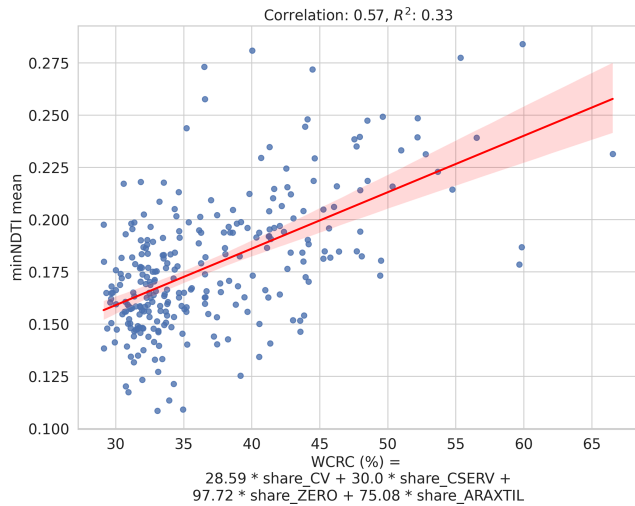


**Figure 14.** Relationship between individual share of tillage practices and average minNDTI of NUTS2 regions recorded in Eurostat. Each point represents a NUTS2 region. Four types of tillage practices shown are TIL\_CV, conventional tillage; TIL\_CSERV, conservation tillage; TIL\_ZERO, zero tillage; and ARAXTIL, land excluding from tillage survey. Coefficient refers to the Pearson correlation coefficient, quantifying the linear correlation between two variables. Sensitivity measures how changes in tillage practices influence minNDTI, obtained as a regression slope coefficient from the OLS model fitting.

Heiden et al., 2022). As a result, these BSC products provide a more direct and reliable representation of soil properties. While its applicability may be limited in regions dominated by permanent vegetation (Hill and Guerschman, 2022), and while it is typically produced as a static snapshot, BSC remains a valuable predictor for digital soil mapping (DSM) (Broeg et al., 2024). Recent studies suggest that combining the strengths of BSF and BSC is feasible. For instance, Zepp et al. (2023) investigated the temporal windows required to generate recurrent BSC products, highlighting the potential to create BSC time series that capture dynamic soil surface properties. Building on this, a natural next step for this study is to apply this approach and generate BSC time series as a new feature, leveraging the extensive Landsat archive at our disposal. Additionally, BSF could be integrated as an indicator to dynamically refine the temporal frame for BSC generation.

Many studies on land surface phenology (LSP), i.e., the descriptors of spatio-temporal patterns of the vegetated land surface derived from EO data (De Beurs and Henebry, 2005), present different methods to quantify crop cycles (Henebry and De Beurs, 2013; Estel et al., 2016; Liu et al., 2020; Meroni et al., 2021). All these methods highlight potential avenues for refining the current CDR and NOS data products in the future. Most LSP studies and products have moderate to coarse spatial resolutions (Ganguly et al., 2010; Es-

tel et al., 2016; Zhang et al., 2018; Caparros-Santiago et al., 2021) due to the trade-off between spatial resolution and revisit frequency (Bolton et al., 2020; Meroni et al., 2021; Caparros-Santiago et al., 2021). Although the Harmonized Landsat Sentinel-2 (HLS) product – created by harmonizing data from Landsat 8/9 and Sentinel-2A/B – offers new opportunities to map LSP indicators at fine spatial resolutions (Bolton et al., 2020), it is only available starting from 2013. Lastly, despite the potential to quantify the intensity of tillage with EO data such as Landsat 7/8/9 and Sentinel-2, the availability of information on tillage practice on a continental scale remains limited. In summary, the Landsat indices we described in this paper can be seen as part of a comprehensive effort to produce useful biophysical indices to support soil, land, water, and land use monitoring. While our methods may be less specialized than those of dedicated thematic studies, this reflects a compromise to balance computational resource demands and the need for complete, gap-free coverage. Despite this, they are designed as a foundational resource, offering a broad and adaptable set of predictors in a comprehensive, gap-free form that requires minimal effort from users for effective application in various environmental modeling and monitoring tasks.



**Figure 15.** Relationship between weighted crop residue cover (WCRC) and mean minNDTI values of NUTS2 regions recorded in Eurostat. Each point stand for a NUTS2 region. Below the figure, the equation for calculating WCRC is presented, where the coefficients correspond to the fitted CRC for each type of tillage practice. Four types of tillage practices considered are TIL\_CV, conventional tillage; TIL\_CSERV, conservation tillage; TIL\_ZERO, zero tillage; and ARAXTIL, land excluding from tillage survey.

#### 4.2 Implications from plausibility checks

Our quality assessment primarily focused on third- and fourth-tier predictors, such as minNDTI, BSF, NOS, and CDR, and their long-term trends due to their complexity and the availability of relevant ground data. We also visualized the annual P50 of NDVI and NDWI to provide basic information on general vegetation and water conditions. Visual examination indicated that annual NDVI, NDWI, and BSF effectively differentiate various landscapes, while NOS and CDR distinguish between different crop patches. Although interpreting these indices without detailed crop management information is challenging, regional statistics of NOS, CDR, and BSF reveal significant differences in crop cycles among different crop types and regions across Europe. We demonstrated that the BSF index effectively measures cropland bareness, with a high negative correlation coefficient of  $-0.73$  with the duration of crop cover in agricultural areas between 2007 and 2016. However, the moderate correlation of minNDTI with tillage practices statistics from EUROSTAT suggests the need for further investigation.

In Fig. 5, a negative NDVI trend is observed in regions characterized by high altitude and latitude. This finding appears to be in conflict with the greening phenomenon in cold/snowy ecosystems, as reported by Obuchowicz et al. (2023). Furthermore, the positive trend observed in the NDWI trend map is inconsistent with Poussin et al. (2021)'s observation of a slight decline in NDWI values in the Alps

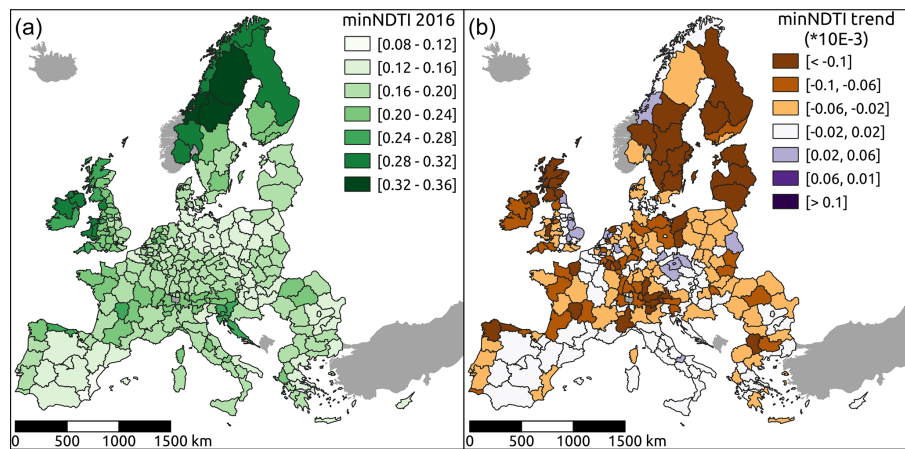
region of Switzerland during the period from 1984 to 2019. Several factors could be influencing these discrepancies, such as the different temporal scales used for trend calculation, as noted by de Jong and de Bruin (2012), or the increased dependence on gap filling due to limited available high-quality EO data in these regions. These contrasting observations call for further investigation to understand the underlying causes and implications of them.

Although these efforts are limited by the availability of relevant ground data and sometimes yield inconclusive results, it is important to note that these indices are proxies for ground objects and processes rather than exact measurements of ground conditions. They are designed to complement, not replace, ground survey data. Therefore, plausibility checks are not expected to provide one-to-one validation but to highlight the advantages and limitations of these indices. In summary, well-established indices such as NDVI and NDWI and the relatively straightforward BSF index are reliable and suitable for direct analysis, as demonstrated in Sect. 3.3. However, for more complex indices like NOS and CDR, robust assessments require careful integration with ground-truth data to ensure accuracy and validity.

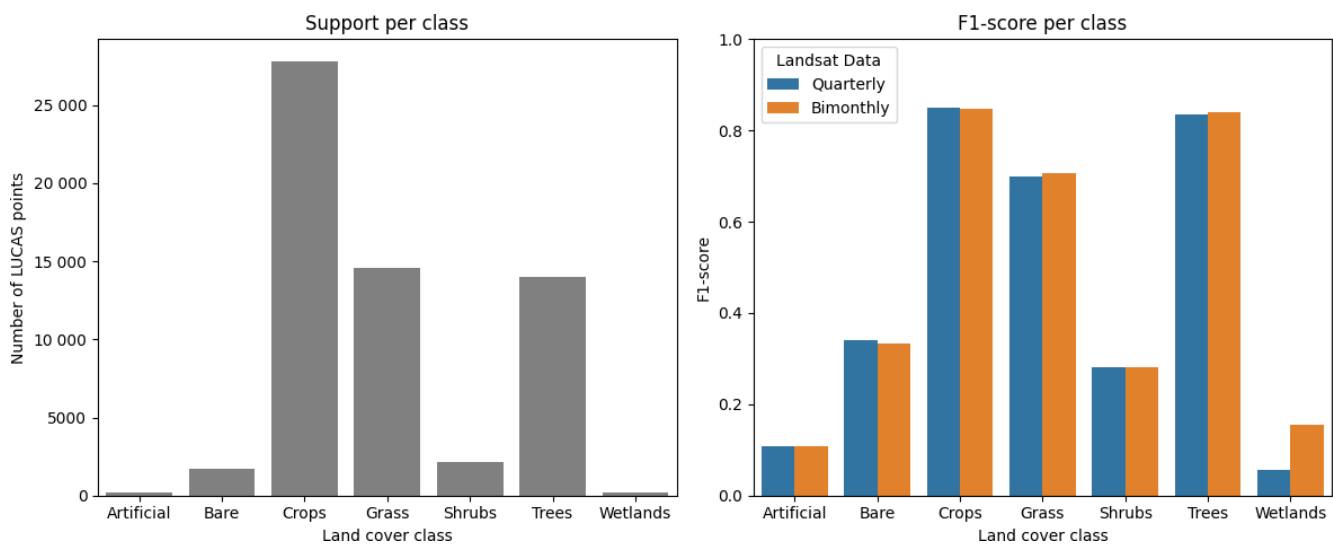
#### 4.3 Insights from modeling experiments

By looking at predictors of different themes' impact on SOC regression and LC classification model, band theme predictors are vital for both models. This is unsurprising as they form the foundation from which other spectral indices are derived and provide a general representation of surface objects and processes (Zhou et al., 2021). Apart from band theme predictors, models trained on single-theme predictors show that vegetation predictors have a stronger positive impact on both LC and SOC modeling than other themes. Vegetation is fundamental in defining land covers, naturally ranking high in feature importance. Its close relationship with living soil through physical, chemical, and biological interactions (Kooch et al., 2022) positions it as a vital "organism" factor in the soil formation process and a key driver of SOC change (McBratney et al., 2003; Wiesmeier et al., 2019). This importance is reflected in numerous SOC modeling studies, where vegetation indices like NDVI and EVI (enhanced vegetation index) consistently rank among the top predictors and are sometimes even used as the only spectral index (Gomes et al., 2019; Sothe et al., 2022; Zeraatpisheh et al., 2022; Wadoux et al., 2020).

Soil and water theme parameters also contribute valuable insights for SOC modeling. The indices of these two themes already differ statistically across most LCs even before considering LC classification (Szabo et al., 2016). For SOC modeling, tillage practices and soil exposure, as indicated by soil theme predictors, negatively alter soil structure and impact carbon dynamics (Chahal et al., 2021; T. Yang et al., 2020; Demattê et al., 2020; Mzid et al., 2021; Wiesmeier et al., 2019). Similarly, water dynamics and climate, represented by



**Figure 16.** Average minNDTI values in the year 2016 (a) and average minNDTI trend values (b) by NUTS2 regions.

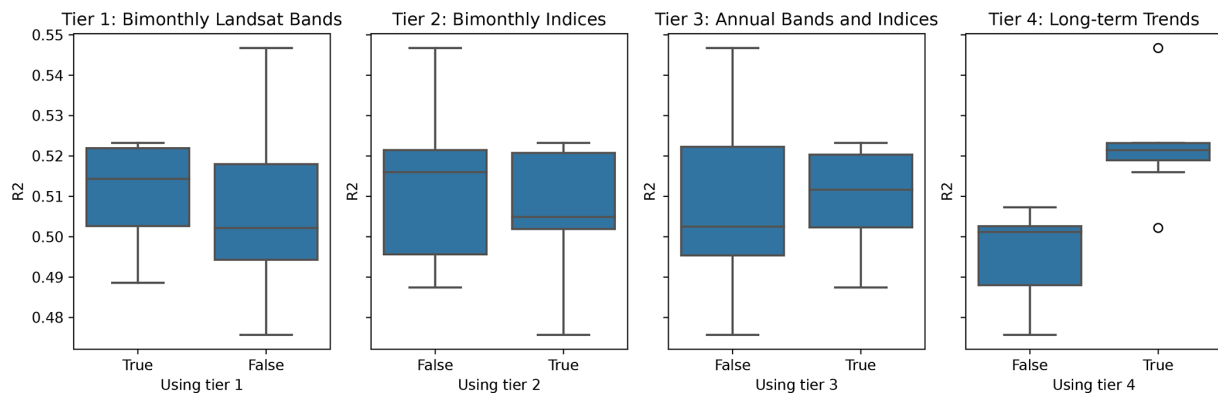


**Figure 17.** Comparison between classification results between Landsat data aggregated to four quarters and three percentiles per quarter (EcoDataCube; Witjes et al., 2023) and the data from this work, aggregated to six bimonthly variables per band, as well as three annual percentiles. Note that the overall accuracy was similar, except for the wetlands class, which was classified more accurately by the model trained on the bimonthly and annual data.

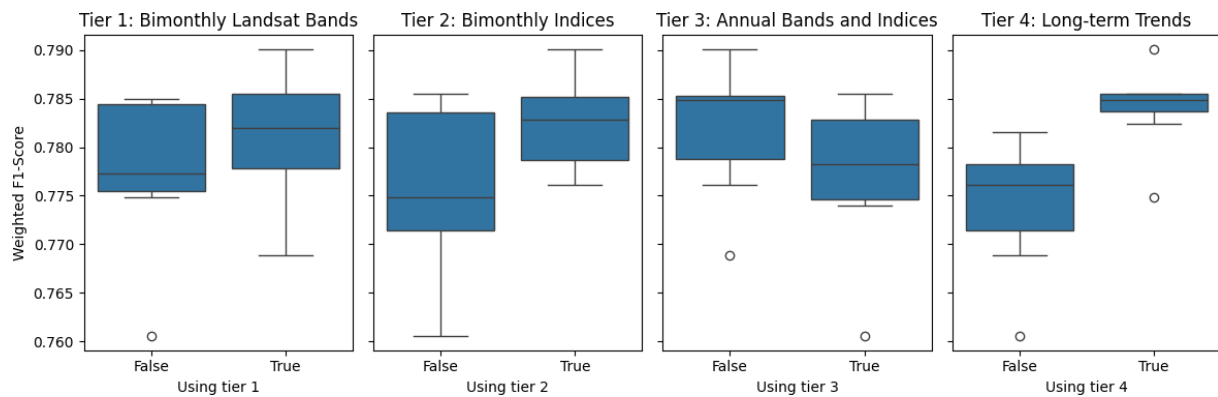
water theme predictors, are key elements of the climate factor in the soil-forming scorpan model (Maynard and Levi, 2017; Wiesmeier et al., 2019; Zeraatpisheh et al., 2022). When direct climate data, such as precipitation and temperature, are unavailable, water theme indices can serve as useful substitutes.

Combining predictors from multiple themes generally enhances model performance. This is consistent with the findings of Witjes et al. (2023) for LC classification. For SOC modeling, most studies in DSM ensure that all soil-forming factors are incorporated within the corpan model (McBratney et al., 2003; Wiesmeier et al., 2019; Wadoux et al., 2020). Zeraatpisheh et al. (2022) demonstrated that while combining feature groups may not always result in higher prediction accuracy, it does produce SOC predictions with lower uncer-

tainty. The exception is crop theme predictors, likely due to limitations in our methods for quantifying NOS and CDR. Numerous studies have shown that phenological parameters can improve soil property and LC modeling, at least for croplands (Maynard and Levi, 2017; L. Yang et al., 2020; He et al., 2021; Xue et al., 2014; Nguyen et al., 2020). However, these studies often focus on small areas with much less data, allowing for localized and detailed extraction of phenological parameters. In contrast, our approach to extracting phenological parameters like NOS and CDR, and assessing their relevance to SOC and LC modeling, remains limited. Further investigation is needed to determine whether phenological parameters can be effectively scaled to large, continental-level studies and, if so, to determine the methods for their application.



**Figure 18.** Accuracy comparison of SOC regression models trained on 15 combinations of four tiers of predictors.



**Figure 19.** Accuracy comparison of LC classification models trained on 15 combinations of four tiers of predictors.

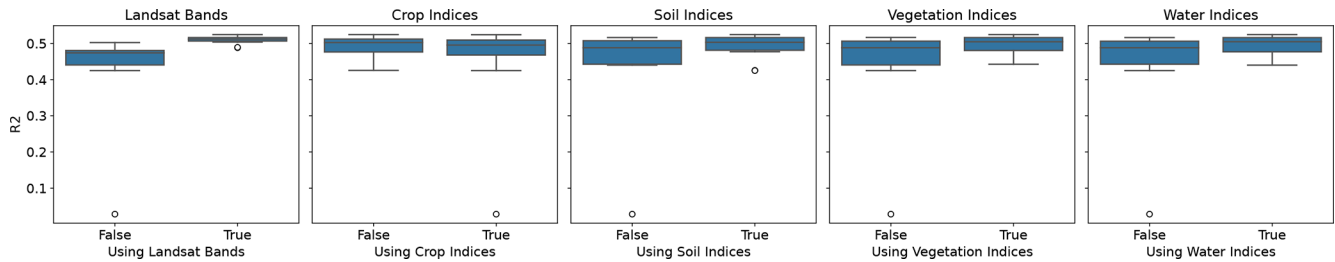
Unlike theme-based predictors, predictors of different tiers exhibit varying levels of importance for SOC regression and LC classification modeling. For SOC regression, predictors that represent longer periods are more important, aligning with the findings of Zepp et al. (2023). This is clear in Fig. 18, where predictors with extended temporal processing significantly enhance model performance. Table 5 shows that model trained solely on tier-4 predictors (long-term characteristics) consistently outperform others across all three metrics. This suggests that long-term characteristics are highly valuable for SOC regression and can sometimes outperform more complex combinations of features. A study by Yang et al. (2022) demonstrates that temporal variables are effective in modeling temporal SOC variation, though their scope spans 30 years, while the LUCAS topsoil survey covers only 9 years. Dynamic information appears less beneficial for SOC modeling, which is not surprising since SOC is a slow-changing soil property that typically varies over decades.

In contrast, incorporating detailed dynamic information (tier 1 and tier 2: bimonthly bands and indices) improves the model performance for LC classification. This improvement may be attributed to the dynamic nature of LC, which changes more frequently, or suddenly compared to SOC, making it better captured with more frequent revisit infor-

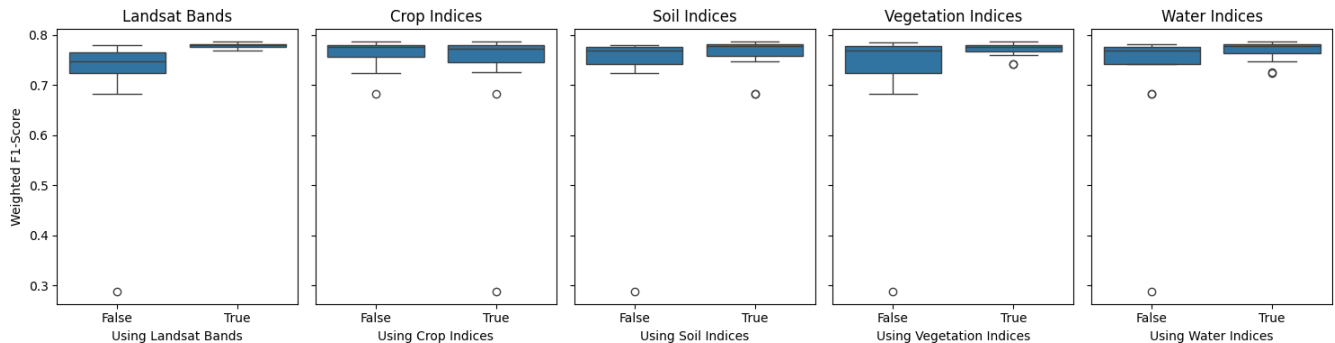
mation. The temporal dimension has always been critical in LC classification. Researchers have emphasized incorporating temporal features into models to enable the mapping of LC at finer temporal scales, such as annual intervals (Gómez et al., 2016; Zhang et al., 2021; Brown et al., 2022). Surprisingly, the tier-3 (annual) indices are shown to be less important for LC classification. This could be because the temporal characteristics within a year neither provide the detailed phenological information that bimonthly predictors do (as phenological processes are not defined strictly by a calendar year) nor offer the long-term characteristics that tier-4 predictors provide.

Combining predictors from multiple themes adds more value than combining predictors from multiple tiers. This is likely because different tiers mainly vary in their level of temporal feature engineering, conveying similar messages. However, more predictors do not always lead to better performance. However, as demonstrated by our SOC regression and LC classification experiments, using more features does not always lead to better performance. This can increase the risk of multicollinearity and overfitting, add complexity, and require greater computational resources (Wadoux et al., 2020). Feature importance analysis highlights that feature selection should be guided by a clear understanding of the pre-





**Figure 20.** Accuracy comparison of SOC regression models trained on 31 combinations of five themes of predictors.



**Figure 21.** Accuracy comparison of LC classification models trained on 31 combinations of five themes of predictors.

dictors' characteristics, modeling objectives, and task constraints. Factors such as the temporal and spatial resolution of covariates, the spatial extent of modeling, and the balance of sample data also affect the determination of the appropriate features (Chen et al., 2022; Shi et al., 2018; Nazari et al., 2020).

#### 4.4 Limitations

As with any product derived from the Landsat ARD V2, users should be aware of inherent limitations. A key step in harmonizing Landsat products into ARD is surface reflectance normalization (Potapov et al., 2020), which differs from physically based atmospheric and surface anisotropy corrections. Our plausibility checks demonstrate that the bi-monthly Landsat reflectance bands (tier-1) are well suited for index derivation and modeling purposes, indicating sufficient product quality. However, similar to the Landsat ARD V2 data, this data cube may not be suitable for very precise studies of land surface reflectance.

As noted by Potapov et al. (2020), the GLAD Landsat ARD algorithm struggles with wintertime image processing in temperate and boreal climates at high latitudes. This is due to low Sun azimuth angles and the similarity between snow cover and clouds during the winter season. As a result, some images and their resulting 16 d composites may still contain snow-covered observations. To address this issue, additional filtering – such as applying the “snow/ice” observation quality flag – was implemented in this study to remove affected observations. However, this process resulted in larger data

gaps in certain regions. Consequently, the TSIRF algorithm, which depends on other images in the sequence for reconstruction, can produce inaccurate results when the time series contains too few valid pixels (Consoli et al., 2024). This is also reflected in the reduced time series reconstruction accuracy of the European dataset compared to the global dataset, which can be attributed to Europe's higher average latitude compared to the global dataset. Consequently, artifacts are observed in high-altitude and high-latitude regions, such as central Sweden and southern Norway, as illustrated in Figs. 5 and 6. While these artifacts do not significantly impact the overall analysis, users should be aware of them and exercise caution when interpreting results, especially in regions with challenging data conditions.

As with many land surface phenology (LSP) studies, our temporally aggregated indices such as BSF, NOS, and CDR rely on threshold-based methods (De Beurs and Henebry, 2010). These thresholds can vary depending on crop types and regional characteristics, often requiring expert knowledge for accurate application, in particular crop-relevant indices (Chen et al., 2016). Given the broad spatial and temporal scope of our data cube, we adopted a generalized threshold approach, which might not capture local variations accurately (Henebry and De Beurs, 2013; Estel et al., 2016; Demattè et al., 2020). The limitation of this method lies in its inability to distinguish between meaningful changes and natural fluctuations exceeding the threshold (De Beurs and Henebry, 2010). Therefore, users are advised to proceed cautiously and conduct preliminary tests when a meaningful

conclusion needs to be drawn from direct analysis with this limited local or domain knowledge.

Another challenge posed by advances in EO technology is the expanding list of satellite spectral indices, each uniquely calibrated for specific environmental attributes. For example, the awesome spectral index (ASI), developed by Montero et al. (2023), lists 127 vegetation indices by its publication date. Incorporating all of these indices into a single data cube is impractical. To balance data volume and complexity, we adopted a hybrid approach. The predictor data cube includes the indices most commonly applied, covering several important aspects of the environment, such as vegetation health, soil conditions, and water dynamics. Additionally, developed in alignment with the principles of the Open Data Cube (ODC), our ARCO data cube includes guided documentation and scripts to assist users in querying, exploring, performing analyses, and tracking the provenance of all contained data. This approach facilitates quality control and updates, allowing users to customize the data cube according to their specific needs or to explore new indices that could offer additional insight into the environment.

The data cube includes a wide array of indices with various applications. However, some indices may yield irrelevant values outside their intended regions. For example, many indices lose significance in snowy environments at high latitudes and altitudes, and minNDTI does not convey useful information about tillage practices outside of arable land. It is crucial not to apply these indices indiscriminately across all land covers. Other indices within the data cube can help delineate appropriate application areas. For example, the NDSI can exclude snowy areas, ensuring a more accurate application of other indices.

#### 4.5 Future work

While NOS and CDR provide useful insights into crop cycles and their duration, their limited contribution to SOC regression models suggests a need for further refinement. Improving the methods for quantifying these indices or exploring alternative indicators might enhance their utility in SOC modeling. As discussed in Sect. 4.1, there is significant value in learning from specialized studies that focus on accurately quantifying specific processes. Future updates to this work will incorporate more localized and domain-specific knowledge could help to tailor these indices to better capture relevant environmental processes.

Although the original Landsat archive has a 16 d temporal revisit, we aggregated the dataset into a bimonthly resolution to address quality issues such as gaps and cloudy pixels. Although bimonthly resolution effectively captures annual dynamics, a finer 16 d resolution could provide richer details, especially for NDTI values observed from harvest windows crucial for mapping tillage practices (Zheng et al., 2012). However, finer temporal resolution requires extensive

gap filling, presenting a significant trade-off between temporal detail and data consistency.

Our dataset currently relies exclusively on Landsat ARD V2 to derive NOS, CDR, and minNDTI. However, integrating multiple sensors could enhance accuracy (Henebry and De Beurs, 2013; Caparros-Santiago et al., 2021). For example, using minimum surface temperature thresholds can prevent false crop cycle identifications (Li et al., 2014), and integrating precipitation data can improve crop residue estimates (Beeson et al., 2020; Dvorakova et al., 2023). Furthermore, combining Harmonized Landsat Sentinel-2 (HLS) data could further improve index usability, especially for years covered by Sentinel-2 (Wu et al., 2022).

As indicated by the survey results of Wagemann et al. (2021), improving the usability of spectral indices requires the participation of various stakeholders, including farmers, policy-makers, and sector-specific experts. Their local knowledge and domain-specific insight can refine algorithms and improve the utility of data. This approach aligns with the objectives of the Horizon Europe AI4SoilHealth project (2023–2027), aiming to bridge the gap between EO data and practical applications.

## 5 Data availability

This dataset is registered with a valid DOI (<https://doi.org/10.5281/zenodo.10776891>; Tian et al., 2024a). While each data bucket has a unique DOI, we kindly ask users to cite the DOI on the Zenodo landing page to ensure consistency. This DOI represents all dataset versions and always resolves to the latest release.

Due to Zenodo's storage limitations and the large size of the dataset, only portions of the data are stored in Zenodo, distributed across multiple buckets. To streamline access, we have created a Zenodo landing page that includes a navigation catalog linking to the individual buckets. Additionally, a catalog ([https://docs.google.com/spreadsheets/d/1QTA6OkkYIZlJfHst\\_inCrkC7DJcMAyHnM9k0iHulwpg/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1QTA6OkkYIZlJfHst_inCrkC7DJcMAyHnM9k0iHulwpg/edit?usp=sharing), Google Sheet link, 2024), provided as a Google Sheet, offers complete links for public access to the data stored in Wasabi Cloud Storage. Both the Zenodo landing page and the central catalog include a detailed explanation of the naming conventions used.

To lower the barrier to data use, metadata are embedded directly into file names. Each dataset layer follows a standardized naming format, structured into 10 key fields: generic variable name, variable procedure combination, position in the probability distribution or variable type, spatial support, depth reference, time reference (including start and end times), bounding box, EPSG code, and version code. Each metadata field serves a specific purpose in assessing the datasets' fitness for use. For the datasets presented in this study, the metadata specify a uniform spatial support

of 30 m resolution, a depth reference denoted by the letter *s* (land surface), a bounding box identified as *eu*, and an EPSG code of EPSG:3035. For the other fields, the generic variable name helps users identify the required predictor layer; the variable procedure combination provides indications of how the data were derived and their source; the time reference, comprising the start and end dates, ensures users can select layers matching their relevant temporal scope; and the version code, which reflects the creation date of the corresponding layer, facilitates tracking and version control. This metadata framework enhances geodata trustworthiness and usability by allowing users to assess dataset relevance and suitability directly from the name.

## 6 Code availability

Complete code used to derive all indices and generate the visualization figures is available at <https://doi.org/10.5281/zenodo.12907281> (Tian et al., 2024b). Even more spectral indices can be derived using the library provided by Montero et al. (2023) (<https://github.com/awesome-spectral-indices/>, last access: 27 June 2024).

## 7 Conclusions

In this paper, we present the production process and quality assessment of a 17 TB ARCO data cube, primarily based on the Landsat indices. It is available at a 30 m resolution from 2000 to 2022, comprising multiple spectral indices. Multiple tiers of predictors are derived to represent different levels of temporal information extraction: bimonthly, annual, and long-term analyses covering continental Europe, including Ukraine, the UK, and Türkiye. The main advantages of our dataset include (1) its fine spatial resolution and extensive temporal coverage, which are crucial for analyzing and tracking the spatial variations and temporal dynamics of different environmental aspects, such as soil; (2) its broad coverage of indices, each providing unique insights into different aspects, together offering a holistic view of environment; and (3) the fact that it is analysis-ready for modeling and mapping applications.

To maximize the effectiveness and reliability of the data, the following usage recommendations are proposed:

- *Established indices.* For well-established indices like NDVI, NDWI, SAVI, FAPAR, NDSI, and those demonstrating high reliability such as BSF, the data cube provides valuable information and insights, especially when ground-truth data are lacking.
- *Complex indices.* For more complex indices that were not validated in this study due to the lack of corresponding ground-truth data, caution is advised. These should be used alongside ground-truth data to ensure accuracy

and validity, and their application should consider domain and local knowledge.

- *Modeling.* All indices in the data cube have proven effective as predictors for modeling purposes, making them useful for various environmental and ecological models.
- *Feature selection.* It is beneficial to perform feature selection to identify the best combinations of indices from the comprehensive predictor sets for specific modeling tasks, optimizing the model's performance and relevance to the study's objectives.

**Author contributions.** XT and TH conceptualized the study. LP, DC, XT, and YH curated the data. Funding was secured by TH and LP. XT and FS conducted the investigation. Methodology was developed by XT, LP, and DC. Project administration was handled by TH and RM. Resources were provided by TH and LP. Software development was performed by LP, DC, XT, and YH. Supervision was carried out by TH, LP, FS, and RM. XT, MW, and FS validated the results. Visualization was done by MŞ, XT, and MW. The original draft was prepared by XT, TH, MW, and DC, with review and editing contributions from TH, MW, and YH.

**Competing interests.** Xuemeng Tian, Davide Consoli, Martijn Witjes, Leandro Parente, Murat Şahin, Yu-Feng Ho, Robert Miňárik, and Tomislav Hengl are employed by OpenGeoHub.

**Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

**Acknowledgements.** The authors are grateful to Sytze de Bruin (Wageningen University & Research), Kirsten de Beurs (Wageningen University & Research), and Bernhard Ahrens (Max Planck Institute for Biogeochemistry) for their valuable advice on the structure, visualization, and phrasing of this paper. The authors are especially grateful to Calogero Schillaci (European Commission Joint Research Center) for providing constructive feedback on the first version of the manuscript. The authors are grateful to Peter Potapov, codirector of the Global Land Analysis and Discovery (GLAD) group and the whole GLAD group for providing assistance with the Landsat ARD 16 d composites.

**Financial support.** This research was funded by the AI4SoilHealth project (EU Horizon Europe research and innovation program; grant agreement no. 101086179) and the Open-Earth-Monitor Cyberinfrastructure project (EU Horizon Europe research and innovation program; grant agreement no. 101059548).

**Review statement.** This paper was edited by Giulio G. R. Iovine and reviewed by three anonymous referees.

## References

- Baumann, P.: On the analysis-readiness of spatio-temporal earth data and suggestions for its enhancement, *Environ. Model. Softw.*, 176, 106017, <https://doi.org/10.1016/j.envsoft.2024.106017>, 2024.
- Beeson, P. C., Daughtry, C. S., and Wallander, S. A.: Estimates of conservation tillage practices using landsat archive, *Remote Sens.*, 12, 2665, <https://doi.org/10.3390/rs12162665>, 2020.
- Bolton, D. K., Gray, J. M., Melaas, E. K., Moon, M., Eklundh, L., and Friedl, M. A.: Continental-scale land surface phenology from harmonized Landsat8 and Sentinel-2 imagery, *Remote Sens. Environ.*, 240, 111685, <https://doi.org/10.1016/j.rse.2020.111685>, 2020.
- Broeg, T., Don, A., Gocht, A., Scholten, T., Taghizadeh-Mehrjardi, R., and Erasmí, S.: Using local ensemble models and Landsat bare soil composites for large-scale soil organic carbon maps in cropland, *Geoderma*, 444, 116850, <https://doi.org/10.1016/j.geoderma.2024.116850>, 2024.
- Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., et al.: Dynamic World, Near real-time global 10 m land use land cover mapping, *Sci. Data*, 9, 251, <https://doi.org/10.1038/s41597-022-01307-4>, 2022.
- Brown, M., De Beurs, K., and Marshall, M.: Global phenological response to climate change in crop areas using satellite remote sensing of vegetation, humidity and temperature over 26 years, *Remote Sens. Environ.*, 126, 174–183, <https://doi.org/10.1016/j.rse.2012.08.009>, 2012.
- Caparros-Santiago, J. A., Rodriguez-Galiano, V., and Dash, J.: Land surface phenology as indicator of global terrestrial ecosystem dynamics: A systematic review, *ISPRS J. Photogram. Remote Sens.*, 171, 330–347, <https://doi.org/10.1016/j.isprsjprs.2020.11.019>, 2021.
- Castaldi, F., Chabrillat, S., Don, A., and van Wesemael, B.: Soil organic carbon mapping using LUCAS topsoil database and Sentinel-2 data: An approach to reduce soil moisture and crop residue effects, *Remote Sens.*, 11, 2121, <https://doi.org/10.3390/rs11182121>, 2019.
- Chahal, I., Hooker, D., Deen, B., Janovicek, K., and Van Eerd, L.: Long-term effects of crop rotation, tillage, and fertilizer nitrogen on soil health indicators and crop productivity in a temperate climate, *Soil Till. Res.*, 213, 105121, <https://doi.org/10.1016/j.still.2021.105121>, 2021.
- Chatenoux, B., Richard, J.-P., Small, D., Roeoesli, C., Wingate, V., Poussin, C., Rodila, D., Peduzzi, P., Steinmeier, C., Ginzler, C., Psomas, A., Schaeppman, M. E., and Giuliani, G.: The Swiss data cube, analysis ready data archive using earth observations of Switzerland, *Sci. Data*, 8, 295, <https://doi.org/10.1038/s41597-021-01076-6>, 2021.
- Chen, Y., Song, X., Wang, S., Huang, J., and Mansaray, L. R.: Impacts of spatial heterogeneity on crop area mapping in Canada using MODIS data, *ISPRS J. Photogram. Remote Sens.*, 119, 451–461, <https://doi.org/10.1016/j.isprsjprs.2016.07.007>, 2016.
- Chen, Y., Ma, L., Yu, D., Zhang, H., Feng, K., Wang, X., and Song, J.: Comparison of feature selection methods for mapping soil organic matter in subtropical restored forests, *Ecol. Indic.*, 135, 108545, <https://doi.org/10.1016/j.ecolind.2022.108545>, 2022.
- Consoli, D., Parente, L., Simoes, R., Şahin, M., Tian, X., Wijes, M., Sloat, L., and Hengl, T.: A computational framework for processing time-series of Earth Observation data based on discrete convolution: global-scale historical Landsat cloud-free aggregates at 30 m spatial resolution, *Peer J.*, 12, e18585, <https://doi.org/10.7717/peerj.18585>, 2024.
- d’Andrimont, R., Yordanov, M., Martinez-Sanchez, L., Eiselt, B., Palmieri, A., Dominici, P., Gallego, J., Reuter, H. I., Jobges, C., Lemoine, G., and van der Velde, M.: Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union, *Sci. Data*, 7, 352, <https://doi.org/10.6084/m9.figshare.12841202>, 2020.
- d’Andrimont, R., Verhegghen, A., Lemoine, G., Kempeneers, P., Meroni, M., and Van der Velde, M.: 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>, 2021.
- Daughtry, C. S., Serbin, G., Reeves III, J. B., Doraiswamy, P. C., and Hunt Jr., E. R.: Spectral reflectance of wheat residue during decomposition and remotely sensed estimates of residue cover, *Remote Sens.*, 2, 416–431, <https://doi.org/10.3390/rs2020416>, 2010.
- De Beurs, K. M. and Henebry, G. M.: Land surface phenology and temperature variation in the International Geosphere–Biosphere Program high-latitude transects, *Global Change Biol.*, 11, 779–790, <https://doi.org/10.1111/j.1365-2486.2005.00949.x>, 2005.
- De Beurs, K. M. and Henebry, G. M.: Spatio-Temporal Statistical Methods for Modelling Land Surface Phenology, in: *Phenological Research*, edited by: Hudson, I. and Keatley, M., Springer, Dordrecht, [https://doi.org/10.1007/978-90-481-3335-2\\_9](https://doi.org/10.1007/978-90-481-3335-2_9), 2010.
- de Jong, R. and de Bruin, S.: Linear trends in seasonal vegetation time series and the modifiable temporal unit problem, *Biogeosciences*, 9, 71–77, <https://doi.org/10.5194/bg-9-71-2012>, 2012.
- Demattê, J. A., Safanelli, J. L., Poppiel, R. R., Rizzo, R., Silvero, N. E. Q., Mendes, W. D. S., Bonfatti, B. R., Dotto, A. C., Salazar, D. F. U., Mello, F. A. D. O., da Silveira Paiva, A. F., Barros Souza, A., dos Santos, N. V., Nascimento, C. M., de Mello, D. C., Belinaso, H., Neto, L. G., Amorim, M. T. A., de Resende, M. E. B., da Souza Vieira, J., de Queiroz, L. G., Gallo, B. C., Sayão, V. M., and da Silva Lisboa, C. J.: Bare earth’s surface spectra as a proxy for soil resource monitoring, *Sci. Rep.*, 10, 4461, <https://doi.org/10.1038/s41598-020-61408-1>, 2020.
- Diek, S., Fornallaz, F., Schaeppman, M. E., and De Jong, R.: Barest pixel composite for agricultural areas using landsat time series, *Remote Sens.*, 9, 1245, <https://doi.org/10.3390/rs9121245>, 2017.
- Dvorakova, K., Heiden, U., Pepers, K., Staats, G., van Os, G., and van Wesemael, B.: Improving soil organic carbon predictions from a Sentinel-2 soil composite by assessing surface conditions and uncertainties, *Geoderma*, 429, 116128, <https://doi.org/10.1016/j.geoderma.2022.116128>, 2023.
- Edlinger, A., Garland, G., Banerjee, S., Degrunne, F., García-Palacios, P., Hallin, S., Herzog, C., Jansa, J., Kost, E., Maestre, F. T., Sánchez-Pescador, D., Philippot, L., Rillig, M., Romdhane, S., Saghai, A., Spor, A., Frossard, E., van der Heijden, M., Hartman, K., and Held, A.: Edlinger\_et\_al\_database.xlsx, Figshare [data set], <https://doi.org/10.6084/m9.figshare.15134328.v4>, 2022.

- Edlinger, A., Garland, G., Banerjee, S., Degrune, F., García-Palacios, P., Herzog, C., Sánchez Pescador, D., Romdhane, S., Ryo, M., Saghaï, A., Hallin, S., Maestre, F. T., Philippot, L., Rillig, M., and van der Heijden, M.: The impact of agricultural management on soil aggregation and carbon storage is regulated by climate. Digging deeper Dataset, Figshare [data set], <https://doi.org/10.6084/m9.figshare.19762114.v5>, 2023.
- Eskandari, I., Navid, H., and Rangzan, K.: Evaluating spectral indices for determining conservation and conventional tillage systems in a vetch-wheat rotation, *Int. Soil Water Conserv. Res.*, 4, 93–98, <https://doi.org/10.1016/j.iswcr.2016.04.002>, 2016.
- Estel, S., Kuemmerle, T., Alcántara, C., Levers, C., Prishchepov, A., and Hostert, P.: Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series, *Remote Sens. Environ.*, 163, 312–325, <https://doi.org/10.1016/j.rse.2015.03.028>, 2015.
- Estel, S., Kuemmerle, T., Levers, C., Baumann, M., and Hostert, P.: Mapping cropland-use intensity across Europe using MODIS NDVI time series, *Environ. Res. Lett.*, 11, 024015, <https://doi.org/10.1088/1748-9326/11/2/024015>, 2016.
- Ettehadi Osgouei, P., Kaya, S., Sertel, E., and Alganci, U.: Separating built-up areas from bare land in mediterranean cities using Sentinel-2A imagery, *Remote Sens.*, 11, 345, <https://doi.org/10.3390/rs11030345>, 2019.
- Eurostat: Agricultural practices (ef\_mp\_prac) [Data set], Eurostat [data set], [https://doi.org/10.2908/EF\\_MP\\_PRAC](https://doi.org/10.2908/EF_MP_PRAC), 2024.
- Ganguly, S., Friedl, M. A., Tan, B., Zhang, X., and Verma, M.: Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product, *Remote Sens. Environ.*, 114, 1805–1816, <https://doi.org/10.1016/j.rse.2010.04.005>, 2010.
- Gao, B.-C.: NDWI – A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sens. Environ.*, 58, 257–266, [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3), 1996.
- Giuliani, G., Chatenoux, B., De Bono, A., Rodila, D., Richard, J.-P., Allenbach, K., Dao, H., and Peduzzi, P.: Building an earth observations data cube: lessons learned from the swiss data cube (sdc) on generating analysis ready data (ard), *Big Earth Data*, 1, 100–117, <https://doi.org/10.1080/20964471.2017.1398903>, 2017.
- Giuliani, G., Cazeaux, H., Burgi, P.-Y., Poussin, C., Richard, J.-P., and Chatenoux, B.: SwissEnveo: A FAIR national environmental data repository for earth observation open science, *Data Sci. J.*, 20, 22–22, <https://doi.org/10.5334/dsj-2021-022>, 2021.
- Gomes, L. C., Faria, R. M., de Souza, E., Veloso, G. V., Schaefer, C. E. G., and Fernandes Filho, E. I.: Modelling and mapping soil organic carbon stocks in Brazil, *Geoderma*, 340, 337–350, <https://doi.org/10.1016/j.geoderma.2019.01.007>, 2019.
- Gomes, V. C., Queiroz, G. R., and Ferreira, K. R.: An overview of platforms for big earth observation data management and analysis, *Remote Sens.*, 12, 1253, <https://doi.org/10.3390/rs12081253>, 2020.
- Gómez, C., White, J. C., and Wulder, M. A.: Optical remotely sensed time series data for land cover classification: A review, *ISPRS J. Photogram. Remote Sens.*, 116, 55–72, <https://doi.org/10.1016/j.isprsjprs.2016.03.008>, 2016.
- Google Sheet link: Landsat-based Spectral Indices Data Cube – Access Catalog, Google Sheet link [data set], [https://docs.google.com/spreadsheets/d/1QTA6OkkYIZlJfHst\\_inCrkC7DJcMAYHnM9k0iHulwpg/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1QTA6OkkYIZlJfHst_inCrkC7DJcMAYHnM9k0iHulwpg/edit?usp=sharing) (last access: 27 June 2024), 2024.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: 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>, 2017.
- He, X., Yang, L., Li, A., Zhang, L., Shen, F., Cai, Y., and Zhou, C.: Soil organic carbon prediction using phenological parameters and remote sensing variables generated from Sentinel-2 images, *Catena*, 205, 105442, <https://doi.org/10.1016/j.catena.2021.105442>, 2021.
- Heiden, U., d’Angelo, P., Schwind, P., Karlshöfer, P., Müller, R., Zepp, S., Wiesmeier, M., and Reinartz, P.: Soil reflectance composites – improved thresholding and performance evaluation, *Remote Sens.*, 14, 4526, <https://doi.org/10.3390/rs14184526>, 2022.
- Henebry, G. M. and De Beurs, K. M.: Remote sensing of land surface phenology: A prospectus, in: *Phenology: An integrative environmental science*, Springer, 385–411, [https://doi.org/10.1007/978-94-007-6925-0\\_21](https://doi.org/10.1007/978-94-007-6925-0_21), 2013.
- Hill, M. J. and Guerschman, J. P.: Global trends in vegetation fractional cover: Hotspots for change in bare soil and non-photosynthetic vegetation, *Agr. Ecosyst. Environ.*, 324, 107719, <https://doi.org/10.1016/j.agee.2021.107719>, 2022.
- Huete, A. R.: A soil-adjusted vegetation index (SAVI), *Remote Sens. Environ.*, 25, 295–309, [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X), 1988.
- Iosifescu Enescu, I., de Espona, L., Haas-Artho, D., Kurup Buchholz, R., Hanimann, D., Rietschi, M., Karger, D. N., Plattner, G.-K., Hägeli, M., Ginzler, C., Zimmermann, N. E., and Pellissier, L.: Cloud optimized raster encoding (core): A web-native streamable format for large environmental time series, *Geomatics*, 1, 369–382, <https://doi.org/10.3390/geomatics1030021>, 2021.
- Kansakar, P. and Hossain, F.: A review of applications of satellite earth observation data for global societal benefit and stewardship of planet earth, *Space Policy*, 36, 46–54, <https://doi.org/10.1016/j.spacepol.2016.05.005>, 2016.
- Killough, B.: Overview of the open data cube initiative, in: *IGARSS 2018 – 2018 IEEE international geoscience and remote sensing symposium*, 22–27 July 2018, Valencia, Spain, 8629–8632, <https://doi.org/10.1109/IGARSS.2018.8517694>, 2018.
- Killough, B.: The impact of analysis ready data in the Africa regional data cube, in: *IGARSS 2019 – 2019 IEEE International Geoscience and Remote Sensing Symposium*, 28 July–2 August 2019, Yokohama, Japan, 5646–5649, <https://doi.org/10.1109/IGARSS.2019.8898321>, 2019.
- Kooch, Y., Amani, M., and Abedi, M.: Vegetation degradation threatens soil health in a mountainous semi-arid region, *Sci. Total Environ.*, 830, 154827, <https://doi.org/10.1016/j.scitotenv.2022.154827>, 2022.
- Lacagnina, C., David, R., Nikiforova, A., Kuusniemi, M.-E., Cappiello, C., Biehlermaier, O., Wright, L., Schubert, C., Bertino, A., Thiemann, H., and Dennis, R.: Towards a data quality framework for EOSC authorship community, PhD thesis, EOSC Association, Zenodo, <https://doi.org/10.5281/zenodo.7515816>, 2022.
- Lawrence, I. and Lin, K.: A concordance correlation coefficient to evaluate reproducibility, *Biometrics*, 45, 255–268, <https://doi.org/10.2307/2532051>, 1989.

- Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., Raevksi, G., Hooke, J., Woodcock, R., Sixsmith, J., Wu, W., Tan, P., Li, F., Killough, B., Minchin, S., Roberts, D., Ayers, D., Bala, B., Dwyer, J., Dekker, A., Dhu, T., Hicks, A., Ip, A., Purss, M., Richards, C., Sagar, S., Trenham, C., Wang, P., and Wang, L.-W.: The Australian geoscience data cube – foundations and lessons learned, *Remote Sens. Environ.*, 202, 276–292, <https://doi.org/10.1016/j.rse.2017.03.015>, 2017.
- Li, L., Friedl, M. A., Xin, Q., Gray, J., Pan, Y., and Froelking, S.: Mapping crop cycles in China using MODIS-EVI time series, *Remote Sens.*, 6, 2473–2493, <https://doi.org/10.3390/rs6032473>, 2014.
- Liu, C., Zhang, Q., Tao, S., Qi, J., Ding, M., Guan, Q., Wu, B., Zhang, M., Nabil, M., Tian, F., Zeng, H., Zhang, N., Bavuudorj, G., Rukundo, E., Liu, W., Bofana, J., Beyene, A. N., and Elnashar, A.: A new framework to map fine resolution cropping intensity across the globe: Algorithm, validation, and implication, *Remote Sens. Environ.*, 251, 112095, <https://doi.org/10.1016/j.rse.2020.112095>, 2020.
- Lokers, R., Knapen, R., Janssen, S., van Randen, Y., and Jansen, J.: Analysis of Big Data technologies for use in agro-environmental science, *Environ. Model. Softw.*, 84, 494–504, <https://doi.org/10.1016/j.envsoft.2016.07.017>, 2016.
- Magdoff, F. and Van Es, H.: Building soils for better crops, in: vol. 4, Sustainable Agriculture Network Beltsville, <https://www.sare.org/wp-content/uploads/Building-Soils-for-Better-Crops.pdf> (last access: 28 November 2024), 2000.
- Maynard, J. J. and Levi, M. R.: Hyper-temporal remote sensing for digital soil mapping: Characterizing soil-vegetation response to climatic variability, *Geoderma*, 285, 94–109, <https://doi.org/10.1016/j.geoderma.2016.09.024>, 2017.
- McBratney, A. B., Santos, M. M., and Minasny, B.: On digital soil mapping, *Geoderma*, 117, 3–52, [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4), 2003.
- Meroni, M., d’Andrimont, R., Vrieling, A., Fasneder, D., Lemoine, G., Rembold, F., Seguini, L., and Verhegghen, A.: 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>, 2021.
- Montero, D., Aybar, C., Mahecha, M. D., Martinuzzi, F., Söchting, M., and Wieneke, S.: A standardized catalogue of spectral indices to advance the use of remote sensing in Earth system research, *Scientific Data*, 10, 197, <https://doi.org/10.1038/s41597-023-02096-0>, 2023 (code available at: <https://github.com/awesome-spectral-indices/>, last access: 27 June 2024).
- Myneni, R. and Williams, D.: On the relationship between FAPAR and NDVI, *Remote Sens. Environ.*, 49, 200–211, [https://doi.org/10.1016/0034-4257\(94\)90016-7](https://doi.org/10.1016/0034-4257(94)90016-7), 1994.
- Mzid, N., Pignatti, S., Huang, W., and Casa, R.: An analysis of bare soil occurrence in arable croplands for remote sensing topsoil applications, *Remote Sens.*, 13, 474, <https://doi.org/10.3390/rs13030474>, 2021.
- Nazari, S., Rostamina, M., Ayoubi, S., Rahmani, A., and Mousavi, S. R.: Efficiency of Different Feature Selection Methods in Digital Mapping of Subgroup and Soil Family Classes with Data Mining Algorithms, *Water Soil*, 34, 973–987, <https://doi.org/10.22067/jsw.v34i4.85748>, 2020.
- Nguyen, L. H., Joshi, D. R., Clay, D. E., and Henebry, G. M.: Characterizing land cover/land use from multiple years of Landsat and MODIS time series: A novel approach using land surface phenology modeling and random forest classifier, *Remote Sens. Environ.*, 238, 111017, <https://doi.org/10.1016/j.rse.2018.12.016>, 2020.
- Obuchowicz, C., Poussin, C., and Giuliani, G.: Change in observed long-term greening across Switzerland – evidence from a three decades NDVI time-series and its relationship with climate and land cover factors, *Big Earth Data*, 8, 1–32, <https://doi.org/10.1080/20964471.2023.2268322>, 2023.
- Patel, J. H. and Oza, M. P.: Deriving crop calendar using NDVI time-series, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-8, 869–873, <https://doi.org/10.5194/isprsarchives-XL-8-869-2014>, 2014.
- Pietrzykowski, M.: Soil and plant communities development and ecological effectiveness of reclamation on a sand mine cast, *J. Forest Sci.*, 54, 554–565, 2008.
- Pietrzykowski, M. and Krzaklewski, W.: An assessment of energy efficiency in reclamation to forest, *Ecol. Eng.*, 30, 341–348, <https://doi.org/10.1016/j.ecoleng.2007.04.003>, 2007.
- Pôças, I., Gonçalves, J., Marcos, B., Alonso, J., Castro, P., and Honrado, J. P.: Evaluating the fitness for use of spatial data sets to promote quality in ecological assessment and monitoring, *Int. J. Geogr. Inform. Sci.*, 28, 2356–2371, <https://doi.org/10.1080/13658816.2014.924627>, 2014.
- Potapov, P., Hansen, M. C., Kommareddy, I., Kommareddy, A., Turbanova, S., Pickens, A., Adusei, B., Tyukavina, A., and Ying, Q.: Landsat analysis ready data for global land cover and land cover change mapping, *Remote Sens.*, 12, 426, <https://doi.org/10.3390/rs12030426>, 2020.
- Poussin, C., Massot, A., Ginzler, C., Weber, D., Chatenoux, B., Lacroix, P., Piller, T., Nguyen, L., and Giuliani, G.: Drying conditions in Switzerland – indication from a 35-year Landsat time-series analysis of vegetation water content estimates to support SDGs, *Big Earth Data*, 5, 445–475, <https://doi.org/10.1080/20964471.2021.1974681>, 2021.
- Poussin, C., Timoner, P., Chatenoux, B., Giuliani, G., and Peduzzi, P.: Improved Landsat-based snow cover mapping accuracy using a spatiotemporal NDSI and generalized linear mixed model, *Sci. Remote Sens.*, 7, 100078, <https://doi.org/10.1016/j.srs.2023.100078>, 2023.
- Radeloff, V. C., Roy, D. P., Wulder, M. A., Anderson, M., Cook, B., Crawford, C. J., Friedl, M., Gao, F., Gorelick, N., Hansen, M., Healey, S., Hostert, P., Hulley, G., Huntington, J. L., Johnson, D. M., Neigh, C., Lyapustin, A., Lymburner, L., Pahlevan, N., Pekel, J.-F., Scambos, T. A., Schaaf, C., Strobl, P., Woodcock, C. E., Zhang, H. K., and Zhu, Z.: Need and vision for global medium-resolution Landsat and Sentinel-2 data products, *Remote Sens. Environ.*, 300, 113918, <https://doi.org/10.1016/j.rse.2023.113918>, 2024.
- Ranghetti, L., Busetto, L., Crema, A., Fasola, M., Cardarelli, E., and Boschetti, M.: Testing estimation of water surface in Italian rice district from MODIS satellite data, *Int. J. Appl. Earth Obs. Geoinf.*, 52, 284–295, <https://doi.org/10.1016/j.jag.2016.06.018>, 2016.
- Reddy, N. M., Peerzada, I. A., Moonis, M., and Singh, O.: Application of biophysical, soil, and vegetation indices to better understand forest dynamics and develop strategies

- for forest conservation, in: *Forest Dynamics and Conservation: Science, Innovations and Policies*, Springer, 399–415, [https://doi.org/10.1007/978-981-19-0071-6\\_19](https://doi.org/10.1007/978-981-19-0071-6_19), 2022.
- Rhyman, P., Norizah, K., Hamdan, O., Faridah-Hanum, I., and Zulfa, A.: Integration of normalised different vegetation index and Soil-Adjusted Vegetation Index for mangrove vegetation delineation, *Remote Sens. Appl.: Soc. Environ.*, 17, 100280, <https://doi.org/10.1016/j.rsase.2019.100280>, 2020.
- Rizzo, R., Wadoux, A. M.-C., Dematté, J. A., Minasny, B., Barrón, V., Ben-Dor, E., Francos, N., Savin, I., Poppiel, R., Silvero, N. E. Q., da Silva, F., Rosin, N. A., Rosas, J. T. F., Greschuk, L. T., Ballester, M. V. R., Gómez, A. M. R., Bellinaso, H., Safanelli, J. L., Chabrilat, S., Fiorio, P. R., Das, B. S., Malone, B. P., Zalidis, G., Tziolas, N., Tsakiridis, N., Karyotis, K., Samarinas, N., Kaloposa, E., Gholizadeh, A., Shepherd, K. D., Milewski, R., Vaudour, E., Wang, C., and Salama, E. S. M.: Remote sensing of the Earth's soil color in space and time, *Remote Sens. Environ.*, 299, 113845, <https://doi.org/10.1016/j.rse.2023.113845>, 2023.
- Robinson, N. P., Allred, B. W., Smith, W. K., Jones, M. O., Moreno, A., Erickson, T. A., Naugle, D. E., and Running, S. W.: Terrestrial primary production for the conterminous United States derived from Landsat 30 m and MODIS 250 m, *Remote Sens. Ecol. Conserv.*, 4, 264–280, <https://doi.org/10.1002/rse2.74>, 2018.
- Rogge, D., Bauer, A., Zeidler, J., Mueller, A., Esch, T., and Heiden, U.: Building an exposed soil composite processor (SCMaP) for mapping spatial and temporal characteristics of soils with Landsat imagery (1984–2014), *Remote Sens. Environ.*, 205, 1–17, <https://doi.org/10.1016/j.rse.2017.11.004>, 2018.
- Safanelli, J. L., Chabrilat, S., Ben-Dor, E., and Dematté, J. A.: Multispectral models from bare soil composites for mapping topsoil properties over Europe, *Remote Sens.*, 12, 1369, <https://doi.org/10.3390/rs12091369>, 2020.
- Salcedo-Sanz, S., Ghamisi, P., Piles, M., Werner, M., Cuadra, L., Moreno-Martínez, A., Izquierdo-Verdiguier, E., Muñoz-Marí, J., Mosavi, A., and Camps-Valls, G.: Machine learning information fusion in Earth observation: A comprehensive review of methods, applications and data sources, *Inform. Fusion*, 63, 256–272, <https://doi.org/10.1016/j.inffus.2020.07.004>, 2020.
- Salomonson, V. V. and Appel, I.: Development of the Aqua MODIS NDSI fractional snow cover algorithm and validation results, *IEEE T. Geosci. Remote*, 44, 1747–1756, <https://doi.org/10.1109/TGRS.2006.876029>, 2006.
- Seifert, C. A. and Lobell, D. B.: Response of double cropping suitability to climate change in the United States, *Environ. Res. Lett.*, 10, 024002, <https://doi.org/10.1088/1748-9326/10/2/024002>, 2015.
- Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, *J. Am. Stat. Assoc.*, 63, 1379–1389, <https://doi.org/10.1080/01621459.1968.10480934>, 1968.
- Sharma, P., Singh, A., Kahlon, C. S., Brar, A. S., Grover, K. K., Dia, M., and Steiner, R. L.: The role of cover crops towards sustainable soil health and agriculture – A review paper, *Am. J. Plant Sci.*, 9, 1935–1951, <https://doi.org/10.4236/ajps.2018.99140>, 2018.
- Shi, J., Yang, L., Zhu, A.-X., Qin, C., Liang, P., Zeng, C., and Pei, T.: Machine-Learning Variables at Different Scales vs. Knowledge-based Variables for Mapping Multiple Soil Properties, *Soil Sci. Soc. Am. J.*, 82, 645–656, <https://doi.org/10.2136/sssaj2017.11.0392>, 2018.
- Siebert, S., Portmann, F. T., and Döll, P.: Global patterns of cropland use intensity, *Remote Sens.*, 2, 1625–1643, <https://doi.org/10.3390/rs2071625>, 2010.
- Sonmez, N. K. and Slater, B.: Measuring intensity of tillage and plant residue cover using remote sensing, *Eur. J. Remote Sens.*, 49, 121–135, <https://doi.org/10.5721/EuJRS20164907>, 2016.
- Sothe, C., Gonsamo, A., Arabian, J., and Snider, J.: Large scale mapping of soil organic carbon concentration with 3D machine learning and satellite observations, *Geoderma*, 405, 115402, <https://doi.org/10.1016/j.geoderma.2021.115402>, 2022.
- Stern, C., Abernathy, R., Hamman, J., Wegener, R., Lepore, C., Harkins, S., and Merose, A.: Pangeo Forge: Crowdsourcing Analysis-Ready, Cloud Optimized Data Production, *Front. Clim.*, 3, 782909, <https://doi.org/10.3389/fclim.2021.782909>, 2022.
- Storey, E. A., Stow, D. A., and O'Leary, J. F.: Assessing postfire recovery of chamise chaparral using multi-temporal spectral vegetation index trajectories derived from Landsat imagery, *Remote Sens. Environ.*, 183, 53–64, <https://doi.org/10.1016/j.rse.2016.05.018>, 2016.
- Szabo, S., Gácsi, Z., and Balazs, B.: Specific features of NDVI, NDWI and MNDWI as reflected in land cover categories, *Acta Geogr. Debrecina Landsc. Environ. Ser.*, 10, 194–202, <https://doi.org/10.21120/LE/10/3-4/13>, 2016.
- Taha, A. A. and Hanbury, A.: Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool, *BMC Med. Imag.*, 15, 1–28, <https://doi.org/10.1186/s12880-015-0068-x>, 2015.
- Theil, H.: A rank-invariant method of linear and polynomial regression analysis, *Indagationes mathematicae, P. Roy. Neth. Acad. Sci.*, 53, Part I: 386–392, Part II: 521–525, Part III: 1397–1412, 1950.
- Tian, X., Ho, Y.-F., Witjes, M., Parente, L., Hengl, T., and Minarik, R.: Pan-EU Landmask: 10 m Resolution Geospatial Land Coverage with Administrative Boundary details on country and regional level (Version v20230722) [Data set], Zenodo [data set], <https://doi.org/10.5281/zenodo.8171860>, 2023.
- Tian, X., Consoli, D., Parente, L., Yufeng, H., and Hengl, T.: Landsat-based spectral indices for pan-EU 2000–2022, Zenodo [data set], <https://doi.org/10.5281/zenodo.10776891>, 2024a.
- Tian, X., Hengl, T., yu-feng-ho, and d-consoli: AI4SoilHealth/SoilHealthDataCube: v20241204-1 (v20241204-1), Zenodo [code], <https://doi.org/10.5281/zenodo.12907281>, 2024b.
- Tucker, C. J.: Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sens. Environ.*, 8, 127–150, [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0), 1979.
- Turmel, M.-S., Speratti, A., Baudron, F., Verhulst, N., and Govaerts, B.: Crop residue management and soil health: A systems analysis, *Agricult. Syst.*, 134, 6–16, <https://doi.org/10.1016/j.agsy.2014.05.009>, 2015.
- Van Deventer, A., Ward, A., Gowda, P., and Lyon, J.: Using thematic mapper data to identify contrasting soil plains and tillage practices, *Photogram. Eng. Remote Sens.*, 63, 87–93, 1997.
- Wadoux, A. M.-C., Minasny, B., and McBratney, A. B.: Machine learning for digital soil mapping: Applications, challenges and suggested solutions, *Earth-Sci. Rev.*, 210, 103359, <https://doi.org/10.1016/j.earscirev.2020.103359>, 2020.

- Wagemann, J., Siemen, S., Seeger, B., and Bendix, J.: Users of open Big Earth data – An analysis of the current state, *Comput. Geosci.*, 157, 104916, <https://doi.org/10.1016/j.cageo.2021.104916>, 2021.
- Wentz, E. A. and Shimizu, M.: Measuring spatial data fitness-for-use through multiple criteria decision making, *Ann. Am. Assoc. Geogr.*, 108, 1150–1167, <https://doi.org/10.1080/24694452.2017.1411246>, 2018.
- White, M. A., Thornton, P. E., and Running, S. W.: A continental phenology model for monitoring vegetation responses to interannual climatic variability, *Global Biogeochem. Cy.*, 11, 217–234, <https://doi.org/10.1029/97GB00330>, 1997.
- Wiesmeier, M., Urbanski, L., Hobbey, E., Lang, B., von Lütow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H.-J., and Kögel-Knabner, I.: Soil organic carbon storage as a key function of soils—A review of drivers and indicators at various scales, *Geoderma*, 333, 149–162, <https://doi.org/10.1016/j.geoderma.2018.07.026>, 2019.
- Witjes, M., Parente, L., Križan, J., Hengl, T., and Antonić, L.: Eco-datacube.eu: analysis-ready open environmental data cube for Europe, *Peer J.*, 11, e15478, <https://doi.org/10.7717/peerj.15478>, 2023.
- Wu, J., Lin, L., Li, T., Cheng, Q., Zhang, C., and Shen, H.: Fusing Landsat 8 and Sentinel-2 data for 10-m dense time-series imagery using a degradation-term constrained deep network, *Int. J. Appl. Earth Obs. Geoinf.* 108, 102738, <https://doi.org/10.1016/j.jag.2022.102738>, 2022.
- Wulder, M. A., Roy, D. P., Radeloff, V. C., Loveland, T. R., Anderson, M. C., Johnson, D. M., Healey, S., Zhu, Z., Scambos, T. A., Pahlevan, N., Hansen, M., Gorelick, N., Crawford, C. J., Masek, J. G., Hermosilla, T., White, J. C., Belward, A. S., Schaaf, C., Woodcock, C. E., Huntington, J. L., Lymburner, L., Hostert, P., Gao, F., Lyapustin, A., Pekel, J.-F., Strobl, P., and Cook, B. D.: Fifty years of Landsat science and impacts, *Remote Sens. Environ.*, 280, 113195, <https://doi.org/10.1016/j.rse.2022.113195>, 2022.
- Xiang, X., Du, J., Jacinthe, P.-A., Zhao, B., Zhou, H., Liu, H., and Song, K.: Integration of tillage indices and textural features of Sentinel-2A multispectral images for maize residue cover estimation, *Soil Till. Res.*, 221, 105405, <https://doi.org/10.1016/j.still.2022.105405>, 2022.
- Xue, Z., Du, P., and Feng, L.: Phenology-driven land cover classification and trend analysis based on long-term remote sensing image series, *IEEE J. Select. Top. Appl. Earth Obs. Remote Sens.*, 7, 1142–1156, <https://doi.org/10.1109/JSTARS.2013.2294956>, 2014.
- Yang, L., He, X., Shen, F., Zhou, C., Zhu, A.-X., Gao, B., Chen, Z., and Li, M.: Improving prediction of soil organic carbon content in croplands using phenological parameters extracted from NDVI time series data, *Soil Till. Res.*, 196, 104465, <https://doi.org/10.1016/j.still.2019.104465>, 2020.
- Yang, R.-M., Liu, L.-A., Zhang, X., He, R.-X., Zhu, C.-M., Zhang, Z.-Q., and Li, J.-G.: The effectiveness of digital soil mapping with temporal variables in modeling soil organic carbon changes, *Geoderma*, 405, 115407, <https://doi.org/10.1016/j.geoderma.2021.115407>, 2022.
- Yang, T., Siddique, K. H., and Liu, K.: Cropping systems in agriculture and their impact on soil health – A review, *Global Ecol. Conserv.*, 23, e01118, <https://doi.org/10.1016/j.gecco.2020.e01118>, 2020.
- Yang, X., Blower, J., Bastin, L., Lush, V., Zabala, A., Masó, J., Cornford, D., Díaz, P., and Lumsden, J.: An integrated view of data quality in Earth observation, *Philos. T. Roy. Soc. A*, 371, 20120072, <https://doi.org/10.1098/rsta.2012.0072>, 2013.
- Zepp, S., Heiden, U., Bachmann, M., Möller, M., Wiesmeier, M., and van Wesemael, B.: Optimized bare soil compositing for soil organic carbon prediction of topsoil croplands in Bavaria using Landsat, *ISPRS J. Photogram. Remote Sens.*, 202, 287–302, <https://doi.org/10.3390/rs13030474>, 2023.
- Zeraatpisheh, M., Garosi, Y., Owliaie, H. R., Ayoubi, S., Taghizadeh-Mehrjardi, R., Scholten, T., and Xu, M.: Improving the spatial prediction of soil organic carbon using environmental covariates selection: A comparison of a group of environmental covariates, *Catena*, 208, 105723, <https://doi.org/10.1016/j.catena.2021.105723>, 2022.
- Zhang, X., Liu, L., Liu, Y., Jayavelu, S., Wang, J., Moon, M., Henebry, G. M., Friedl, M. A., and Schaaf, C. B.: Generation and evaluation of the VIIRS land surface phenology product, *Remote Sens. Environ.*, 216, 212–229, <https://doi.org/10.1016/j.rse.2018.06.047>, 2018.
- Zhang, X., Liu, L., Chen, X., Gao, Y., Xie, S., and Mi, J.: GLC\_FCS30: global land-cover product with fine classification system at 30 m using time-series Landsat imagery, *Earth Syst. Sci. Data*, 13, 2753–2776, <https://doi.org/10.5194/essd-13-2753-2021>, 2021.
- Zheng, B., Campbell, J. B., and de Beurs, K. M.: Remote sensing of crop residue cover using multi-temporal Landsat imagery, *Remote Sens. Environ.*, 117, 177–183, <https://doi.org/10.1016/j.rse.2011.09.016>, 2012.
- Zheng, B., Campbell, J., Serbin, G., and Daughtry, C.: Multitemporal remote sensing of crop residue cover and tillage practices: A validation of the minNDTI strategy in the United States, *J. Soil Water Conserv.*, 68, 120–131, <https://doi.org/10.2489/jswc.68.2.120>, 2013.
- Zhou, J., Jia, L., Menenti, M., and Gorte, B.: On the performance of remote sensing time series reconstruction methods – A spatial comparison, *Remote Sens. Environ.*, 187, 367–384, <https://doi.org/10.1016/j.rse.2016.10.025>, 2016.
- Zhou, T., Geng, Y., Ji, C., Xu, X., Wang, H., Pan, J., Bumberger, J., Haase, D., and Lausch, A.: Prediction of soil organic carbon and the C : N ratio on a national scale using machine learning and satellite data: A comparison between Sentinel-2, Sentinel-3 and Landsat-8 images, *Sci. Total Environ.*, 755, 142661, <https://doi.org/10.1016/j.scitotenv.2020.142661>, 2021.