European Journal of Remote Sensing

ISSN: 2279-7254 (Online) Journal homepage: www.tandfonline.com/journals/tejr20

Tillage and biomass detection for estimating winter-time cropland management practices with satellite remote sensing

Maria Yli-Heikkilä, Arto Klami, Samantha Wittke, Markku Luotamo, Pinja Mero, Petri Pellikka, Janne Heiskanen, Mwaba Hiltunen, Kari Luojus, Golda Prakasam, Mikko Strahlendorff, Markus Törmä & Mi Sulkava

To cite this article: Maria Yli-Heikkilä, Arto Klami, Samantha Wittke, Markku Luotamo, Pinja Mero, Petri Pellikka, Janne Heiskanen, Mwaba Hiltunen, Kari Luojus, Golda Prakasam, Mikko Strahlendorff, Markus Törmä & Mi Sulkava (2025) Tillage and biomass detection for estimating winter-time cropland management practices with satellite remote sensing, European Journal of Remote Sensing, 58:1, 2525967, DOI: 10.1080/22797254.2025.2525967

To link to this article: https://doi.org/10.1080/22797254.2025.2525967

EUROPEAN JOURNAL OF REMOTE SENSING
2025, VOL. 58, NO. 1, 2525967
https://doi.org/10.1080/22797254.2025.2525967

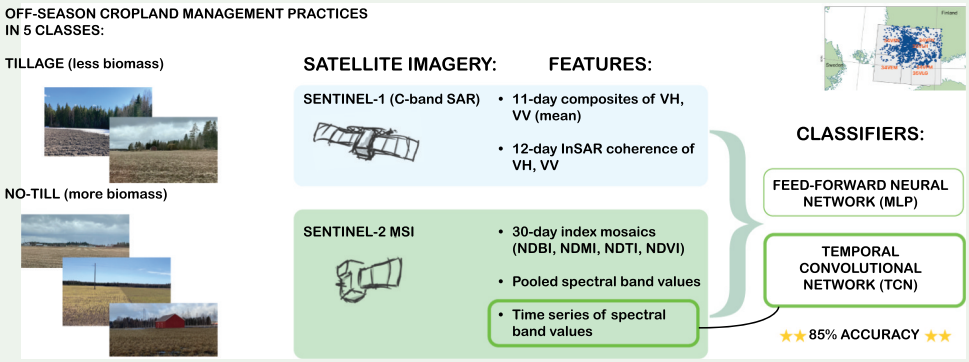
Tillage and biomass detection for estimating winter-time cropland management practices with satellite remote sensing

Maria Yli-Heikkilä[a,b], Arto Klami[c], Samantha Wittke[d,e], Markku Luotamo[c], Pinja Mero[a], Petri Pellikka[b], Janne Heiskanen[b,f], Mwaba Hiltunen[f], Kari Luojus[f], Golda Prakasam[f], Mikko Strahlendorff[f], Markus Törmä[g] and Mi Sulkava[a]

[a]Natural Resources Institute Finland, Helsinki, Finland; [b]Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland; [c]Department of Computer Science, University of Helsinki, Helsinki, Finland; [d]Finnish Geospatial Research Institute in the National Land Survey of Finland, Espoo, Finland; [e]Department of Built Environment, Aalto University, Espoo, Finland; [f]Finnish Meteorological Institute, Helsinki, Finland; [g]Finnish Environment Institute, Helsinki, Finland

ABSTRACT

Supportive policies to promote sustainable agriculture have been implemented across countries and regions. For example, continuous vegetative groundcover and reduced tillage have been advocated for sustainable post-harvest biomass management. Accurate and timely information on cropland management practices is needed for agricultural policy evaluations, evidence-based planning, and agri-environmental assessments. We show that a satellite-based approach can yield off-season cropland management information on preferred spatial and temporal scales from a narrow window of opportunity in early spring after snow melt and before seedbed preparation. Agricultural parcel geometries from an administrative registry were used to extract information on Sentinel-1 backscatter and coherence, and Sentinel-2 spectral reflectance. Based on a large survey-based dataset of 6,623 fields, we show that the highest impact on model performance comes from the spectral regions of near-infrared and upper red edge of the Sentinel-2 mission, whereas Sentinel-1–based features had a relatively small contribution to classification performance. Our proposed method for tillage and biomass detection generalises well in the study area of boreal environmental zone with dominantly mineral soils, as confirmed by the high test set classification accuracy of 85%. The supporting dataset and codes are stored in a publicly accessible repository.



ARTICLE HISTORY
Received 23 August 2024
Revised 24 March 2025
Accepted 18 June 2025

KEYWORDS
Object-based; satellite remote sensing; temporal convolutional neural network; random forest; satellite image time series; agricultural monitoring

Introduction

Croplands and pastures comprise 36% of the global land surface and this share is slowly decreasing, albeit with an increase in sub-Saharan Africa and elsewhere in the tropics (FAO, 2022). Agricultural land use presents us with a dilemma: while the persistent rise in the world's population calls for intensified food production, modern cropland management practices are degrading the ecosystems and ecosystem services upon which we depend (Amundson et al., 2015; Foley et al., 2005). To confront the diminishing capacity of global ecosystems to sustain food production, supportive policies to promote sustainable agriculture have been implemented across countries and regions. One of the target areas of agricultural support programmes has been regionally specific sustainable off-season post-harvest biomass management (e.g. European Commission 2023). Generally, cropland management practices such as continuous vegetative groundcover and reduced tillage have been advocated to support resilient farming systems.

Accurate and timely information on post-harvest biomass management is needed for agricultural policy

CONTACT Maria Yli-Heikkilä maria.yli-heikkila@luke.fi Natural Resources Institute Finland, Latokartanonkaari 9, Helsinki 00790, Finland
Supplemental data for this article can be accessed online at https://doi.org/10.1080/22797254.2025.2525967.

evaluations and evidence-based planning. For example, variables within the European Union's Integrated Farm Statistics regulation strive to depict the interlinks between winter-time cropland management practices and soil productivity and biophysical processes (European Union, 2018). Furthermore, agricultural monitoring and many agricultural and environmental assessments involve information on off-season croplands. For example, tillage affects the nutrient and terrestrially fixed carbon runoff from the soil into water bodies (Bechmann & Bøe, 2021; Van Oost et al., 2005) as well as soil organic carbon stocks and fluxes (Dlugoß et al., 2012). Although current parcel-level information on off-season cropland management practices can to some extent be extracted from the existing agricultural registries, e.g. in the European Union Member States, the information need has not been completely fulfilled (Matthews et al., 2023).

Space-borne remote sensing has been successfully applied to the mapping of post-harvest practices (see reviews in Zheng et al. (2015) and Bégué et al. (2018)). Several authors have used both Synthetic Aperture Radar (SAR) and optical sensor data to monitor typically wintertime cropland management practices in specific geographical areas using various machine learning methods. For instance, Azzari et al. (2019) used optical and radar data to map tillage practices across the north central region of the US. For optical data, they used the Landsat 5/7/8 collections and computed several spectral indices per pixel. For radar data, they used SAR Sentinel-1 (S1) imagery. For each pixel and band, they also computed surface texture metrics. The individual annual observations were converted to an analysis-ready format using two different approaches to temporal compositing. A random forest classifier was trained on a survey-based dataset to produce large-scale tillage intensity maps on a binary scale (high-tillage, low-tillage) at 30 m resolution. The best model had an overall accuracy between 75% and 79%, depending on the validation approach.

Thieme et al. (2020) estimated the winter cover crop performance for agricultural monitoring using optical Landsat 5/7/8 and harmonised Sentinel-2 (S2) imagery to compute field-level Normalised Difference Vegetation Indices (NDVIs) in the state of Maryland in the US. The models for cover crop conservation performance were calibrated using a linear regression between the seasonal maximum NDVI and in situ data collected on Maryland farms. The models were reported to have a significant correlation with the cover crop biomass ($R^2 = 0.56$), and with the observed percent vegetative ground cover ($R^2 = 0.68$). Denize et al. (2019), in turn, evaluated object- and pixel-based approaches to identify wintertime agricultural land use. Their study addressed the high diversity of off-season farming strategies and practices divided into multiple classes: winter crops, grasslands, catch crops, crop residues, and bare soil. Several differential index, biophysical, and polarimetric features were calculated from the time series of S1 and S2 images over a study site in France in the winter of 2016–2017. The best performing classifier, random forest, combined both S1 and S2 features and used an object-based approach with an overall accuracy of 81%.

Acknowledging the potential of remote sensing to complement the evidence base provided by agricultural registries, our study area in northern Europe nevertheless imposed limitations on the applicability of the previous research results. Most of the off-season (winter) is unsuitable for optical remote sensing due to darkness and snow cover, leaving a narrow window of opportunity for monitoring in early spring. Driven by the interests of stakeholders, information on off-season cropland management is needed on a more granular level than previous solutions have encompassed. Therefore, for a sound base, our approach builds on new, unique in-situ data that covers a large geographical area in southwestern Finland for the period 2020–2023. Although the study is located in one of the northern-most agricultural systems in the world, the methodology should be feasible in other parts of the Northern Hemisphere, especially in similar industrial agricultural systems of northern Europe and North America.

Our main objective was to find the best performing classification model for estimating diverse off-season cropland management practices from the perspective of an operational application for agroenvironmental monitoring. More specifically, by grounding on an extensive in-situ dataset, we:

- Studied which cropland management practices can be distinguished by a classifier to a satisfactory degree of accuracy.
- Challenged the established vegetation index-based research by introducing band-based optical features.
- Tested several ways to incorporate optical and radar features for modelling.
- Elaborated on the impact of features on the classification accuracy.

To identify the best-performing model, we considered alternative feature representations and information sources and evaluated classifier methods from classic to deep machine learning. New information on the most informative and discriminative sensing data in our study setting contributes to the advancement of knowledge in remote sensing of tillage practices and post-harvest biomass and can furthermore assist in designing an

operational application. The resulting recommendations of the study pave the way to effectively harnessing remote sensing and machine learning for mapping off-season cropland management practices over large agricultural areas needed for, e.g. agricultural statistics, policy evaluations, or environmental research.

Materials and methods

Study area

The study was carried out in southwestern Finland between latitudes 59°83'N and 61°49'N spanning from the towns of Hanko in the south, to Pori in the north and Lahti in the east. The area belongs to the northernmost European annual cropping system. Geomorphologically, the croplands in the area are characterised by low plateaus and undulating plains with clay soils. According to the Köppen climate classification, the area belongs mainly to the cold summer humid continental climate zone (Dfc) with a snow-dominated winter and short growing season (Beck et al., 2018). According to the environmental stratification of Europe, the study area belongs to the boreal environmental zone in its northern parts and hemiboreal zone in its southern parts close to the Gulf of Finland (Metzger et al., 2005).

The region of interest (ROI) was set to comprise a geographical area in which spring occurs and the croplands are free of snow at approximately the same time. The time of interest (TOI) was the narrow window between snow melt and the beginning of the growing season, followed by spring tillage. Based on meteorological data, the growing season typically begins around the middle of April in the southwestern parts of Finland (Finnish Meteorological Institute, 2023). Any earlier observation would most probably suffer from snow cover, at least in the northern and eastern parts of the ROI, due to variations in the local climate. In 2022, the growing season began a few weeks later, at the end of April. See Table 1 for the selected annual observation windows.

Six S2 tiles (34VEN, 34VEM, 35VLH, 34VFM, 34VFN, 35VLG) outline the ROI, as shown in Figure 1. The area of the ROI is ~36,000 km2. While only 8% of the soil of the sampled fields within the ROI is organic, the dominant mineral soils (92%) are clay soils (42%), and rough mineral silt (20%) or till (13%) soils (Lilja et al., 2017).

Table 1. Annual observation windows (dates) for each type of satellite imagery. S1 stands for Sentinel-1 products; S1 IW GRD for a Sentinel-1 ground range detected product type in interferometric wide swath mode; S2 for Sentinel-2; S2 L2A for a Sentinel-2 product type of level 2A; S1 COH12 for a Sentinel-1 12-day interferometric coherence product. Additionally, there are annual minimum and maximum numbers of observations per parcel in the region of interest for Sentinel-1 composites, coherences, and Sentinel-2 time series, and the number of parcels in the in-situ dataset.

Year	Month	Dates: S1 composite	# S1 IW GRD images	Dates: S2 images	# S2 L2A images	Dates: S2 indices	Dates: S1 COH12	# parcels
2020	April	11–20	2... 3	1–21	2... 8	1–30	—	69
2021	April	11–20	3... 5	1–21	2... 7	1–30	—	100
2022	April	21–30	2	11–30	2... 8	1–30	—	413
2023	April	11–20	3... 4	1–21	1... 10	1–30	8th, 20th	6,041

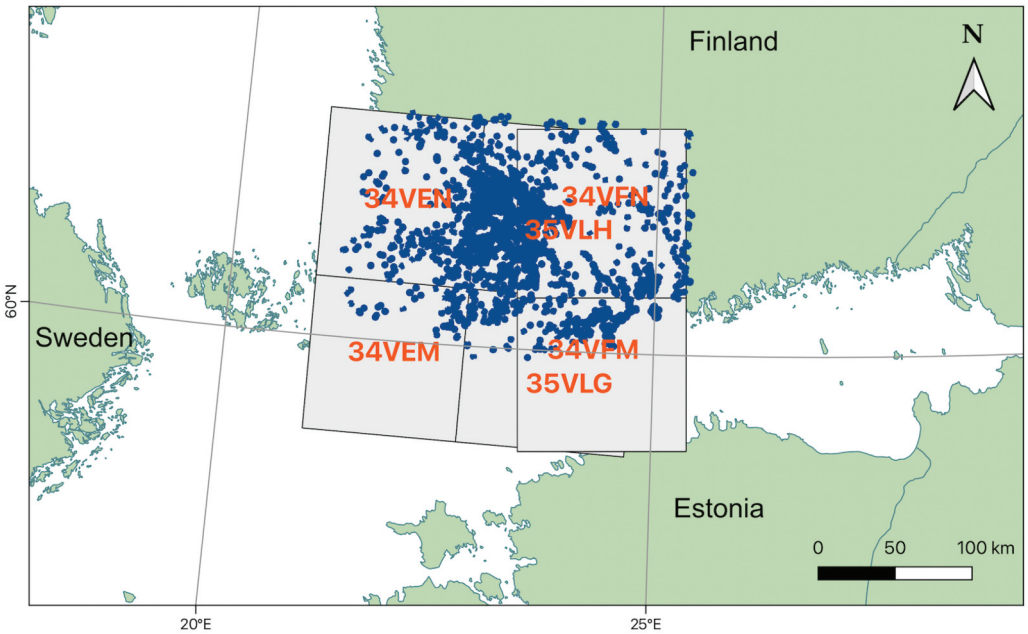


Figure 1. The region of interest (ROI) of the study was determined on agrometeorological grounds. Six Sentinel-2 tiles cover the ROI. The blue dots indicate the sampled fields in the ROI.

Data

Reference data

The reference data hold the correct parcel-wise cropland management labels and field parcel geometries. The data originate from both field surveys and agricultural registries. The labels were primarily recorded in a targeted roadside data collection campaign conducted by the Natural Resources Institute Finland in 2020–2023. In addition, we received cropland management data from farms, which volunteered to participate in the data collection. Our study aimed to minimize subjectivity by employing trained data collectors and carefully selecting labels that meet information needs while remaining visually distinguishable. To ensure a sufficient number of labels, we scaled the campaign up in 2023 to collect as much data as possible. In the roadside survey, the fields were chosen according to their proximity to suitable routes in the agricultural landscapes and observed from the edges of fields.

The roadside data collection was complemented by selected parts of the data in the agricultural registries of the National Integrated Administration and Control System (IACS), operated by the Finnish Food Authority. The IACS contains farmer declarations on parcel-wise agri-environmental measures and the crops cultivated. There are several measures that commit to off-season vegetation cover. For each field parcel, our rule-based algorithm assigned a vegetation cover type based on the information on the farmers' environmental commitments. The overall accuracy of the rule-based algorithm was rather low, especially with tillage classes, while the overall balanced accuracy was 72%. However, the winter crop declarations were 99% accurate and hence could be safely included in the reference data (see details in Supplement A.1). As a result, this study had a reference dataset of 6,623 parcels (see Table 1 for the annual numbers of parcels in the dataset). In addition, IACS field parcel geometries were used to map parcel-level intensity values from the satellite images.

Cropland management classes

We specified eight cropland management classes to encompass all the off-season cropland management categories in Finland, namely:

(1) **High till** High-intensity conventional tillage technique that typically involves mould-board ploughing to a depth of 20–25 cm, leaving the soil bare from the post-harvest season in September–October until the pre-seeding season in late April–May. See Figure 2(a).
(2) **Low till** Low-intensity conservation tillage includes various tilling methods that mechanically disturb the soil to a depth of less than 15 cm while retaining most of the crop residues on the surface. See Figure 2(b).
(3) **Winter crop** Cereal crops (mainly rye and winter wheat, but also winter barley) and oil crops (turnip rape and rape) sown in August–September. See Figure 2(c).

(a) High till (b) Low till (c) Winter crop (d) Grass (e) Stubble (f) Stubble with companion crop

Figure 2. The six cropland management classes used in the study.

(4) **Grass** Dense vegetation cover of herbaceous plants. See Figure 2(d).
(5) **Cereal crop stubble** Post-harvest cereal crop stubble and crop residues partly covering the soil. See Figure 2(e).
(6) **Other stubble** Post-harvest other than cereal crop (such as oil crop) stubble and crop residues partly covering the soil.
(7) **Stubble with companion crop** Stubble fields where a companion crop (catch crop or cover crop) vegetates the soil after the cash crop is harvested. See Figure 2(f).
(8) **Biomass burning** Burning the excessive crop residues soon after harvest to prepare the land for sowing.

From the remote sensing perspective, cereal crop stubble exhibits different reflectance patterns compared to other crop stubble such as oil seed crops. However, cereal crops constitute a clear majority of the stubble crops, and hence we combined the classes. Because burning the biomass is seldom practiced within the study area, we excluded this class. The remaining six classes sufficiently fulfill the information needs for cropland management practices.

Optical remote sensing data

For optical remote sensing data, we used Copernicus Sentinel-2 imagery from the multi-spectral instrument aboard the S2A and S2B satellites. The imagery was pre-processed in three different ways. In the first two approaches, geometrically and atmospherically corrected bottom-of-the-atmosphere Sentinel-2 Level-2A products were downloaded from the Copernicus Open Access Hub (European Space Agency, 2023). We excluded scenes with a cloud cover of over 95%. We then used 10 spectral bands suitable for environmental monitoring with the following central wavelengths: Band 2 (492 nm), Band 3 (560 nm), Band 4 (665 nm), Band 5 (705 nm), Band 6 (740 nm), Band 7 (783 nm), Band 8 (842 nm), Band 8A (865 nm), Band 11 (1610 nm), and Band 12 (2190 nm) (European Space Agency, 2015). In the first approach, we selected single Level-2A products within the given time frame and treated the observations as time series. In the second approach, we took all the individual images within the time window and pooled the parcel-wise pixel values. See Table 1 for the acquisition dates and the numbers of images.

For the third approach, we used the S2 30-day image index mosaics provided by the Finnish Environment Institute SYKE (Törmä, 2020). The 30-day mosaic of Normalised Difference Vegetation Index (NDVI), Tillage Index (NDTI), Built-up Index (NDBI), and Moisture Index (NDMI) are operationally computed in CalFin-processing cluster (Fomferra et al., 2012) at the National Satellite Data Center (NSDC) of the Finnish Meteorological Institute (FMI). The monthly mosaics are based on maximum NDVI, i.e. the pixel values of the mosaics are selected from the image where the pixel-wise NDVI value is the highest within the 30-day period. See details in Supplement A.2. The index mosaics of NDBI, NDMI, NDTI and NDVI for April from the years 2020–2023 were used in this study, as shown in Table 1. The formulae for the indices are shown in Table 2.

In the first two approaches, we used the scene classification product that the Sen2Cor processor calculates alongside the S2 Level-2A product for cloud masking (Richter et al., 2012). We chose to filter out saturated or defective pixels, cloud shadows, clouds on medium and high probability, thin cirrus, and snow (classes 0, 1, 3, 8, 9, 10, and 11) from the images. In the index mosaics, the Idepix algorithm (Wevers et al., 2021) was used for cloud masking.

Synthetic aperture radar

As an active data acquisition technique, synthetic aperture radar (SAR) provides complementary information to optical remote sensing data by providing data at a different wavelength range, which is less affected by the atmosphere or variations in illumination from the Sun. In this study, operationally produced Level 2 S1 polarimetric and coherence products were used. Up until December 2021, two satellites of the Sentinel-1 constellation were orbiting 180° part, each acquiring data in the C-band (5.405 GHz) and with two polarisations: cross-polarisation VH and co-polarisation VV (Torres et al., 2012). After the failure of S1B, the number of S1 images available within the TOI dropped in 2022 and 2023, as the revisit frequency decreased by half (see Table 1). This partly guided our study design.

Table 2. The formulae of indices employed in the Sentinel-2 monthly mosaics.

Index	Formula with Sentinel-2 bands	Reference
NDBI	$\frac{B11_{SWIR1}-B08_{NIR}}{B11_{SWIR1}+B08_{NIR}} - \frac{B08_{NIR}-B04_{RED}}{B08_{NIR}+B04_{RED}}$	Zha et al. (2003)
NDMI	$(B08_{NIR} - B11_{SWIR1})/(B08_{NIR} + B11_{SWIR1})$	Gao (1996)
NDTI	$(B11_{SWIR1} - B12_{SWIR2})/(B11_{SWIR1} + B12_{SWIR2})$	Van Deventer et al. (1997)
NDVI	$(B08_{NIR} - B04_{RED})/(B08_{NIR} + B04_{RED})$	Rouse et al. (1974)

For polarimetric features, we used 11-day mean composites of VV and VH polarizations (see the acquisition dates, Table 1). Seasonal compositing harmonises the data across years. The composites were operationally produced at FMI with the sen1mosaic toolbox s (Bowers, 2020) on dual-polarised S1 imagery in the interferometric wide-swath mode. In pre-processing, the sen1mosaic uses the Sentinel Application Platform software (SNAP) (European Space Agency, 2022). The pre-processing steps included calibration, speckle noise reduction with a refined Lee filter, and geometric terrain correction with the 10 m resolution digital elevation model (DEM) from the National Land Survey of Finland (National Land Survey of Finland, 2019). For an improvement in radiometric resolution, multi-looking was applied. This, however, increased the pixel size from 10 to 20 m. See details in Supplement A.2.

For coherence features, we used 12-day coherence products from 2023 (see the acquisition dates in Table 1). The Finnish Meteorological Institute Sentinel-1 Copernicus Analysis Ready Data Coherence Processor operationally generates interferometric coherence, which indicates the phase change of the radar signal from a pair of consecutive overlapping measurements, 12 days apart. The Level 1 single-look complex products were pre-processed with SNAP software through the steps of applying an orbit file, TOPSAR[1] splits, geocoding, coherence calculation, and TOPSAR deburst (European Space Agency, 2022). The result was then merged and orthorectified using the local DEM derived from the Copernicus GLO-30 digital surface model (European Space Agency, 2019). Both the S1 ground range detected products for the composites and single-look complex products for the coherence were downloaded from the Finnish Copernicus Sentinel Collaborative Ground Station (National Satellite Data Centre, 2023) hosted by NSDC at FMI.

Feature engineering

Similar to our previous work (Luotamo et al., 2022), we considered fields as homogeneous objects, i.e. a soil management practice is applied across the whole area within the field boundaries. The mapped sensor values of a parcel were treated as a spectral distribution from which we formed a 32-bin histogram. Thus, each sensor/feature type was represented by 32 features. The range of values covered approximately 5%–95% of the total distribution per feature.

In addition to having a different feature base, the time of sensing also differs. We tested three time representations. 1) Index features were calculated from the observation date with the highest NDVI value per sensing period of 30 days (the best-available-pixel approach). 2) The band features were pooled from a 21-day sensing period as described in Table 1. Due to overlapping orbits in the ROI, parcels could be observed by more than one S2 tile within the time frame. However, we only had one observation from S2 mosaics. Therefore, we randomly chose observations parcel-wise from one S2 tile only. See class-wise numbers of observations for the mosaic (index features) and pooling (band features) approaches in Table 3 (Experiment 1). In both cases, the data were in 2D (observations, features) and the S2 data were simply merged with the S1 data. 3) Lastly, the sequence of sensing observations per parcel was treated as a time series. Each parcel would have 1. . .10 observations from a 21-day period of sensing (see Table 1). The data were sparse time series of shape 3D (observations, time, features). The S1 data were artificially reshaped into 3D by zero-padding $time - 1$ time points.

For the rest of the experiments, we used multiple S2 tiles over a parcel. Thus we might have had several observations from S2 per parcel, but only one from the S1 mosaic. As a result, the S1 observations would be

Table 3. Class-wise number of observations in the training/testing set in each experiment setting. The datasets are S1—Sentinel-1; S2—Sentinel-2; COH12—Sentinel-1 12-day interferometric coherence. CC stands for "with companion crop", ND for "no duplicates".

Setting	High till	Low till	Winter crop	Grass	Stubble	Stubble CC	Total
Experiment 1:							
S1S2, mosaic	468/156	837/279	1,276/426	1,075/358	1,178/393	131/43	4,965/1,655
S1S2, pooling	450/150	805/268	1,209/403	1,024/341	1,170/391	130/44	4,788/1,597
Experiment 2:							
S1S2 3D	626/299	1,245/458	1,670/810	1,905/463	1,694/573	–	7,140/2,603
S2 3D	694/231	1,277/426	1,860/620	1,776/592	1,700/567	–	7,307/2,436
S2 3D (ND)	450/150	805/268	1,209/403	1,023/342	1,047/349	–	4,534/1,512
S1S2, pooling	626/299	1,245/458	1,670/810	1,905/463	1,694/573	–	7,140/2,603
S2, pooling	694/231	1,277/426	1,860/620	1,777/592	1,700/567	–	7,308/2,436
Experiment 3:							
S1S2 3D	626/299	1,245/458	1,670/810	1,905/463	1,694/573	–	7,140/2,603
Experiment 4:							
S1S2+COH12	536/177	996/329	1,327/354	1,775/336	1,402/436	–	6,036/1,632

[1]Terrain Observation with Progressive Scans SAR.

duplicated within parcel-wise observations. Therefore, all duplicates were placed in the training set, so that the testing set included only truly unseen data. For comparison purposes, we also prepared 3D time series data where duplicates were completely removed. See Table 3 for the data amounts in (Experiments 2–4).

Models

We employed supervised machine learning, where the goal is to approximate a function that learns a mapping between an input **x** and an output y based on examples for which the outputs are known. In our case, the input **x** consists of satellite image data, processed as detailed in Section 2.3. and summarized in Table 3, while y represents the cropland management class. We considered several alternative models that differed in how the features are represented and combined, but all models were trained for the same task.

Neural network models

We used a feed-forward neural network (Multi-Layer Perceptron; MLP) as a baseline classifier, following the example of Luotamo et al. (2022). MLP is typically composed of multiple chained functions, i.e. hidden layers. The number and dimensionality of these hidden layers determine the depth and the width of the model (Goodfellow et al., 2016). MLP was applied to static data where the time representation is reduced to one time point.

We used a temporal convolutional neural network (TCN) (Bai et al., 2018; Lea et al., 2017) to capture the temporal nature of the data. Similar to MLP, convolutional networks consist of multiple layers that effectively learn representations of the data. TCN consists of a configuration of one-dimensional fully convolutional network and causal convolutions. The causality constraint implies that, at each layer, the transformation at a given time step relies solely on the current and past units of the preceding layer. For each experiment, the hyperparameters and network architectures were optimised heuristically. The Tensorflow (version 2.7.0, Abadi et al. (2015)) backend framework was used both for MLP (Keras version 2.7.0, Chollet et al. (2015)) and TCN (Keras implementation version 3.5.0, Remy (2020)). For interested readers, we refer to Supplement A.3, where we provide a detailed overview of the network architectures and model implementations. The supporting source codes are stored in a publicly accessible repository (see Supplemental Online Material).

Feature importance

Albeit achieving high accuracy, the non-parametric machine learning models used in this study lack interpretability due to a complex set of non-linear functions. For quantifying the strength of the relationship between the spectral features and the cropland management type in neural networks, we used a simple leave-one-out approach as a measurement of feature relevance. Each feature group is permuted, and the loss in performance is assessed when the effect of the feature group is negated. In this approach, a substantial drop in performance is indicative of an important feature group.

Another widely used method for estimating the impact of features is to use an ensemble classifier random forest (RF) which has intrinsic measurements of predictor importance (e.g. Barnes et al. (2021); Denize et al. (2019); Wei et al. (2023)). RF is a decision-tree classifier which builds on bagging (Breiman, 1996) and random feature selection at each node of a tree. Successive decision trees are induced by sampling training data with replacement, and selecting a small number of features as predictors at each split. In an ensemble of randomised decision trees, each casts a vote for plurality voting for the predicted class. The best cut at each node is selected by optimising the splitting criterion, i.e. in our case, the decrease of Gini impurity in the descendent nodes. The decreased impurity due to splits over a given feature is recorded. The total amount of decreased impurity in the ensemble of trees is then the importance measure of that feature (Breiman, 2001).

We used the RF implementation in sklearn-library version 1.1.2 (Pedregosa et al., 2011). The tuning parameter of the number of randomly selected features at each node was set to the square root of the number of features. We set the number of trees in an ensemble to 500. The trees were grown to the maximal depth.

Class imbalance

The distribution of the classes in the dataset was highly imbalanced (see Table 3). In the classification task, we addressed this issue using stratified sampling in dataset splitting and incorporating class weights in the model fitting (see C. Chen et al. (2004) for RF and Pedregosa et al. (2011) for MLP and TCN). The accuracy of the classifiers on an unseen test set was gauged by the balanced classification accuracy (the mean of true positives).

Results

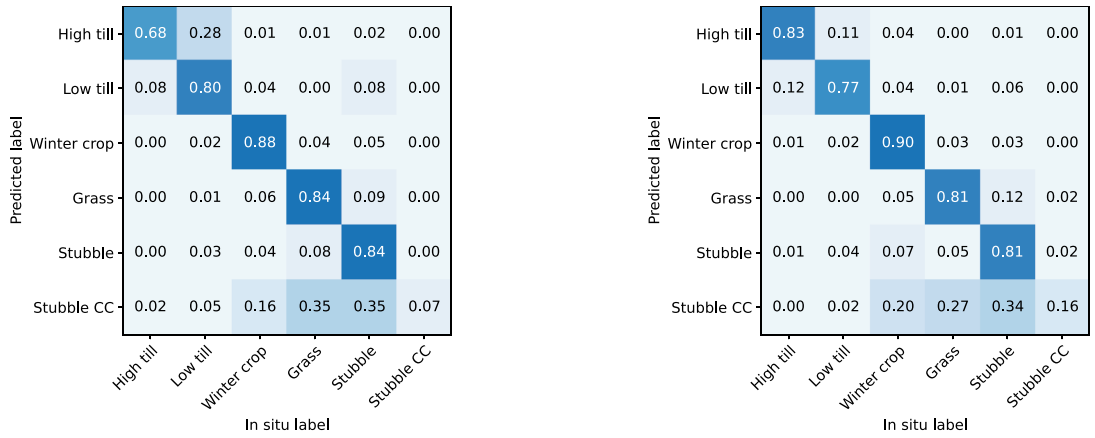
Preliminary data investigation (Experiment 1)

In the preliminary experiment, we instantiated the simplest possible model, MLP, with two slightly different but simple feature representations that both ignored time. The purpose was to see how the two alternative feature representations could identify classes. At the model level, the balanced accuracy of the spectral band – based MLP model (0.71) outperformed that of the pre-calculated S2 differential

index – based MLP model (0.68). However, at the class level, model superiority varied (see the confusion matrices in Figure 3(a,b)).

Differential indices were calculated from the date where a pixel had the highest NDVI value within the time period under observation (1–30 April). Figure 4(a) shows the distribution of the pixel-wise image acquisition dates in April 2023 delineated by the parcels in the sample. Presumably, the sensing dated from clear sky conditions at the end of the sensing period. However, the acquisition dates already began on April 6, and were quite evenly distributed over the following three weeks. More specifically, 68.3% of the image acquisition dates fell within the period of April 1–21, which overlapped with that of the spectral band – based approach. Figure 4(b) shows the 20th–80th percentile distribution of the parcel-wise temporal distances between the earliest and latest image acquisition dates. This confirms that, within a parcel, the majority of observations may have been dated several days apart, causing high within-parcel variability in the differential index values.

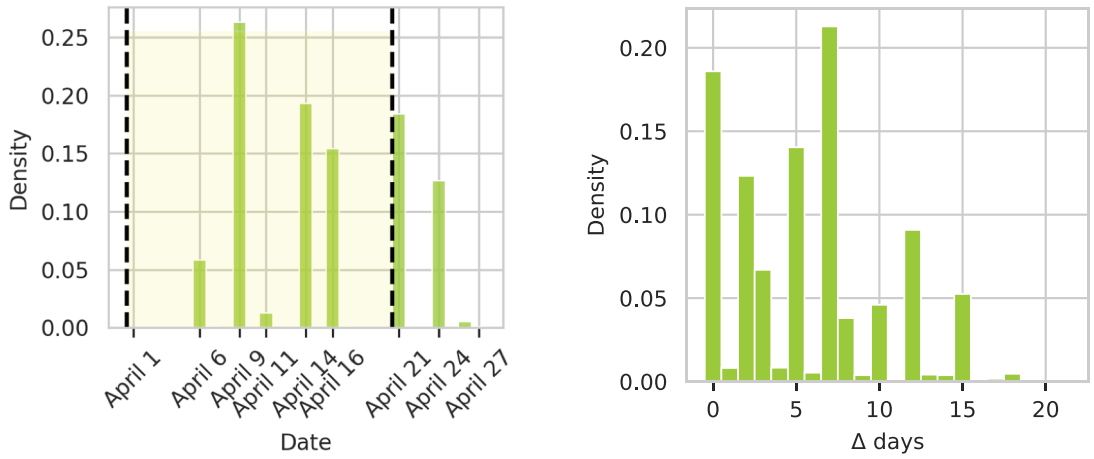
Nevertheless, the photosynthetically active classes winter crop and grass had a similar accuracy under both model settings (see Figure 3(a,b)). Stubble was slightly more confused with grass by the index-based model. Stubble can also have quite extensive weed growth underneath, which make it resemble the photosynthetically active classes. The index-based model had strikingly lower class-wise accuracy for the high till, where the model confused low till, but also winter crop, grass, and stubble, with high till.



(a) With differential index–based features. The balanced accuracy was 0.68.

(b) With spectral band–based features. The balanced accuracy was 0.71.

Figure 3. Confusion matrices for multilayer perceptron models in six-class case (Experiment 1). CC stands for "with companion crop".



(a) The dates used in calculating the 30-day Sentinel-2 index mosaics.

(b) The 20–80% distribution of the parcel-wise temporal distances between the earliest and latest image acquisition dates.

Figure 4. The normalised histograms present pixel-wise image acquisition dates in April delineated by the field parcels of the study. The shaded area between April 1–21 in Figure (a) marks the sensing period of spectral band–based features for comparison.

Along with the advancement of the growing season, the tilled soil may also have extensive weed growth, beginning to resemble winter crop growth. On the other hand, winter crop growth may suffer from winter-kill and have large areas with no photosynthetic growth.

In both settings, the stubble with companion crop had more false positives than accurate classifications. The misclassification was probably primarily due to the lack of samples in that class (see Table 1). Secondly, the class consisted of non-uniform samples resembling either stubble or photosynthetically active classes. The companion crop could also be either growing excessively and totally covering the crop residues, or the crop residues might dominate the thin dead and decomposed biomass of the companion crop.

To conclude, when the time dimension is excluded, both feature representations yielded high-performing models on classes other than high till and stubble with companion crop. For the rest of the analyses, we simplified the study setting by excluding the stubble with companion crop resulting in 5 distinct classes: high till, low till, winter crop, grass, and stubble.

Static or time series model? (Experiment 2 and 3)

Next, we tested whether the time series of S2 observations outperformed the pooled observations. TCN was harnessed to test the time series against MLP with 2D pooled data. Table 4 shows the results for the mean balanced accuracy and standard deviation for 20 iterations. The TCN-based time series slightly outperformed MLP-based pooling method (0.78±0.01 and 0.77±0.01 respectively).

We also noticed that when using only S2 observations, the accuracy was the highest of all tested settings for both TCN (0.85±0.01) and MLP (0.84±0.01). When using only S2, the modeling benefited from the full use of the data, incl. multiple observations per parcel from overlapping tiles (see Experiment 2 in Table 3). Merging S2 with S1 dropped some observations with no match in the borders of the ROI. Therefore, we also tested whether the S2-based time series still outperformed the fusion of S1 and S2 features when multiple observations (i.e. duplicates) were excluded. As a result, the S2 model outperformed the fusion model (0.80±0.01 and 0.78±0.01, respectively, see Table 4).

We also tested improving the model architecture to better utilise both types of time representations in the data (Experiment 3). The tested model architecture (TCN & MLP) did not outperform the simple fusion of two sensors based on zero-padding the single S1 observation into an artificial time series (0.76±0.01 and 0.78±0.01 respectively, see Table 4).

What is the impact of the features on accuracy? (Experiment 4)

Lastly, we had doubts as to whether our choice of S1 product (the seasonal composite) was losing important information. Therefore, we added S1 12-day interferometric coherence features to the analysis. The model scored better than the fusion model of S1 and S2 (0.82±0.01). However, the S2-based TCN remained the best-performing model.

We also studied feature-set-wise leave-one-out iterations with TCN to find out which of the feature sets affected the class-wise accuracy and importance scores from RF to more precisely determine the most important spectral ranges. The bands were grouped by their wavelength ranges as follows: red-edge and near-infrared (RENIR): S2 bands B05 red edge, B06 red edge, B07 red edge, B08 NIR, and B8A narrow NIR; visible (VIS): S2 bands B2, B3, and B4; short-wave infrared (SWIR): S2 bands B11 and B12; InSAR: S1 interferometric coherence VV and VH; SAR: S1 VV and VH.

In the leave-one-out method, we excluded one group of features at a time and monitored the class-wise accuracy metrics. Table 5 shows the results for a leave-one-out analysis with TCN. When all features were included in the model, the accuracy was 0.82, indicating that including the COH12 features improved the model performance, but did not outperform the purely S2-based model (at best 0.85 in Table 4).

Table 4. The mean balanced accuracy and standard deviations of $n = 20$ training iterations in 3 experiments: time series vs. pooled data (Experiment 2); testing different network architecture to include Sentinel-2 (S2) time series and static Sentinel-1 (S1) data (Experiment 3); fusion of 12-day interferometric coherence features (COH12) (Experiment 4). TCN with S2 as a baseline. MLP stands for multilayer perceptron, TCN for temporal convolutional network, ND for "no duplicates".

Experiment	Time	Classifier	Sensor	Accuracy
2	Pooled 2D	MLP	S2	0.84 ± 0.01
2		MLP	S1S2	0.77 ± 0.01
2	3D time series	TCN	S2	0.85 ± 0.01
2		TCN, ND	S2	0.80 ± 0.01
2		TCN	S1S2	0.78 ± 0.01
3		TCN & MLP	S1S2	0.76 ± 0.01
4		TCN	S1S2+COH12	0.82 ± 0.01

Table 5. The class-wise accuracy and total balanced accuracy results from $n = 10$ iterations of a temporal convolutional network classifier (TCN), when leaving one group of bands out at a time. The dataset is as in Experiment 4.

Omitted band	High till	Low till	Winter crop	Grass	Stubble	Accuracy
RENIR	0.67	0.78	0.86	0.85	0.76	0.78
SAR	0.76	0.77	0.90	0.81	0.76	0.80
SAR&COH12	0.77	0.76	0.89	0.81	0.77	0.80
VIS	0.76	0.79	0.85	0.86	0.78	0.81
SWIR	0.78	0.77	0.89	0.87	0.77	0.82
COH12	0.78	0.75	0.90	0.87	0.79	0.82
—	0.76	0.78	0.91	0.88	0.79	0.82

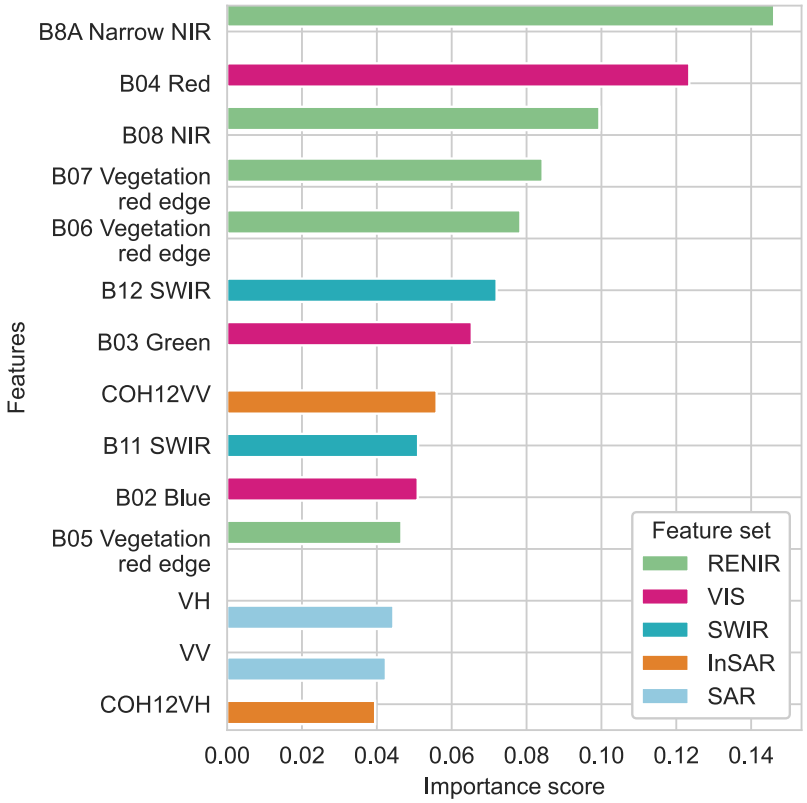


Figure 5. Variable importance from a random forest model based on the gini impurity index. Features were grouped into spectral regions. The dataset was as in Experiment 4 (2D).

In a class-wise comparison, we see that RENIR bands were essential for detecting high till soil; if the RENIR bands were excluded from the model, the accuracy in terms of detecting high till fields was the lowest (0.67). The SAR, COH12, nor SWIR bands did not make a difference in detecting high till fields. However, when coherence bands were excluded, low till was less accurately detected (0.76 or 0.75), but the overall accuracy was still quite high (0.80 or 0.82). Omitting RENIR or VIS bands reduced the accuracy of winter crop classification. In fact, VIS bands seemed to have the only notable effect on class-wise accuracy for detecting winter crops.

Figure 5 shows the order of importance for the spectral bands in the RF model. The most important bands were the NIR and upper red edge bands spanning the wavelengths ~740–865 nm and the red band at the mean wavelength of 665 nm. B12 (2190 nm) is more important than the B11 (1610 nm) part of the SWIR bands. The COH12VV band of all the SAR features had the highest importance score. VH, VV, and COH12VH were the least important features of all. The balanced accuracy of the RF model was 0.77.

Discussion

Studies on monitoring off-season cropland management practices typically focus on one aspect, e.g. discriminating the soil tillage intensity from no-tilling (Liu et al., 2022; Watts et al., 2011) or the extent of post-harvest biomass cover, such as the adoption of cover crops or the amount of crop residues (Hively et al., 2015; Nowak et al., 2021; Seifert et al., 2018). Our study strove for high granularity in cropland management classes driven by information needs. In similar classification settings, Luotamo et al. (2022) and Denize et al. (2019) have shown the highest overall classification accuracy results of 70% and 81%,

respectively, in comparison to our best result of 85% accuracy for 5 classes.

Both index-based approaches (Matthews et al., 2023; Sonmez & Slater, 2016) and band-based approaches (Wang et al., 2023) have recently been used in similar cropland management monitoring tasks. In this study, the band-based classifier outperformed the differential index – based classifier especially in discriminating bare soil from other classes, as shown in Figure 3. This suggests that the band-based features captured information that was omitted in the index-based approach. In this respect, our study was limited to four indices, none of which utilised the red edge bands. However, a vast number of indices and tools for smooth application (e.g. Montero et al. 2023) exist for vegetation monitoring. A more comprehensive set of indices could yield a higher class accuracy in this task.

A limitation of indices in themselves is that they are typically simple linear functions of different spectral regions from a sensor. In machine learning, we let the model learn these functions. Machine learning algorithms thus have the capacity to model complex relationships from higher-dimensional data. The differential indices have a long history in the remote sensing of vegetation (Bannari et al., 1995; Zeng et al., 2022). The development of indices has contained years of modelling and experimental campaigns (see, for example, Steele-Dunne et al. (2017); Xue and Su (2017)). Indices compress spectral information into interpretable proxies of a given quantity which can be used as standalone indicators (Hively et al., 2015; Nowak et al., 2021; Yue et al., 2022; Zheng et al., 2013) or as variables in statistical or process models (Dong et al., 2020; Thieme et al., 2020). They also show utility for harmonising data between different sensors, especially in long time series (Jönsson et al., 2018; Sonmez & Slater, 2016). Although our index-based machine learning model had a lower overall balanced accuracy, it had a higher class-wise accuracy for low till, grass, and stubble than the band-based model (Figure 3). This suggests that pre-processed features such as indices can alleviate the learning task, especially with shallow learners and/or small volumes of training data.

Mosaicking approaches can have limitations when applied across long time spans. The challenge of maintaining class disparity becomes evident as the within-parcel variance grows under rapidly changing growing conditions. As shown in this study, 30-day index – based mosaicking was suboptimal for the monitoring task in rapidly changing conditions, although it has been successfully applied in longer monitoring tasks such as crop type mapping (Defourny et al., 2019) or land cover classification (Nasiri et al., 2022).

Exploring the benefits of temporal data structures in machine learning can enhance classification performance. Building on the convention of minimal feature engineering in machine learning, we further examined whether the learner benefited from an asynchronous Sentinel-2 time series over pooled observations. Leveraging the recent developments in neural networks for sequence data, our results showed that the simplest single source sensor (S2) time-series approach outperformed other attempts (Table 4). This approach benefits from multiple observations per parcel from overlapping tiles. Conceptually, this is related to explicit data transformations (e.g. rotation of image patches) used in computer vision to generate diverse samples, but instead of augmenting existing samples, we expand the total data volume by extracting multiple observations from the same object. To the best of our knowledge, the literature has not revealed similar results utilising multiple time series per object in a similar task.

The fusion of radar and optical sensor data has been a topic of study. Interestingly, and contrary to our results, several studies on tillage detection have shown that the fusion of radar and optical sensor data is superior to a single data source (Denize et al., 2019; Liu et al., 2022; Whyte et al., 2018). Nevertheless, these studies are limited to static 2D datasets. As an exception, a recent tillage detection study by Azzari et al. (2019) using temporal composites showed that optical Landsat-based models outperformed S1-based composite models in accuracy and robustness. Furthermore, their feature importance analysis confirmed that S1 features had low contribution to the task. Similarly, the sensitivity analyses in our study showed that radar features had little effect on classification accuracy (see Table 5 and Figure 5).

The techniques of data fusion from different data platforms (Wang et al., 2023) or spaceborne sensors (Ienco et al., 2019; Jiao et al., 2022) is a vivid research area (see reviews, e.g. in Joshi et al. (2016); Mena et al. (2024); Sengani et al. (2023)). Our attempt to combine S2 time series and static S1 data was limited to two types of network architectures. For future studies, advances in neural networks for similar tasks (such as Zhao et al. (2020) and Gbodjo et al. (2020)) could be applied.

Recognising the benefits of SAR data in monitoring vegetation dynamics (see review in Steele-Dunne et al. (2017)), we also tested incorporating interferometric coherence features. The choices made in producing both backscatter and coherence features imposed limitations on their informativeness. Radar features are noisy due to scattering and geometric distortions (Meyer, 2019). We chose to use 11-day temporal composites as a simple noise-reduction solution. The underlying assumption was that in a lapse of 11 days the state of the actual soil cover only slightly changes. In practice, the diurnal variation in, for example, the moisture of the target can greatly vary from dry snow

to a very dry soil surface and everything moist in between (frost, dew, rain). In addition, in the ROI, the Sentinel-1 acquisitions occur either in the morning at 4–5 am or in the afternoon at 4–5 pm depending on the orbit, causing further adverse variation in the dataset. The great variation in the moisture of the target hampers the backscatter sensing (McNairn et al., 2001).

Interferometric coherence, on the other hand, is produced from two consecutive images on the same orbit, escaping the different subdaily conditions of dawn and afternoon sensing. Coherence is typically used to detect a drastic change in the target: e.g. a recent study by Voormansik et al. (2020) used S1 VV and VH 6-day interferometric coherence time series coupled with NDVI for mowing and tillage detection. They found that the VV coherence values were remarkably higher than VH coherence values after ploughing. They suggested that the structure of the exposed soil favours VV backscatter. In this study, the target was presumptively in a stable state, i.e. a drastic change was not expected. However, bare clay soil in particular tends to dry out rapidly after snow melt in the spring. As shown in Figure 5, the COH12VV rises notably high in the variable importance ranking, giving faint confidence that it can detect subtle true differences in changes between classes despite the confounding effect of the local fluctuation in moisture conditions. However, in the leave-one-out sensitivity analysis omitting both COH features was invariant to classification accuracy.

Improving SAR feature quality may enhance the ability to detect tillage practices. In order to leverage the complementary information from SAR for tillage detection, we envisage focusing on improving the quality of S1 features. Recent advances in de-speckling techniques (Singh et al., 2021) lend confidence to improved SAR image processing. Instead of temporal composites, we could apply self-supervised learning to the single-look complex SAR images within the TOI, following the work of, e.g. Molini et al. (2020).

Our study underscored the significance of spectral bands in classification performance. The sensitivity analysis (Table 5) and variable importance analysis (Figure 5) highlighted the crucial role of RENIR features. Additionally, the importance analysis revealed that the red band significantly contributed to overall accuracy (Figure 5). This result corroborates previous research where NIR, red edge, and VIS bands are included in several differential indices particularly sensitive to green vegetation, as shown in, for example, cover crop biomass estimation (Swoish et al., 2022) or leaf area index estimation (Sun et al., 2020).

The SWIR region plays a crucial role in distinguishing vegetation; however, moisture conditions can interfere with its effectiveness. The SWIR region of the electromagnetic spectrum is known to be sensitive to senescent vegetation (Zheng et al., 2015). Proximal imaging spectroscopy has shown that dry crop residues have absorption features in the 2100–2350 nm wavelength region associated with cellulose and lignin (Kokaly & Clark, 1999). However, water alters the reflectance spectra of wet crop residues, hampering their discrimination from soils (Daughtry, 2001). Therefore, highly dynamic moisture conditions in the spring may have attenuated the utility of the SWIR region for discriminating high till, low till, and stubble in our study. Nevertheless, the B12 band of the Sentinel-2 SWIR bands was ranked high in variable importance scores (Figure 5) supporting previous research as it has the central wavelength (~2190 nm) within the range of interest. However, its bandwidth of ~180 nm is rather wide (European Space Agency, 2015). Although B12 ranked highly in the importance analysis, the mutual relevance of B11 (~1610 nm with the bandwidth ~90 nm) and B12 was nullified in the sensitivity analysis. Nevertheless, narrower bandwidths in the upper-SWIR region would provide valuable discriminative information for crop residue detection (Hively et al., 2018; Quemada & Daughtry, 2016) and presumably improve the class-wise accuracy for high till, low till, and stubble.

There was an interannual imbalance in dataset size. The heavy emphasis on year 2023 in the dataset may have reduced prediction accuracy in 2020–2022, as the model had limited examples from earlier years with varying climatic and soil moisture conditions to learn from. Addressing spatiotemporal heterogeneity requires data from multiple years and locations, and a larger dataset would enhance the model's robustness.

The scope of the study was limited to monitoring soil cover status over a specific period of interest. By extending the monitoring to cover the whole off-season from harvest to the onset of the new growing season, we would not only track changes in soil cover but also the extent of the class, e.g. vegetation covered soil during off-season. Such parcel-level information would be valuable for applications such as nutrient leakage assessments (Mårtensson et al., 2023) and more accurate greenhouse gas inventories (Palosuo et al., 2025). Monitoring the amount of crop residues would lead to more precise estimates of carbon sequestration, as crop residues constitute the primary source of SOC input in agricultural soils (Jacobs et al., 2020). Moreover, data on the quantity of crop residues would improve the accuracy of nitrous oxide emission estimates (Olesen et al., 2023).

Looking ahead, a promising area for future research would be to apply the model to new geographical regions, especially those where post-harvest biomass burning is a widely practiced cropland management technique (Hall et al., 2024). This application could be of particular interest to air pollution scientists, as biomass burning significantly impacts air quality (J. Chen et al., 2017), and the method could contribute to

a more comprehensive understanding of its environmental effects. Expanding the model's geographical applicability would enhance its utility in global environmental monitoring.

Conclusions

We showed that the remote sensing-based approach can yield information on off-season cropland management practices on preferred spatial and temporal scales with 85% accuracy. The study used agricultural parcel delineations to extract information on S1 backscatter and coherence and S2 spectral reflectance. Based on a large in-situ dataset of 6,623 parcels, we showed that the highest impact on model performance comes from the spectral regions of the near-infrared and upper red edge. The best-performing model TCN included only S2 features. The classification performance was presumably hampered by rapidly changing soil moisture conditions, high local variation in photosynthetic activity, highly nonuniform units within classes, and highly uneven distribution of classes in the data.

Despite these caveats discussed, our proposed method for tillage and biomass detection generalises well in the study area of the hemiboreal and boreal environmental zones with dominantly mineral soils, as confirmed by the high classification accuracy. Further studies are needed to see how the model transfers to other environmental zones and soil types in similar agricultural systems and beyond. The model can be deployed in operational agro-environmental monitoring to support informed decision-making in various environmental, scientific, and societal domains.

Acknowledgments

We thank Anna Auvinen for data curation, Pedro Neto for software development, Dr. Matieu Henry for inspiration, Dr. Eetu Puttonen, and Dr. Mika Karjalainen for their invaluable comments at the outset of the research, and the members of the field survey team—Elias Mantere, Sanna Maula, Matts Nysand, and Ville Ruohonen—for the data collection. We are also grateful to Sari Rämö, Heikki Jalli, Juha Honkaniemi, Samuli Klemelä, Tuomo Heikkilä, and Hanna Kastikainen for their contributions to the data collection. Our gratitude also extends to Kenneth Quek for kindly proofreading this article. We would like to express our gratitude to the editor and the two anonymous referees for their valuable feedback, which greatly enhanced the manuscript through their detailed comments and suggestions. Furthermore, we would like to acknowledge the CSC—IT Center for Science Finland for computational resources and user support. We also made use of geocomputing services provided by the Open Geospatial Information Infrastructure for Research (Geoportti, urn: nbn:fi:research-infras-2016072513) funded by the Academy of Finland, CSC—IT Center for Science Finland, and other Geoportti consortium members. Sentinel-1 and -2 imagery originates from the European Copernicus Sentinel Programme.

Author contributions

Conceptualisation, M.Y., J.H., and A.K.; methodology, M. Y. and P.M.; analysis ready data preprocessing software, M. Y., P.M., M.L., and S.W.; S1 composite data preparation, Mikko S., S1 COH12 data preparation, M.H. and G.P.; S2 index data preparation, M.T.; reference data curation, M. Y. and P.M.; funding application, Mi S. and M.Y.; writing, all; supervision, J.H., A.K., P.P., and Mi S.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the European Union under Grant [101033957] and Open access funded by Helsinki University Library.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., & Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*. https://www.tensorflow.org/

Amundson, R., Berhe, A. A., Hopmans, J. W., Olson, C., Sztein, A. E., & Sparks, D. L. (2015). Soil and human security in the 21st century. *Science*, *348*(6235), 1261071. https://doi.org/10.1126/science.1261071

Azzari, G., Grassini, P., Edreira, J. I. R., Conley, S., Mourtzinis, S., & Lobell, D. B. (2019). Satellite mapping of tillage practices in the North Central US region from 2005 to 2016. *Remote Sensing of Environment*, *221*, 417–429. https://doi.org/10.1016/j.rse.2018.11.010

Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. https://doi.org/10.48550/arXiv.1803.01271

Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, *13*(1–2), 95–120. https://doi.org/10.1080/02757259509532298

Barnes, M. L., Yoder, L., & Khodaee, M. (2021). Detecting winter cover crops and crop residues in the Midwest US using machine learning classification of thermal and optical imagery. *Remote Sensing*, *13*(10), 1998. https://doi.org/10.3390/rs13101998

Bechmann, M. E., & Bøe, F. (2021). Soil tillage and crop growth effects on surface and subsurface runoff, loss of soil, phosphorus and nitrogen in a cold climate. *The Land*, *10*(1), 77. https://doi.org/10.3390/land10010077

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 1–12. https://doi.org/10.1038/sdata.2018.214

Bégué, A., Arvor, D., Bellon, B., Betbeder, J., De Abelleyra, D., Ferraz, P. D., Lebourgeois, V., Lelong, C., Simões, M. R., & Verón, S. (2018). Remote Sensing and cropping practices: A review. *Remote Sensing*, *10*(1), 99. https://doi.org/10.3390/rs10010099

Bowers, S. (2020). *SMFM Sen1mosaic*. GitHub. https://github.com/smfm-project/sen1mosaic

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1007/BF00058655

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chen, C., Liaw, A., & Breiman, L. (2004). *Using random forest to learn imbalanced data*. (Technical report No. 666). Department of Statistics, UC Berkley.

Chen, J., Li, C., Ristovski, Z., Milic, A., Gu, Y., Islam, M. S., Wang, S., Hao, J., Zhang, H., He, C., Guo, H., Fu, H., Miljevic, B., Morawska, L., Thai, P., Lam, Y. F., Pereira, G., Ding, A., Huang, X., & Dumka, U. C. (2017). A review of biomass burning: Emissions and impacts on air quality, health and climate in China. *Science of the Total Environment*, *579*, 1000–1034. https://doi.org/10.1016/j.scitotenv.2016.11.025

Chollet, F., et al. (2015). *Keras*. https://keras.io

Daughtry, C. S. (2001). Discriminating crop residues from soil by shortwave infrared reflectance. *Agronomy Journal*, *93*(1), 125–131. https://doi.org/10.2134/agronj2001.931125x

Defourny, P., Bontemps, S., Bellemans, N., Cara, C., Dedieu, G., Guzzonato, E., Hagolle, O., Inglada, J., Nicola, L., Rabaute, T., Savinaud, M., & Koetz, B. (2019). Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sensing of Environment*, *221*, 551–568. https://doi.org/10.1016/j.rse.2018.11.007

Denize, J., Hubert-Moy, L., Betbeder, J., Corgne, S., Baudry, J., & Pottier, E. (2019). Evaluation of using Sentinel-1 and -2 time-series to identify winter land use in agricultural landscapes. *Remote Sensing*, *11*(1), 37. https://doi.org/10.3390/rs11010037

Dlugoß, V., Fiener, P., Van Oost, K., & Schneider, K. (2012). Model based analysis of lateral and vertical soil carbon fluxes induced by soil redistribution processes in a small agricultural catchment. *Earth Surface Processes and Landforms*, *37*(2), 193–208. https://doi.org/10.1002/esp.2246

Dong, T., Liu, J., Qian, B., He, L., Liu, J., Wang, R., Jing, Q., Champagne, C., McNairn, H., Powers, J., Shi, Y., & Shang, J. (2020). Estimating crop biomass using leaf area index derived from Landsat 8 and Sentinel-2 data. *ISPRS Journal of Photogrammetry & Remote Sensing*, *168*, 236–250. https://doi.org/10.1016/j.isprsjprs.2020.08.003

European Commission. (2023). *Report from the Commission to the European Parliament and the Council—Summary of CAP strategic plans for 2023-2027: joint effort and collective ambition. COM/2023/707 final*. Retrieved April 26, 2024, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52023DC0707

European Space Agency. (2015). *Sentinel-2 user handbook [Computer software manual]*. Retrieved December 1, 2023, from https://sentinel.esa.int/documents/247904/685211/Sentinel-2User Handbook

European Space Agency. (2019). *Copernicus DEM-global and European digital elevation model (COP-DEM)*. Retrieved November 26, 2023.

European Space Agency. (2022). *SNAP-ESA Sentinel application platform v9.0.0*. GitHub. Retrieved March 26, 2024, from https://github.com/senbox-org/

European Space Agency. (2023). *Copernicus open access hub (SciHub)*. Retrieved September 1, 2023, from https://scihub.copernicus.eu/

European Union. (2018). Regulation (EU) 2018/1091 of the European Parliament and of the Council of 18 July 2018 on integrated farm statistics. *Official Journal of the European Union*, *L200*(1). Retrieved April 26, 2024, from http://data.europa.eu/eli/reg/2018/1091/oj

FAO. (2022). *FAOSTAT: Land use*. Retrieved November 26, 2023, from https://www.fao.org/3/CC2211EN/online/CC2211EN.html#tab49

Finnish Meteorological Institute. (2023). *Terminen kasvukausi*. Retrieved September 01, 2023, from https://www.ilmatieteenlaitos.fi/terminen-kasvukausi

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F.S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., & Snyder, P. K. (2005). Global consequences of land use. *Science*, *309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Fomferra, N., Böttcher, M., Zühlke, M., Brockmann, C., & Kwiatkowska, E. (2012). Calvalus: Full-mission EO cal/val, processing and exploitation services. In *2012 IEEE international geoscience and remote sensing symposium* (pp. 5278–5281). IEEE / Institute of Electrical and Electronics Engineers Incorporated.

Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Gbodjo, Y. J. E., Ienco, D., Leroux, L., Interdonato, R., Gaetano, R., & Ndao, B. (2020). Object-based multi-temporal and multi-source land cover mapping leveraging hierarchical class relationships. *Remote Sensing*, *12*(17), 2814. https://doi.org/10.3390/rs12172814

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. Retrieved March 26, 2024, from http://www.deeplearningbook.org

Hall, J. V., Argueta, F., Zubkova, M., Chen, Y., Randerson, J. T., & Giglio, L. (2024). GloCAB: Global cropland burned area from mid-2002 to 2020. *Earth System Science Data*, *16*(2), 867–885. https://doi.org/10.5194/essd-16-867-2024

Hively, W., Duiker, S., McCarty, G., & Prabhakara, K. (2015). Remote sensing to monitor cover crop adoption in southeastern Pennsylvania. *Journal of Soil and Water Conservation*, *70*(6), 340–352. https://doi.org/10.2489/jswc.70.6.340

Hively, W., Lamb, B. T., Daughtry, C. S. T., Shermeyer, J., McCarty, G. W., & Quemada, M. (2018). Mapping crop residue and tillage intensity using WorldView-3 satellite shortwave infrared residue indices. *Remote Sensing*, *10*(10), 1657. https://doi.org/10.3390/rs10101657

Ienco, D., Interdonato, R., Gaetano, R., & Ho Tong Minh, D. (2019). Combining Sentinel-1 and Sentinel-2 satellite image time series for land cover mapping via a multi-source deep learning architecture. *ISPRS Journal of Photogrammetry and Remote Sensing*, *158*, 11–22. https://doi.org/10.1016/j.isprsjprs.2019.09.016

Jacobs, A., Poeplau, C., Weiser, C., Fahrion-Nitschke, A., & Don, A. (2020). Exports and inputs of organic carbon on agricultural soils in Germany. *Nutrient Cycling in Agroecosystems*, *118*(3), 249–271. https://doi.org/10.1007/s10705-020-10087-5

Jiao, X., McNairn, H., Yekkehkhany, B., Robertson, L. D., & Ihuoma, S. (2022). Integrating Sentinel-1 SAR and Sentinel-2 optical imagery with a crop structure dynamics model to track crop condition. *International Journal of Remote Sensing*, *43*(17), 6509–6537. https://doi.org/10.1080/01431161.2022.2142077

Jönsson, P., Cai, Z., Melaas, E., Friedl, M. A., & Eklundh, L. (2018). A method for robust estimation of vegetation seasonality from Landsat and Sentinel-2 time series data. *Remote Sensing*, *10*(4), 635. https://doi.org/10.3390/rs10040635

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P., Mitchard, E. T., Reiche, J., & Waske, B. (2016). A review of the application of optical and radar remote Sensing data fusion to land use mapping and monitoring. *Remote Sensing*, *8*(1), 70. https://doi.org/10.3390/rs8010070

Kokaly, R. F., & Clark, R. N. (1999). Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment*, *67*(3), 267–287. https://doi.org/10.1016/S0034-4257(98)00084-4

Lea, C., Flynn, M. D., Vidal, R., Reiter, A., & Hager, G. D. (2017). Temporal convolutional networks for action segmentation and detection. In *2017 IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1003–1012). IEEE / Institute of Electrical and Electronics Engineers Incorporated.

Lilja, H., Uusitalo, R., Yli-Halla, M., Nevalainen, R., Väänänen, R., Tamminen, P., & Tuhtar, J. (2017). *Suomen maannostietokanta: Käyttöopas versio 1.1 (Finnish Soil Database: Manual, version 1.1) (No. 6)*. Luonnonvaraja biotalouden tutkimus. Luonnonvarakeskus (LUKE) (Natural Resources Institute Finland).

Liu, Y., Rao, P., Zhou, W., Singh, B., Srivastava, A. K., Poonia, S. P., Van Berkel, D., & Jain, M. (2022, 11). Using Sentinel-1, Sentinel-2, and Planet satellite data to map field-level tillage practices in smallholder systems. *PLOS ONE*, *17*(11), 1–17. https://doi.org/10.1371/journal.pone.0277425

Luotamo, M., Yli-Heikkilä, M., & Klami, A. (2022). Density estimates as representations of agricultural fields for remote sensing-based monitoring of tillage and vegetation cover. *Applied Sciences*, *12*(2), 679. https://doi.org/10.3390/app12020679

Mårtensson, K., Johnsson, H., & Kyllmar, K. (2023). Estimated nutrient leakage from arable land in different bioeconomy scenarios for two areas in central Sweden, determined using a leaching coefficient method. *CATENA*, *226*, 107102. https://doi.org/10.1016/j.catena.2023.107102

Matthews, F., Verstraeten, G., Borrelli, P., & Panagos, P. (2023). A field parcel-oriented approach to evaluate the crop cover-management factor and time-distributed erosion risk in Europe. *International Soil and Water Conservation Research*, *11*(1), 43–59. https://doi.org/10.1016/j.iswcr.2022.09.005

McNairn, H., Duguay, C., Boisvert, J., Huffman, E., & Brisco, B. (2001). Defining the sensitivity of multi-frequency and multi-polarized radar backscatter to post-harvest crop residue. *Canadian Journal of Remote Sensing*, *27*(3), 247–263. https://doi.org/10.1080/07038992.2001.10854941

Mena, F., Arenas, D., Nuske, M., & Dengel, A. (2024). Common practices and taxonomy in deep multiview fusion for remote sensing applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 4797–4818. https://doi.org/10.1109/JSTARS.2024.3361556

Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Mücher, C. A., & Watkins, J. W. (2005). A climatic stratification of the environment of Europe. *Global Ecology and Biogeography*, *14*(6), 549–563. https://doi.org/10.1111/j.1466-822X.2005.00190.x

Meyer, F. (2019). Spaceborne synthetic aperture radar: Principles, data access, and basic processing techniques. In Africa Ixmucane Flores-Anderson, Kelsey E. Herndon, Rajesh Bahadur Thapa, Emil Cherrington (Eds.), *Synthetic aperture radar (SAR) handbook: Comprehensive methodologies for forest monitoring and biomass estimation* (pp. 21–64). SERVIR Global Science Coordination Office Huntsville.

Molini, A. B., Valsesia, D., Fracastoro, G., & Magli, E. (2020). *Speckle2Void: Deep self-supervised SAR despeckling with blind-spot convolutional neural networks*.

Montero, D., Aybar, C., Mahecha, M. D., Martinuzzi, F., Söchting, M., & Wieneke, S. (2023). A standardized catalogue of spectral indices to advance the use of remote sensing in Earth system research. *Scientific Data*, *10*(1). https://doi.org/10.1038/s41597-023-02096-0

Nasiri, V., Deljouei, A., Moradi, F., Sadeghi, S. M. M., & Borz, S. A. (2022). Land use and land cover mapping using Sentinel-2, Landsat-8 satellite images, and Google Earth Engine: A comparison of two composition methods. *Remote Sensing*, *14*(9), 1977. https://doi.org/10.3390/rs14091977

National Land Survey of Finland. (2019). *Elevation model, 10 m x 10 m*. http://urn.fi/urn:nbn:fi:csc-kata00001000000000000622

National Satellite Data Centre. (2023). *Copernicus Sentinel collaborative ground station (FinHub)*. Retrieved September 1, 2023, from https://finhub.nsdc.fmi.fi

Nowak, B., Marliac, G., & Michaud, A. (2021, May). Estimation of winter soil cover by vegetation before spring-sown crops for mainland France using multispectral satellite imagery. *Environmental Research Letters*, *16*(6), 064024. https://doi.org/10.1088/1748-9326/ac007c

Olesen, J. E., Rees, R. M., Recous, S., Bleken, M. A., Abalos, D., Ahuja, I., Butterbach-Bahl, K., Carozzi, M., De Notaris, C., Ernfors, M., Haas, E., & Topp, C. F. E. (2023). Challenges of accounting nitrous oxide emissions from agricultural crop residues. *Global Change Biology*, *29*(24), 6846–6855. https://doi.org/10.1111/gcb.16962

Palosuo, T., Heikkinen, J., Hilasvuori, E., Kulmala, L., Launiainen, S., Lehtilä, A., Leinonen, I., Liimatainen, M., Salminen, M., Shurpali, N., Silfver, T., & Liski, J. (2025). Demands and possibilities for field-scale estimation of agricultural greenhouse gas balances. *Catena*, *249*, 108649. https://doi.org/10.1016/j.catena.2024.108649

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. https://doi.org/10.48550/arXiv.1201.0490

Quemada, M., & Daughtry, C. S. T. (2016). Spectral indices to improve crop residue cover estimation under varying moisture conditions. *Remote Sensing*, *8*(8), 660. https://doi.org/10.3390/rs8080660

Remy, P. (2020). *Temporal convolutional networks for Keras*. GitHub. Retrieved March 26, 2024, from https://github.com/philipperemy/keras-tcn

Richter, R., Louis, J., & Müller-Wilm, U. (2012). Sentinel-2 MSI-Level 2A products algorithm theoretical basis document. *European Space Agency, (Special Publication) ESA SP: Paris, France, 2012*, *49*(0), 1–72.

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publications*, *351*(1), 309.

Seifert, C. A., Azzari, G., & Lobell, D. B. (2018, June). Satellite detection of cover crops and their effects on crop yield in the Midwestern United States.

Environmental Research Letters, *13*(6), 064033. https://doi.org/10.1088/1748-9326/aac4c8

Sengani, D., Ramoelo, A., & Archer, E. (2023). A review of fusion framework using optical sensors and synthetic aperture radar imagery to detect and map land degradation and sustainable land management in the semi-arid regions. *Geocarto International*, *38*(1), 2278325. https://doi.org/10.1080/10106049.2023.2278325

Singh, P., Diwakar, M., Shankar, A., Shree, R., & Kumar, M. (2021). A review on SAR image and its despeckling. *Archives of Computational Methods in Engineering*, *28*(7), 4633–4653. https://doi.org/10.1007/s11831-021-09548-z

Sonmez, N. K., & Slater, B. (2016). Measuring intensity of tillage and plant residue cover using remote sensing. *European Journal of Remote Sensing*, *49*(1), 121–135. https://doi.org/10.5721/EuJRS20164907

Steele-Dunne, S. C., McNairn, H., Monsivais-Huertero, A., Judge, J., Liu, P.-W., & Pap- Athanassiou, K. (2017). Radar remote sensing of agricultural canopies: A review. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, *10*(5), 2249–2273. https://doi.org/10.1109/JSTARS.2016.2639043

Sun, Y., Qin, Q., Ren, H., Zhang, T., & Chen, S. (2020). Red-edge band vegetation indices for leaf area index estimation from Sentinel-2/MSI imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(2), 826–840. https://doi.org/10.1109/TGRS.2019.2940826

Swoish, M., Da Cunha Leme Filho, J. F., Reiter, M. S., Campbell, J. B., & Thomason, W. E. (2022). Comparing satellites and vegetation indices for cover crop biomass estimation. *Computers and Electronics in Agriculture*, *196*, 106900. https://doi.org/10.1016/j.compag.2022.106900

Thieme, A., Yadav, S., Oddo, P. C., Fitz, J. M., McCartney, S., King, L., Keppler, J., McCarty, G. W., & Hively, W. D. (2020). Using NASA Earth observations and Google Earth Engine to map winter cover crop conservation performance in the Chesapeake Bay watershed. *Remote Sensing of Environment*, *248*, 111943. https://doi.org/10.1016/j.rse.2020.111943

Törmä, M. (2020). *Sentinel-2 image index mosaics (s2ind)*. Retrieved February 16, 2024, from https://ckan.ymparisto.fi/dataset/sentinel-2-image-index-mosaics-s2ind-sentinel-2-kuvamosaiikit-s2ind

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I.N., & Rostan, F. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, *120*, 9–24. https://doi.org/10.1016/j.rse.2011.05.028

Van Deventer, A., Ward, A., Gowda, P., & Lyon, J. (1997). Using thematic mapper data to identify contrasting soil plains and tillage practices. *Photogrammetric Engineering and Remote Sensing*, *63*(1), 87–93.

Van Oost, K., Govers, G., Quine, T. A., Heckrath, G., Olesen, J. E., De Gryze, S., & Merckx, R. (2005). Landscape-scale modeling of carbon cycling under the impact of soil redistribution: The role of tillage erosion. *Global Biogeochemical Cycles*, *19*(4). https://doi.org/10.1029/2005GB002471

Voormansik, K., Zalite, K., Sünter, I., Tamm, T., Koppel, K., Verro, T., Brauns, A., Jakovels, D., & Praks, J. (2020). Separability of mowing and ploughing events on short temporal baseline Sentinel-1 coherence time series. *Remote Sensing*, *12*(22), 3784. https://doi.org/10.3390/rs12223784

Wang, S., Guan, K., Zhang, C., Zhou, Q., Wang, S., Wu, X., Jiang C., Peng B., Mei W., Li K., Li Z., & Ma, Z. (2023). Cross-scale sensing of field-level crop residue cover: Integrating field photos, airborne hyperspectral imaging, and satellite data. *Remote Sensing of Environment*, *285*, 113366. https://doi.org/10.1016/j.rse.2022.113366

Watts, J. D., Powell, S. L., Lawrence, R. L., & Hilker, T. (2011). Improved classification of conservation tillage adoption using high temporal and synthetic satellite imagery. *Remote Sensing of Environment*, *115*(1), 66–75. https://doi.org/10.1016/j.rse.2010.08.005

Wei, P., Ye, H., Qiao, S., Liu, R., Nie, C., Zhang, B., Song, L., & Huang, S. (2023). Early crop mapping based on Sentinel-2 time-series data and the random forest algorithm. *Remote Sensing*, *15*(13), 3212. https://doi.org/10.3390/rs15133212

Wevers, J., Müller, D., Scholze, J., Kirches, G., Quast, R., & Brockmann, C. (2021, December). *IdePix for Sentinel-2 MSI algorithm theoretical basis document*. Zenodo.

Whyte, A., Ferentinos, K. P., & Petropoulos, G. P. (2018). A new synergistic approach for monitoring wetlands using Sentinels -1 and 2 data with object-based machine learning algorithms. *Environmental Modelling & Software*, *104*, 40–54. https://doi.org/10.1016/j.envsoft.2018.01.023

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, *2017*, 1–17. https://doi.org/10.1155/2017/1353691

Yue, J., Fu, Y., Guo, W., Feng, H., & Qiao, H. (2022). Estimating fractional coverage of crop, crop residue, and bare soil using shortwave infrared angle index and Sentinel-2 MSI. *International Journal of Remote Sensing*, *43*(4), 1253–1273. https://doi.org/10.1080/01431161.2022.2032454

Zeng, Y., Hao, D., Huete, A., Dechant, B., Berry, J., & Chen, J. M., Joiner, J., Frankenberg, C., Bond-Lamberty, B., Ryu, Y., & Xiao, J. (2022). Optical vegetation indices for monitoring terrestrial ecosystems globally. *Nature Reviews Earth and Environment*, *3*(7), 477–493. https://doi.org/10.1038/s43017-022-00298-5

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, *24*(3), 583–594. https://doi.org/10.1080/01431160304987

Zhao, W., Qu, Y., Chen, J., & Yuan, Z. (2020). Deeply synergistic optical and SAR time series for crop dynamic monitoring. *Remote Sensing of Environment*, *247*, 111952. https://doi.org/10.1016/j.rse.2020.111952

Zheng, B., Campbell, J. B., Serbin, G., & Daughtry, C. (2013). Multitemporal remote sensing of crop residue cover and tillage practices: A validation of the minNDTI strategy in the United States. *Journal of Soil and Water Conservation*, *68*(2), 120–131. https://doi.org/10.2489/jswc.68.2.120

Zheng, B., Campbell, J. B., Serbin, G., Daughtry, C. S., McNairn, H., & Pacheco, A. (2015). Remote sensing of tillage status. In P. Thenkabail (Ed.), *Land resources monitoring, modeling, and mapping with remote sensing* (1st ed., pp. 179–200). CRC Press.