



Article

Mapping Crop Types of Germany by Combining Temporal Statistical Metrics of Sentinel-1 and Sentinel-2 Time Series with LPIS Data

Sarah Asam ^{1,*} , Ursula Gessner ¹, Roger Almengor González ¹, Martina Wenzl ¹, Jennifer Kriese ¹
and Claudia Kuenzer ^{1,2}

¹ German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), 82234 Wessling, Germany; ursula.gessner@dlr.de (U.G.); rogeralmengor@gmail.com (R.A.G.); martina.wenzl@dlr.de (M.W.); jennifer.kriese@dlr.de (J.K.); claudia.kuenzer@dlr.de (C.K.)

² Institute of Geography and Geology, University of Wuerzburg, 97074 Wuerzburg, Germany

* Correspondence: sarah.asam@dlr.de

Abstract: Nationwide and consistent information on agricultural land use forms an important basis for sustainable land management maintaining food security, (agro)biodiversity, and soil fertility, especially as German agriculture has shown high vulnerability to climate change. Sentinel-1 and Sentinel-2 satellite data of the Copernicus program offer time series with temporal, spatial, radiometric, and spectral characteristics that have great potential for mapping and monitoring agricultural crops. This paper presents an approach which synergistically uses these multispectral and Synthetic Aperture Radar (SAR) time series for the classification of 17 crop classes at 10 m spatial resolution for Germany in the year 2018. Input data for the Random Forest (RF) classification are monthly statistics of Sentinel-1 and Sentinel-2 time series. This approach reduces the amount of input data and pre-processing steps while retaining phenological information, which is crucial for crop type discrimination. For training and validation, Land Parcel Identification System (LPIS) data were available covering 15 of the 16 German Federal States. An overall map accuracy of 75.5% was achieved, with class-specific F1-scores above 80% for winter wheat, maize, sugar beet, and rapeseed. By combining optical and SAR data, overall accuracies could be increased by 6% and 9%, respectively, compared to single sensor approaches. While no increase in overall accuracy could be achieved by stratifying the classification in natural landscape regions, the class-wise accuracies for all but the cereal classes could be improved, on average, by 7%. In comparison to census data, the crop areas could be approximated well with, on average, only 1% of deviation in class-specific acreages. Using this streamlined approach, similar accuracies for the most widespread crop types as well as for smaller permanent crop classes were reached as in other Germany-wide crop type studies, indicating its potential for repeated nationwide crop type mapping.

Keywords: agriculture; random forest classification; multispectral data; radar data; spectral statistics; temporal statistics; IACS



Citation: Asam, S.; Gessner, U.; Almengor González, R.; Wenzl, M.; Kriese, J.; Kuenzer, C. Mapping Crop Types of Germany by Combining Temporal Statistical Metrics of Sentinel-1 and Sentinel-2 Time Series with LPIS Data. *Remote Sens.* **2022**, *14*, 2981. <https://doi.org/10.3390/rs14132981>

Academic Editors: Anna Jarocińska, Adriana Marcinkowska-Ochtyra and Adrian Ochtyra

Received: 13 May 2022

Accepted: 20 June 2022

Published: 22 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

About 50% of the German land area is dedicated to agriculture, including managed and unmanaged grassland [1]. Intensification of agriculture, i.e., the overuse of fertilizers, pesticides and herbicides, agricultural expansion, and the focus on few crop types, leads to decline in soil fertility and biodiversity, to increased surface runoff, greenhouse gas emissions, water consumption, and to nitrogen and chemicals leaching to the groundwater. The agricultural sector is of high economic importance in Germany, but recent years have shown its vulnerability to climate change effects, such as more frequent heat waves, spring frosts, floods and droughts [2–6]. Therefore, in order to support sustainable land management, and as a basis for assessing climate change-related ecological, economic and

societal effects on agriculture, timely, reliable and area-wide information on agricultural land use is necessary. Even though information on grown crop types is collected for Germany at the Federal State level in the context of managing agricultural subsidies through the Land Parcel Identification System (LPIS) of the Integrated Administration and Control System (IACS), this data does not provide information on all agricultural land (areas without subsidy application are not included). It is usually neither openly accessible for all German Federal States, nor available in a timely manner.

The availability of optical and Synthetic Aperture Radar (SAR) satellite data with high spatial detail of up to 10 m and a temporal resolution of approximately 5–6 days in cloud-free conditions has improved the capacity to map crops at field level. Many approaches rely on optical satellite time series with its strong capability to track vegetation phenology in a comprehensible way. Sentinel-2 data have been used in a number of studies to map agricultural areas worldwide (e.g., [7,8]), in Europe (e.g., [9–20]), and in Germany (e.g., [21–24]) with promising accuracies, sometimes in combination with Landsat or other optical satellite data (e.g., for Europe [25] and for Germany [26–28]). A major drawback of optical data, however, is that observations of crucial phenological crop stages can be lost due to cloud cover. In this context, and particularly since the availability of Sentinel-1 acquisitions, advances have been made in the analysis of radar data for crop mapping, and in its combination with optical images, to create robust classification approaches for agricultural areas [29]. Various studies have shown the potential to map crop types based on Sentinel-1 time series without additional optical data input in Europe (e.g., [30–41]) and in Germany (e.g., [42–44]). However, particularly promising are approaches combining optical and SAR data (e.g., for Asia [45], Africa [46,47], Europe [48–55], and Germany [56–58]). Such synergistic approaches consistently showed better overall accuracies than mono-sensor classifications (e.g., [48,51,56,57]), particularly in the case of challenging cloud situations [49].

Regarding the image classification approaches used for crop type mapping, there is a tendency over the last decade towards non-parametric machine learning classifiers, such as support vector machines, Random Forest (RF), and other decision trees which do not make any a priori statistical assumption on data distribution [29]. Classifications generated with such approaches generally achieve higher accuracies than parametric classifiers (see e.g., [59]). Other promising approaches for crop type mapping comprise deep learning techniques for fusion and application to radar and optical time series (see e.g., [23,55]).

When focusing on national-scale crop type mapping in Germany, previous studies have looked into different time periods and class detail. In two studies the crop type distribution for all of Germany was mapped based on optical data for the year 2016. The first study [26] discriminated seven crop types and five non-cropland classes at 30 m spatial resolution based on best-pixel 10-day, monthly and seasonal composites of a harmonized Landsat and Sentinel-2A time series. With reference data for three German Federal States, they achieved overall accuracies of up to 81%. In the other national-scale mapping study, the authors of [22] differentiated 18 crop types and 5 non-cropland classes in Germany at 20 m spatial resolution. Input data of this study were Sentinel-2A composites with dynamically varying data-driven composite periods, which required generating multiple regionalized input datasets and classification models in order to account for the inhomogeneous optical data availability throughout Germany [22]. Overall accuracies range between 86% and 90%. A first study jointly using optical and radar time series for crop type mapping in Germany was conducted by [57], combining bi-weekly features of Sentinel-1 and Sentinel-2 data, and differentiating 15 crop types in the Federal State of Brandenburg at 10 m spatial resolution with an overall accuracy of 72%. More recently, the authors of [56] have mapped 22 crop types and 2 non-cropland classes for Germany for 2017–2019 at 10 m spatial resolution based on Landsat-8, Sentinel-2, Sentinel-1, and 39 environmental variables related to topography, temperature, and precipitation. Equidistant 5-day time series were generated from the optical data sources by filling gaps with a radial basis convolution function filter ensemble. Sentinel-1 data (VV, VH, Radar Vegetation Index (RVI), VH/VV) was used in form of

monthly mean composites. With reference data for 4–5 German Federal States, the authors of [56] achieved overall map accuracies of 78%, 79%, and 80% for the three investigated years.

The mentioned previous crop type mapping studies for Germany or other large study areas revealed that major challenges are related to strong within-class phenological variabilities that are caused by varying management practices and meteorological crop growing conditions in different years and subregions [56]. Different strategies to overcome this obstacle have been presented. A stratification or regionalization of the study area was suggested to reduce the intra-class variability [7,22,60]. Other studies suggest the additional use of meteorological variables describing crop phenology [56] or the correction of the spectral-temporal profiles of crop types to the agro-meteorological conditions [61–63]. However, these latter two approaches either rely on additional input information, local expert knowledge, or on pixel-wise adapted time series requiring complex workflows, making them computationally heavy or difficult to transfer in space and time.

In this context, the objective of this study is to investigate how large-area crop type mapping approaches can be simplified in comparison to the above-mentioned studies while maintaining high map accuracies, making them applicable for efficient, repeated and timely national-scale crop type mapping in Germany. We present a classification approach for identifying 17 main crop types in Germany in 2018 based on monthly temporal metrics of SAR from the Sentinel-1 mission and optical data from Sentinel-2. Grassland, a class that usually can be mapped with a high accuracy, is not part of the selected classification legend, potentially decreasing overall accuracy in comparison to other studies. It is tested if monthly predictors are able to distinguish crop types throughout their growing season and to pinpoint differences in their vegetative development. As further objective, the potential benefit of combining optical and radar data and the respective contributions of the two signals, so far assessed on a national scale for Germany only by [56], should be analyzed for overall mapping accuracies as well as in detail for individual crop classes, as this is only rarely assessed. To still account for the regional growth condition differences that inevitably prevail in an area as large as Germany, the effect of simple study area stratification is tested to potentially improve national scale crop type mapping. Specifically, the following research questions are assessed:

1. How does the combination of monthly multispectral and SAR features influence the overall and class-wise accuracy of a Germany-wide crop type map?
2. Could the overall or class-wise accuracies of the generated crop type map be improved through regional stratification?
3. Is it possible to rely on a simple and processing efficient approach for national-scale crop type mapping while maintaining good classification accuracies?

A special aspect of this study is the use of high-quality reference data from the LPIS parcel system for 15 of the 16 German Federal States, making the training robust and the results more representative for the entire country compared to former studies. Furthermore, the synergistic yet simplified and exclusive use of radar and multispectral data for large-area crop type mapping in Germany is investigated for the first time.

2. Materials and Methods

In this study, optical and radar data are used synergistically in a RF approach, firstly to delineate the cropland area, and, secondly, to classify crop types. An overview on the workflow is given in Figure 1.

2.1. Study Area

Of its total area of 357,587 km², approximately 52% of Germany was under agricultural use and 32.8% was arable land in 2018, with maize, winter wheat, rapeseed, and winter barley being the main crops [1]. Despite its small share in total economic output, with a total standard output of EUR 46,503 million in 2020 [64], agricultural production is of great importance in Germany for ensuring an adequate food supply for the population.

Crop type diversity is high within, as well as among, German regions. Due to maritime, continental, and alpine climates, as well as due to heterogeneous soil conditions influenced by diverse geomorphological processes, the growing conditions vary strongly. This results in regionally varying dominance of crop types, diverse cropping calendars (see Figure S1), shifts in phenophases, and different landscape structures, with, for example, larger fields in the north and smaller parcels in the south [65]. The former East German Federal States show the same north–south gradient in parcel size, but have, on average, much larger fields due to land reforms in the former German Democratic Republic. All of these aspects pose special challenges to a Germany-wide cropland mapping as they lead to high variance of the spectral-temporal signal within one crop class.

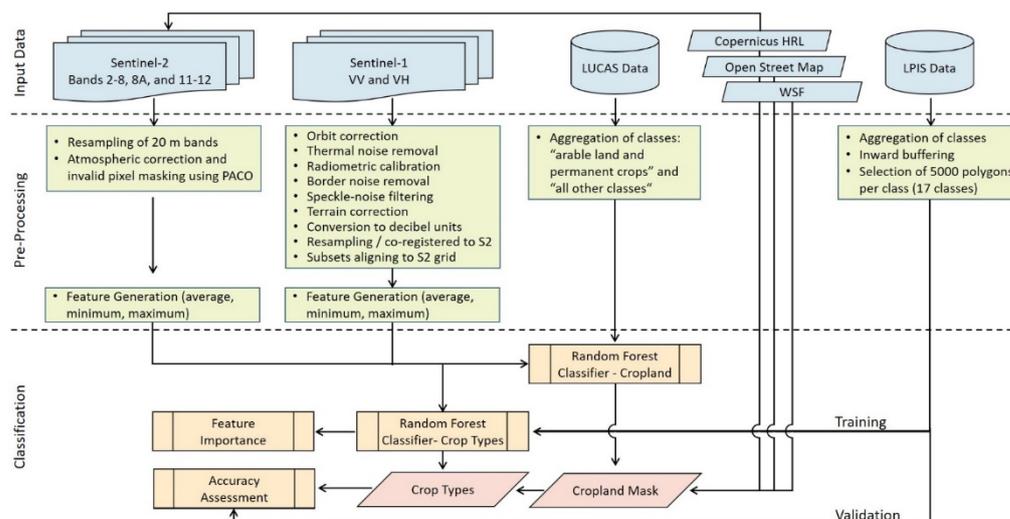


Figure 1. Workflow for the crop type classification in Germany for the year 2018 based on multispectral Sentinel-2, SAR Sentinel-1, LPIS, LUCAS, and auxiliary data.

As with most Central European agricultural systems, agricultural production and agricultural landscapes in Germany face a range of challenges, such as biodiversity conservation, the reduction of fertilizer use, and the preservation of a high quality of surface waters and groundwater, as well as achieving higher resilience to climate change [66]. Although a sound nationwide data basis is crucial to meet these challenges, the federal structure in Germany, in some parts, hampers a timely and easy compilation of such information, and, for example, LPIS data has to be gathered via the individual federal institutions. Access to ground-truth data is, hence, regarded as one of the most important constraints to crop classification at a national scale in Germany.

2.2. Satellite Data and Pre-Processing

The main input data for the classification are time series of the Sentinel-1 and Sentinel-2 missions for the year 2018. The Multi-Spectral Instrument (MSI) on Sentinel-2 provides pixel reflectances in 10, 20, and 60 m spatial resolution [67]. The red edge (bands 5, 6, 7), near infrared (band 8A), and shortwave infrared bands (bands 11 and 12), available in 20 m spatial resolution, were resampled to 10 m using the nearest neighbor algorithm. The atmospheric correction algorithm “Python-Based Atmospheric CORrection” (PACO) [68] was applied to obtain Level-2A Bottom-Of-Atmosphere (BOA) reflectance from the Level-1C data. Given only low altitudinal differences within the cropland regions of Germany, a digital elevation model was not used in the atmospheric correction process. PACO produces an additional layer for each image indicating water, clouds, cloud shadows, snow, and invalid pixels, which was used in the further processing to mask invalid data.

For Sentinel-1, all available Ground Range Detected SAR datasets for 2018 were used. We decided to use Sentinel-1 data in descending mode only, as previous studies on crop development have shown very similar behavior of ascending and descending mode

data [48]. The steps to obtain Sigma_0 backscatter values included orbit correction, thermal noise removal, radiometric calibration, border noise removal, speckle-noise filtering, terrain correction, and conversion from linear values to decibel units. These pre-processing steps were applied using the SNAP Sentinel-1 toolbox [69]. From this data base, VV and VH polarization modes were used. Other backscatter-based ratios, such as RVI and VH/VV ratio, were not considered as previous studies showed that they did not improve crop classifications as long as VV and VH were integrated [48,56]. The Sentinel-1 data were further resampled and co-registered to Sentinel-2 resolution using nearest neighbor, as suggested in [51,56,57,70], and Sentinel-1 subsets aligning to the Sentinel-2 tile grid were generated. Computationally more expensive co-registration/resampling methods (e.g., [71,72]), were not applicable given the large number of input data sets.

Subsequently, temporal features were calculated from the time series of Sentinel-1 and Sentinel-2. Temporal feature generation allows to capture average spectral changes in crops over a certain period of time, and (for optical data) to avoid data gaps due to cloud coverage. Per-pixel monthly statistics (average, minimum and maximum) were calculated for 10 Sentinel-2 bands (bands 2–8, 8A, and 11–12) and the Normalized Difference Vegetation Index (NDVI), ranging from March to October 2018. In order to avoid data gaps in few locations due to persistent cloud coverage during a month, all pixels with no observation were linearly interpolated using the respective feature value of the preceding and following months. This procedure was chosen due to its computational efficiency compared to other commonly used filters (e.g., Savitzky–Golay). To enable the interpolation for March and October, the features for February and November have also been calculated. Data gaps in these features have been filled class-wise using the mean feature value of the respective Copernicus high-resolution layer (HRL) class, i.e., imperviousness, water, grassland, and forests, or of the “incora” cropland class (see Section 2.4). For Sentinel-1 data (VV and VH polarizations), monthly per-pixel statistics (average, minimum and maximum) were calculated from January to December 2018. As C-band data are hardly affected by clouds, there was always a high number of valid observations available for calculating the backscatter temporal features within the monthly periods. Overall, a streamlined data set with a total of 336 features was generated, 264 optical and 72 radar features, based on a relatively easy procedure involving only few pre-processing steps.

2.3. Reference Data

LUCAS (Land Use/Cover Area Frame Survey) data of 2018 provided by the Statistical Office of the European Commission (Eurostat) [73] were used in a first step for the crop mask generation (see Section 2.7). In the LUCAS survey, data are collected every three years on-site at points distributed in a regular 2 km grid over the entire EU within 74 land cover and 40 land use classes (however, not differentiating crop type classes). For Germany, more than 88,000 samples are available, grouped into eight main categories. LUCAS data was used as reference to discriminate five classes: cropland (“arable land” and “permanent crops”), grassland, forests (“wooded areas” and “shrub”), inland water, and all other classes (“bare surfaces” and “artificial constructions”). For each class, 4000 reference samples were randomly selected from the LUCAS points over Germany.

Independent of the LUCAS data-based cropland mask estimation, information on cultivated crop types from the LPIS data sets was used in the main classification (see Section 2.7). Training and validation data for the crop type mapping were extracted from individual parcels of the LPIS 2018 data, that was available for 15 of the 16 German Federal States (all but North Rhine-Westphalia). Given this unprecedented exhaustive coverage of LPIS data for a Germany-wide crop classification, we expect that the procedure is also valid for the Federal State of North Rhine-Westphalia, assuming that the training and validation data from the surrounding states are sufficiently representative. LPIS data were provided for our research purposes by the Federal authorities in form of anonymized field boundaries with information on the crop types cultivated in the year 2018. The LPIS polygon data as provided individually by the German Federal States was harmonized and

prepared for its integration into the classification procedure. This included the elimination of geometric inconsistencies such as overlapping polygon areas and the harmonization of crop type class codes.

2.4. Ancillary Data

The Copernicus Pan-European HRLs Imperviousness Built-up, Water & Wetness, Grassland, and Forest Type for the year 2018 in 10 m spatial resolution [74], as well as the cropland class of the *incora* (“Inwertsetzung von Copernicus-Daten für die Raumbearbeitung”) product for the year 2019 [75], were used for the gap filling of the shoulder month (see Section 2.2).

Apart from the LUCAS data, further inputs to the crop mask were OpenStreetMap [76] layers over Germany as well as the World Settlement Footprint (WSF) map to improve the detection of all settlement areas. The WSF product [77] is a binary mask outlining the extent of human settlements at 10 m resolution based on Landsat and Sentinel-1 images.

For the stratification of the Germany-wide classification, the landscape regions (“Großlandschaften”) as defined by the German Federal Agency for Nature Conservation (by Hauke and Ssymank (not published) based on [78]) have been used. From this data set, six regions (Northwestern Lowlands—NWL, Northeastern Lowlands—NEL, Western Uplands—WUL, Eastern Uplands—EUL, Southwestern Uplands—SWUL, and Alpine Foreland—AFL) have been selected, leaving out only the marine areas as well as the alpine area, since the German Alps comprise almost no cropping areas (see Figure 2).

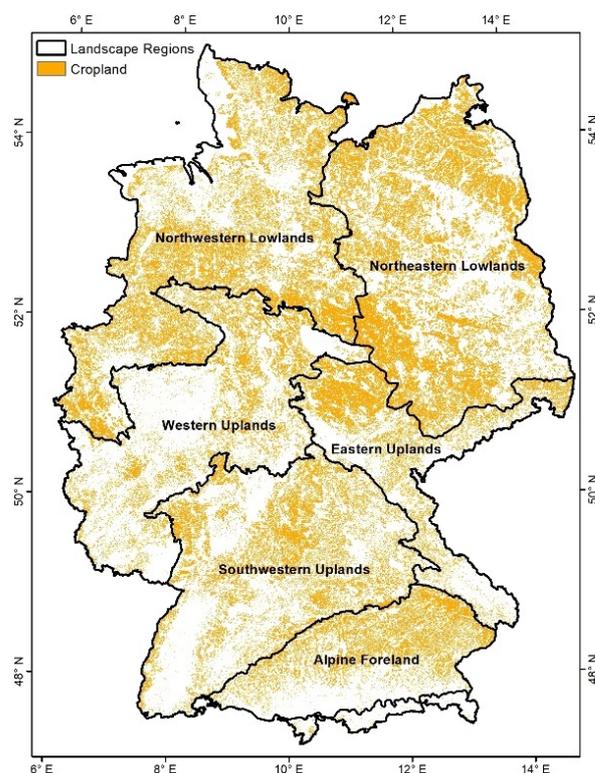


Figure 2. Cropland mask (orange) generated to delineate the crop type classification, and extent of the six landscape regions (black borders) used for the regional classification runs.

In addition to the field-level validation measures calculated with LPIS data, the areas covered by crop types according to the classification map are compared to national statistics on crop acreage published by the Federal Statistical Office [1,79–81] and the Bavarian State Research Centre for Agriculture [82].

2.5. Crop Type Classes

From the very detailed LPIS data (see Section 2.3), a selection and aggregation of LPIS classes and subclasses was conducted (see Table 1) in order to approximate the classes usually used for statistical reporting (e.g., by the Federal Statistical Office). The focus of this study is on the mapping of arable land crop types; hence, managed and unmanaged permanent grasslands (meadows and pastures), as well as fallow areas, were not included. Arable grass and clover/alfalfa, however, are kept as individual classes. Further, only crops with an area that exceeded 0.5% of the total agricultural area recorded by the LPIS data were included in the sample dataset, leading to the exclusion of industrial crops such as sunflower or vegetable classes. Although hops, vineyards, and stone fruits did not reach this threshold, they were included due to their relevance at regional level, exceeding a coverage of 10% for specific Federal States. Sub-categories of maize, winter wheat, clover/alfalfa, fruit trees, and potato, for which no potential for differentiation with multispectral and SAR data was expected, were merged, and the classes peas, beans, soy, and lupins, as well as winter triticale and winter spelt, respectively, were combined. This resulted in a total of 17 crop classes that were selected for classification (Table 1).

Table 1. Crop types analyzed in this study and their translation from LPIS classes.

Crop Type	Class Code	LPIS Classes Included
winter wheat	11	winter wheat, durum wheat
winter barley	12	winter barley
winter rye	13	winter rye
other winter cereals	14	winter triticale, winter spelt
spring wheat	21	spring wheat
spring barley	22	spring barley
spring oat	23	spring oat
maize	30	maize, maize (biogas), maize (silage), grain maize, maize with flower strip/hunting aisle
legumes	40	peas, beans, peas-beans mixtures, soy, lupins
potato	50	potatoes, starch potato, seed potato
sugar beet	60	sugar beet
rapeseed	70	winter rapeseed
clover/alfalfa	81	clover sorts, alfalfa, clover/alfalfa-grass-mixtures
arable grass	82	arable grass
vineyard	90	vineyard
fruit trees	100	stone fruits, pomaceous fruits, orchard meadows
hops	110	hops

2.6. Crop Type Sampling Methodology

For each of the above defined LPIS classes, data for training and validation were randomly sampled. A 10 m inward buffer, as well as a minimum size threshold of 0.5 ha, was applied to each LPIS parcel to exclude very small and overlapping parcels, to ensure a minimum distance between the samples, as well as to reduce noise in the spectral and backscatter information due to mixed pixel effects. From the remaining elements, a total of 5000 LPIS parcels were randomly selected per class. As shown in previous studies, training data class imbalance has a strong effect on the classification accuracy, as large classes are mapped on the expense of smaller classes [22], while equal stratified sampling avoids underrepresentation of small classes in the classification and favors the estimation of user's accuracy by reducing its standard error [83]. Therefore, an equal sampling scheme was used to avoid potential bias when training the RF models with the existing large class imbalance. From each parcel one single pixel was selected, resulting in 85,000 samples overall. At each sample point, the information from the Sentinel-1 and Sentinel-2 temporal feature raster datasets was extracted. These reference data were randomly split for each class into 50% for validation and 50% for training.

Additionally, six subsets were generated from the German-wide data set for the landscape regions [78] in order to test the effect of stratification on the classification result. Since the distribution over the landscape regions for some crop types are quite heterogeneous (such as wine or hops), classes for which fewer than 300 samples per class could be found, or which are cultivated on less than 1% of the respective region, were excluded from the regional classifications. This led to the exclusion of clover/alfalfa, vineyards, and hops from the North-western Lowlands; of spring barley, clover/alfalfa, vineyards, and hops from the North-eastern Lowland; of clover/alfalfa, arable grass, potatoes, and hops from the Western Uplands; of winter rye, potatoes, sugar beet, arable grass, vineyards, and hops from the Eastern Uplands; of arable grass and hops from the Southwestern Uplands; and to the exclusion of winter rye, spring wheat, arable grass, and vineyards from the Alpine Foreland region. For these data sets, a 50% split between training and validation data was applied as well.

2.7. Crop Type Classification Approach

The crop type classification workflow is shown in Figure 1. A supervised pixel-based RF classifier [12] was used. RF has been extensively tested for different land cover and crop type classifications (e.g., [22,26,41,56,57,60,84]) and has proven to be robust and efficient when dealing with different scales or data ranges [85], making it suitable for the joint use of SAR and optical data. Model building and classification was done using the Python “scikit-learn” package (version 0.24.1) [86].

For identifying cropland areas all over Germany, a basic land cover classification was generated as a first step that will be used as cropland mask. The spectral-temporal features of Sentinel-1 and Sentinel-2, as described in Section 2.2, as well as the LUCAS data as described in Section 2.3, served as input data to a RF with 500 trees. From the resulting classification, the classes were aggregated to two classes, “cropland” and “other”. The resulting map was further refined by masking pixels from the cropland class that were assigned to other land cover classes according to the Copernicus HRLs 2018 and the WSF. Additionally, the OpenStreetMap vector layers “shopping centers”, “airport and airfields”, “cemetery”, “commercial areas”, “industrial area”, “military fields”, “public recreational green spaces”, “parks”, “schools and universities”, “playing fields”, “sport centers”, and “residential areas” were used to eliminate remaining non-agricultural areas from the cropland class. The generated mask was validated against all available LUCAS points over Germany.

In the subsequent main classification crop types were identified. In scikit-learn, RF relies on the probabilistic prediction of a number of individual classification trees, where each tree is built using a random bootstrap sample of the training data and random subsets of features (`max_features`) at each node for identifying the best split. After testing multiple ranges of parameters, it was finally decided to build the RFs with 500 trees and to set `max_features` to the square root of the total number of features.

The RFs were built using the training data sampled from the LPIS data as described in Section 2.4, and the remote sensing features as described in Section 2.1. From the 5000 samples per class, 2500 training and 2500 validation pixels were used. To assess the effect of multispectral and SAR data on the classification accuracies, three experiments have been conducted: (i) Using all available Sentinel-1 and Sentinel-2 features ($n = 336$ input features, named “S12”), (ii) using only the Sentinel-1 features ($n = 72$, “S1”), and (iii) using only the Sentinel-2 features ($n = 264$, “S2”).

In order to assess the relevance of the input features, we calculated impurity-based feature importance scores (Gini feature importance) as implemented in scikit-learn [86]. Here, the importance of a feature is calculated as the total reduction of the Gini impurity criterion that is induced by the feature, averaged over all trees in the RF. The feature importance scores are normalized such that the scores of all used features sum to 1. The Gini feature importance is widely used to assess the relevance of features for RF approaches [51,87,88].

Additionally, six individual regional RF models were trained for the German landscape zones (Figure 2) with the respective six training and validation data subsets described in Section 2.6, but otherwise with the same parameters (model settings and input features) as used for the entire country.

A minimum mapping unit of 0.25 ha was applied by assigning small patches to the largest surrounding class using the GDAL sieve function ($n\text{Connectedness} = 4$). As last step, the crop type classification was restricted to agricultural areas by applying the cropland mask.

2.8. Accuracy Assessment

We used independent data for the validation of the resulting crop type maps. Of the randomly sampled LPIS data, 50%, i.e., for the Germany-wide classification 2500 pixels per class, was used to generate confusion matrices and to calculate the overall accuracy (OA), as well as the class-specific metrics precision (i.e., user's accuracy UA), recall (i.e., producer's accuracy PA) [83,89], and the F1-score. The F1-score is the harmonic mean of precision and recall and hence an adapted measure for imbalanced classes in a multi-class classification [90]:

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (1)$$

where precision is a measure of exactness, recall is a measure of completeness, TP are true positives, FP are false positives, and FN are false negatives.

Following the suggestions of [56,83,91], we calculated model accuracy and map accuracy for each model run, to assess both the RF model validity and the map reliability, as well as to allow for comparisons to previous studies. The model accuracy is useful to quantify the separability of classes in each RF model and, thereby, weighing the class accuracy estimates irrespectively of the class sizes, i.e., without reducing the significance of accuracy estimates for small classes. For that purpose, the above-mentioned accuracy measures are calculated based on the sample counts, i.e., the absolute number of pixels used for validation which are assigned to the different classes in the confusion matrix.

However, to estimate the probability of any given pixel in a map to be correctly classified [89], additionally the map accuracy was calculated. The map accuracy is a useful measure to quantify the reliability of the generated map, for which the area covered by each class is important. For example, if large classes are classified with a low accuracy, the probability of a given pixel being correctly classified in the map is smaller than the model accuracy might suggest. As detailed by [83], the error matrix used as basis for estimating map accuracy is, therefore, reported in terms of estimated area proportions per class. Through this approach, the class accuracies are put in relation with each crop's acreage, i.e., the proportions of each class area are used as weights to standardize their respective UAs, and, hence, to adapt the F1-scores and OA estimates. As suggested by [83], we did not derive the proportions of each class used as strata from the resulting classification map but from the LPIS data, in order to avoid introducing bias due to classification errors. This procedure was pursued for the Germany-wide as well as for the regional classifications. Class coverage information for North Rhine-Westphalia, which was not accessible for 2018, was taken as approximation from the available 2019 LPIS data.

In addition, the crop class acreages derived from the resulting crop type map are compared to the official national statistical census data on acreage estimates published by the Federal Statistical Office [79–81] and the Bavarian State Research Institute for Agriculture [82] (see Section 2.4), in order to assess the plausibility of the mapped crop areas.

3. Results

3.1. Cropland Mask

The cropland mask derived for this study is displayed in Figure 2. The generated mask achieved an OA of 85%. The relative low recall value of 0.7 of the cropland class points to a slight underestimation of cropland area, and a class-wise assessment (not shown) indicates

that most of the missing areas were assigned to the grassland class. Nevertheless, the cropland area of this mask of 115,245.82 km² compares well to the official statistics, according to which cropland and permanent crops in Germany 2018 covered 119,303 km² [1].

3.2. Overall Accuracy for Different Input Feature Sets and Overall Feature Importance

To assess the effect of optical and radar input data, three experiments using different combinations of radar and optical features have been conducted for the Germany-wide classification (see Table 2). The highest overall map accuracy (75.5%) could be achieved with the S12 run employing Sentinel-1 and Sentinel-2 data together, while using only the Sentinel-2 monthly features resulted in a map OA of 69.7%. Building the model only with Sentinel-1 features achieved the lowest map OA of 66.6%. The same tendency was observed for the estimated model accuracies with 69.9% (S12), 64.9% (S2) and 61.2% (S1). The consistently higher map vs. model accuracies indicate that the large classes could be modeled more accurately, irrespectively of the input features. Overall, the joint use of Sentinel-1 and -2 data clearly outperformed the mono-sensor approaches.

Table 2. Overall map (area adjusted) and model accuracies of the crop type classifications for the Germany-wide RF model using Sentinel-1 and Sentinel-2 features combined as input feature sets (“S12”) and using the two sensors independently (“S1” and “S2”, respectively).

	S12	S1	S2
Map accuracy [%]	75.5	66.6	69.7
Model accuracy [%]	69.9	61.2	64.9

RF feature importance quantifies the impact of individual features on the classification accuracy. Given the high number (336) of features used in this study, the discussion of importance scores of individual features is not meaningful. We, therefore, focus the presentation of our results on the importance of groups of features (Figure 3). Features are grouped considering their membership to a sensor, temporal interval, or variable (i.e., band or index) and the importance scores of the features in each group are added up. It must be noted that feature importance per category was summed over different numbers of variables, e.g., when grouping per month, 33 Sentinel-2 features were used per month, but only 6 Sentinel-1 features. Nevertheless, comparing the different groups is meaningful, as at each split in a RF only a single feature is used; hence, features with similar information content will rather share the variable importance than each having as high importance scores as if only one of them would be used. Adding them up should, therefore, approximate the feature group importance well.

Figure 3a illustrates the importance of features when regarding the mono-sensor classification approach with only Sentinel-2 data. NDVI (integrating information of the red and the near-infrared bands) is identified being the most important Sentinel-2 variable, followed by the red-edge bands (band 6 and band 5) and the green (band 3), red (band 4) and near infrared (band 8). The least importance was calculated for the blue band around 490 nm (band 2), which is usually most sensitive to atmospheric conditions so that distortions are still likely even after atmospheric correction. In the inner circle of Figure 3a, it is shown that the months of major vegetation growth (April to August) are more relevant for crop type discrimination than early spring (March), and late summer/autumn (September–October). May is the most important month, followed by July and August.

The feature importance analysis for Sentinel-1 temporal features (Figure 3b, outer circle) shows that—similar to Sentinel-2—with the start of the growing season in April, importance scores rise, peak in May, and decrease towards the late growing season in September. Sentinel-1 features of autumn, winter, and early spring (October–March) are of lower relevance (around 6%). VH-based features show a slightly higher contribution (53%) to reducing node impurity, compared to VV (47%, Figure 3c inner circle).

When analyzing the feature importance for the combined Sentinel-1 and Sentinel-2 feature space (Figure 3c), similar patterns regarding the importance of the temporal intervals (outer circle) can be observed as when using the sensors separately. However, the relative importance of the Sentinel-2 features from August and July increase slightly at the expense of the Sentinel-2 May feature importance, and the Sentinel-1 June features become more important than the radar features from May. A possible explanation could be that both sensors are able to detect the important growth stage changes occurring in May, but being used together, must share the relevance for class discrimination during that period of time, which, hence, reduces their individual importance compared to other times. Overall, Sentinel-2 feature importance (73%) outweighs the contribution of the Sentinel-1 features (27%) when used in combination (Figure 3c, inner circle).

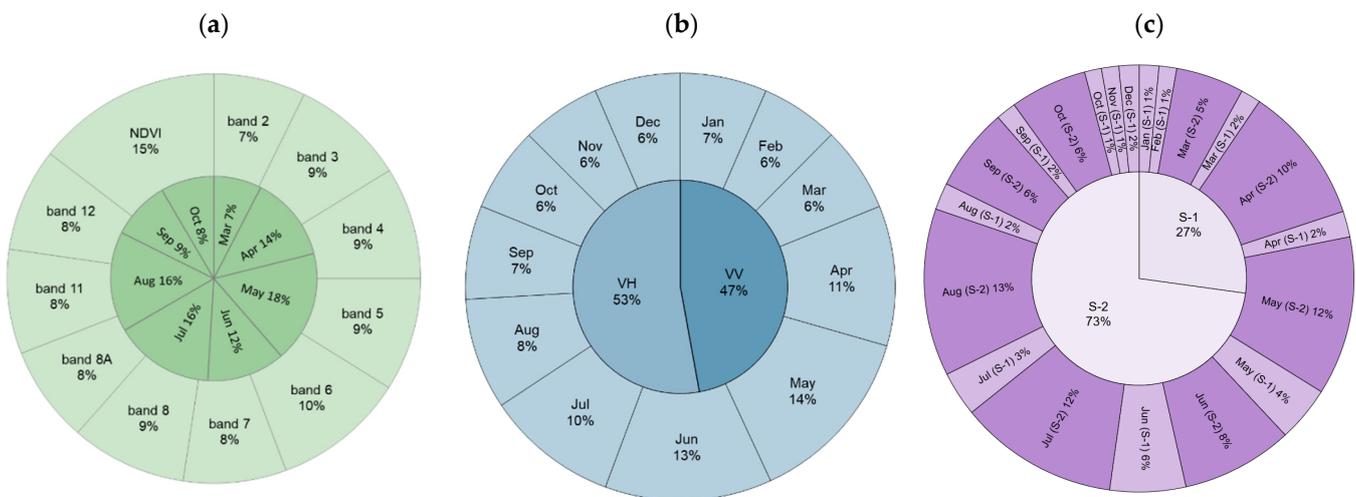


Figure 3. Sums of feature importance scores for groups of features (a) showing the importance of temporal intervals and bands/indices of the model runs with only Sentinel-2 data (S2), (b) showing the importance of intervals and polarization of the model runs with only Sentinel-1 data (S1), and (c) showing the importance of the intervals and sensor type features in the combined S12 model.

3.3. Class-Wise Accuracies and Influence of Sentinel-1 and Sentinel-2 Features on Class Separability

As shown in the previous section, the combined use of Sentinel-1 and Sentinel-2 data achieved highest overall accuracies. Accordingly, the approach combinedly using optical and radar features reached highest accuracies for most classes (Figure 4). Considering the map accuracies (dark colors in Figure 4), the S12 model achieved highest accuracies for all classes except for the classes other winter cereals and fruit trees, for which the S1 model and the S2 model achieved higher F1-scores, respectively. The pattern of the model accuracies is very similar in this regard (bright colors in Figure 4).

The S12 model reached high (≥ 0.8) map F1-scores for the classes winter wheat (0.82), maize (0.90), sugar beet (0.91), and rapeseed (0.93). Although slightly lower, the individual S1 and S2 models also achieve high (≥ 0.8) map F1-scores for these classes (apart from winter wheat, for which both mono-sensor models reach 0.75). The classes winter barley (0.76), winter rye (0.65), spring barley (0.65), potato (0.66), clover/alfalfa (0.54), arable grass (0.50), and vineyards (0.62) could be mapped with medium high (0.5–0.8) map F1-scores. Only the classes other winter cereals (0.41), spring wheat (0.32), spring oat (0.35), legumes (0.48), fruit trees (0.23), and hops (0.33) were model with low map accuracies (F1-score ≤ 0.5). When looking only at the mono-sensor approaches, in the cases of winter rye, other winter cereals, and rapeseed, the S1 model delivered higher map accuracies than the S2 model. For the remaining 14 classes, the S2 classification reached equal or higher accuracies than when using only Sentinel-1 data (Figure 4).

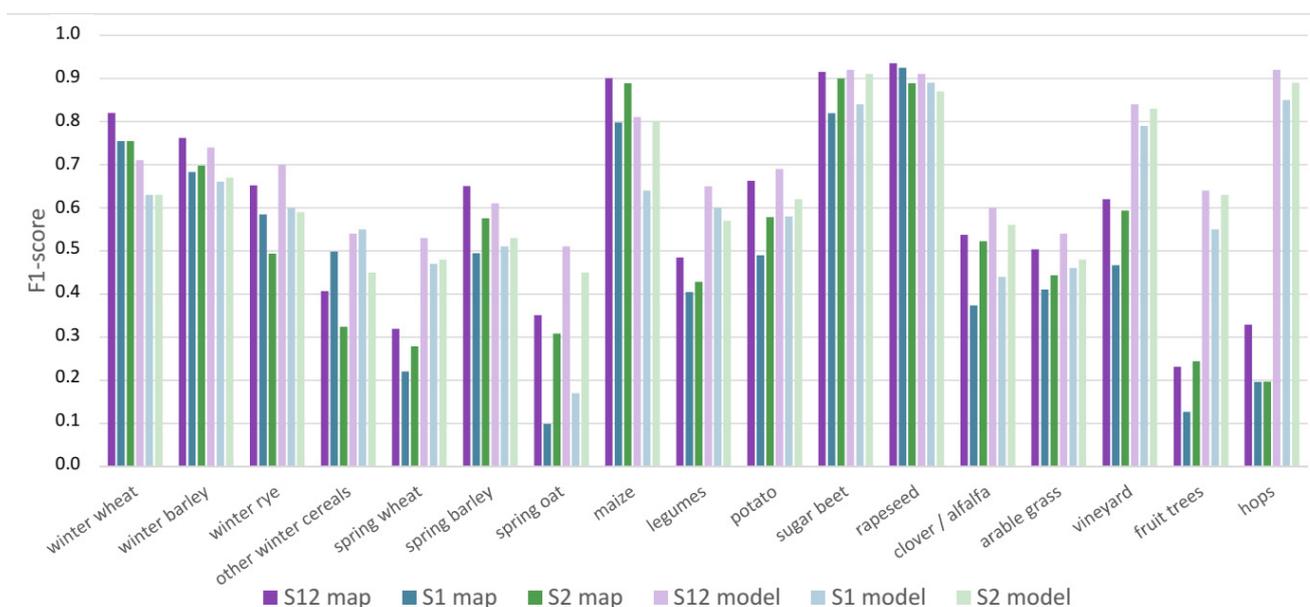


Figure 4. Class-specific map (dark colors) and model (light colors) accuracies achieved for classifications with combined use of Sentinel-1 and Sentinel-2 data (violet), of only Sentinel-1 data (blue), and of only Sentinel-2 data (green).

In addition to Figure 4, Table S1 gives an overview of the class-specific model and map accuracies, detailing precision, recall, and F1-score for the three crop type classification experiments. The class-wise differences between map and model accuracies are related to the class size (see Figure 5a for class proportions) and the type of error committed. For large classes which are rather overestimated (winter wheat and maize), the map accuracy is much higher (around 10%) than when weighing all classes equally. For large classes that are well estimated or rather underestimated (winter barley and rapeseed), the difference between the two measures is not large (around 2%). For all classes covering less than 10% of the of the agricultural area, model accuracy is higher than map accuracy. This pattern is especially striking for the three smallest classes vineyard, fruit trees and hops. While achieving high (>0.6, fruit trees) and very high (>0.8, vineyards and hops) model F1-scores, which underlines their overall good separability from other classes, all of these classes are rather overestimated on the expense of classes such as legumes, clover/alfalfa, arable grass, and cereals. While an overestimation of 10–15% in sample counts is not reducing the model UA much, a four- to six-fold area proportion overestimation (cf. fruit trees and hops) is reducing map UA of these classes drastically (see dark and bright violet colors in Figure 4, and Figure 5a for area proportions). The only deviation from that pattern is spring barley, for which a higher map accuracy was calculated. In that case, the relatively high confusion with the other spring cereals (see below) leads to a low model accuracy, but being the largest of these spring cereals classes, this confusion is not that important in the spatial domain.

To better understand the separability and confusion of single classes in the presented classification approach, model accuracy is analyzed in more detail in the following. The sample-based CM for the 17 crops classified over Germany using the S12 RF is shown in Figure 5b. Certain patterns can be observed in terms of higher commission and omission errors, highlighted by shades of orange. The most confusion occurred among the four winter cereal classes (winter wheat, winter barley, winter rye, and other winter cereals), the three spring cereals (spring wheat, spring barley, and spring oats), and the two forage crops (clover/alfalfa and arable grass). Apart from these blocks, a commission error between winter and spring wheat is apparent. Further, the potato class has a high rate of omission errors with all spring cereals and legumes, while legumes, rather, have commission errors

with the spring cereals and potato. The class of fruit trees has many commission and omission errors with the forage crops as well as with vineyards.

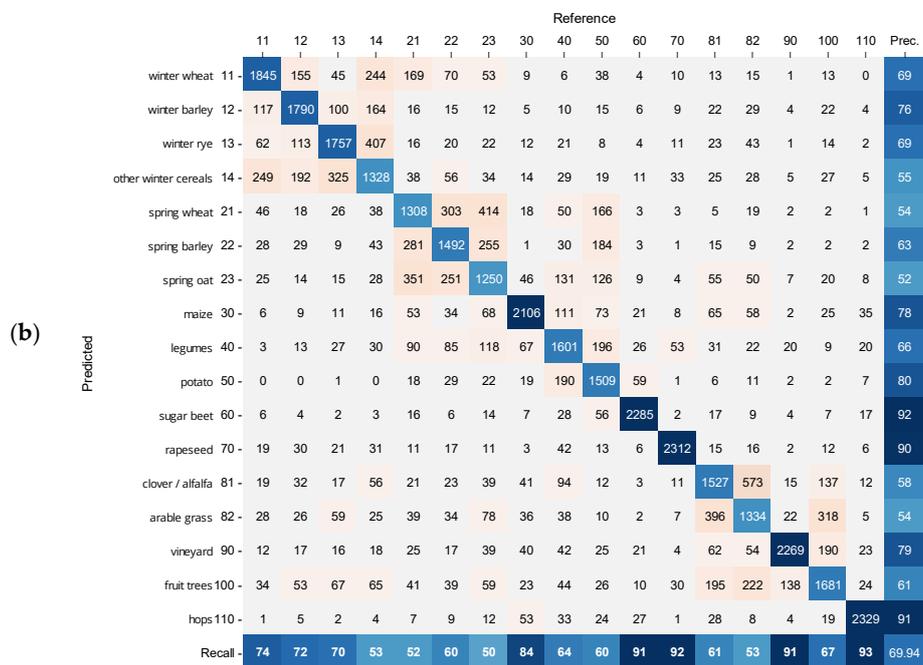
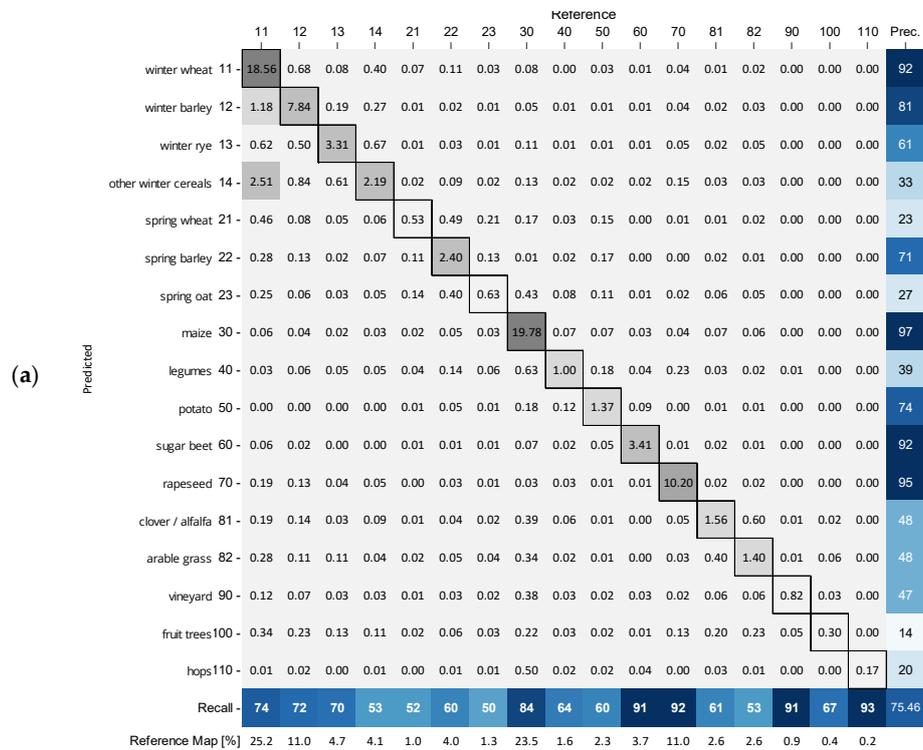


Figure 5. Confusion matrices for the “S12” crop type classification for Germany 2018 based on (a) area proportion and (b) sample counts.

Figure 6 highlights more clearly where the joint use of optical and radar data reduced confusion between crop types. The matrices in Figure 6 show the differences between pairs of the confusion matrices of the S12, S2 and S1 models. Figure 6a displays the differences in the confusion matrices for S12 minus S2. Violet colors mark cells where confusions were reduced (off-diagonal), or—for the diagonal—where correct assignments were increased by

using the S12 feature set, compared to using only optical data, i.e., where the use of Sentinel-1 benefits the result. In contrast, green colors mark cells where better results were identified for S2. The numbers in the off-diagonal cells indicate by how many validation samples the confusions were reduced, or—for the diagonal—by how many validation samples the correct assignments were increased. Likewise, Figure 6b highlights improvements in the classification for using S1 (blue colors) and for using S12 (violet colors), i.e., additionally including Sentinel-2. Figure 6c compares the two mono-sensor approaches and shows where S1 (blue colors) and S2 (green colors) delivered better results.

Figure 6a shows that confusions within all winter cereals were considerably reduced by using Sentinel-1 and -2 data instead of only Sentinel-2 (darker violet colored square in the top left corner). In contrast, Sentinel-2 based confusions of winter cereals with other crop types were not as significantly reduced by the S12 approach. For a number of class combinations, no, or even a slightly negative, effect when including Sentinel-1 data was observed (green colors). Figure 6a shows similar effects as for winter cereals for confusions within the group of spring cereals, but also for legumes and potatoes and their confusion with spring cereals, and for rapeseed (confused with cereal classes based on S2 only) and the forage classes (confusion with almost all other classes but spring cereals). Moreover, in these cases, confusions resulting from optical data could be reduced integrating Sentinel-1 data.

The improvements of the synergistic approach when compared to the S1 model show different patterns (Figure 6b). Here, the reductions in class-wise confusions are not as concentrated on phenologically similar groups of crops, but rather spread over the confusion matrix. Still, particularly strong improvements by including Sentinel-2 could be achieved in increasing the accuracies of the winter cereals (wheat, barley, rye), and of spring barley, maize, legumes, potato, and clover/alfalfa. A distinct deviation from this pattern is the class other winter cereals, where the confusion with winter rye, barley, and wheat classes increased through the inclusion of Sentinel-2 (see accordingly the F1-scores of that class in Figure 4). Further, the classes spring oat and rapeseed stand out. Spring oat is strongly underestimated in the S1 model; hence, including Sentinel-2 data increased the true positive, as well as—to a lower degree—the false positive detections, leading to a significantly higher number of correct assignments (dark violet in Figure 6b) while also increasing commission errors (dark blue colors on the off-diagonal in Figure 6b). The other way around, the class rapeseed is strongly overestimated using only Sentinel-1; hence, reducing a large share of the commission errors when including Sentinel-2 comes at the cost of a (nevertheless lower) omission error, resulting in an overall slightly lower number of correct assignments.

Figure 6c, finally, illustrates the class-wise benefit of using either only Sentinel-1 (blue colors) or Sentinel-2 (green colors). For most classes, a higher number of correctly assigned samples or fewer commission/omission errors could be achieved when using Sentinel-2. However, the translation of this confusion matrix to overall F1-score metrics is more complex, as partly the off-diagonal changes are stronger than the increase in correct class assignments (see e.g., class winter rye or spring wheat). Overall, less distinct patterns can be identified for the mono-sensor comparison.

To illustrate the influence of the sensor selection on the spatial representation of crop classes, Figure S2 displays two subsets of the maps generated in the three experiments in the Magdeburger Börde region in Saxony-Anhalt and in the Eichsfeld region in Thuringia. While the above revealed differences in class separability and frequent class confusions are not as clearly visible on such small patches as in the condensed form of the confusion matrices, it still can be seen that through adding Sentinel-1 or Sentinel-2 data, largest differences occur within the winter and spring cereal classes (see different shades of blue in the upper row subsets of Figure S2, and different shades of brown in the bottom row subsets, respectively). Large classes, such as winter wheat, maize, sugar beet, or rapeseed, are, however, not as much influenced by the selection of sensors. Most apparently, it can be concluded that the combined use of Sentinel-1 and Sentinel-2 (subsets on the right-hand), overall, increases the within field homogeneity of the crop type maps for all classes.

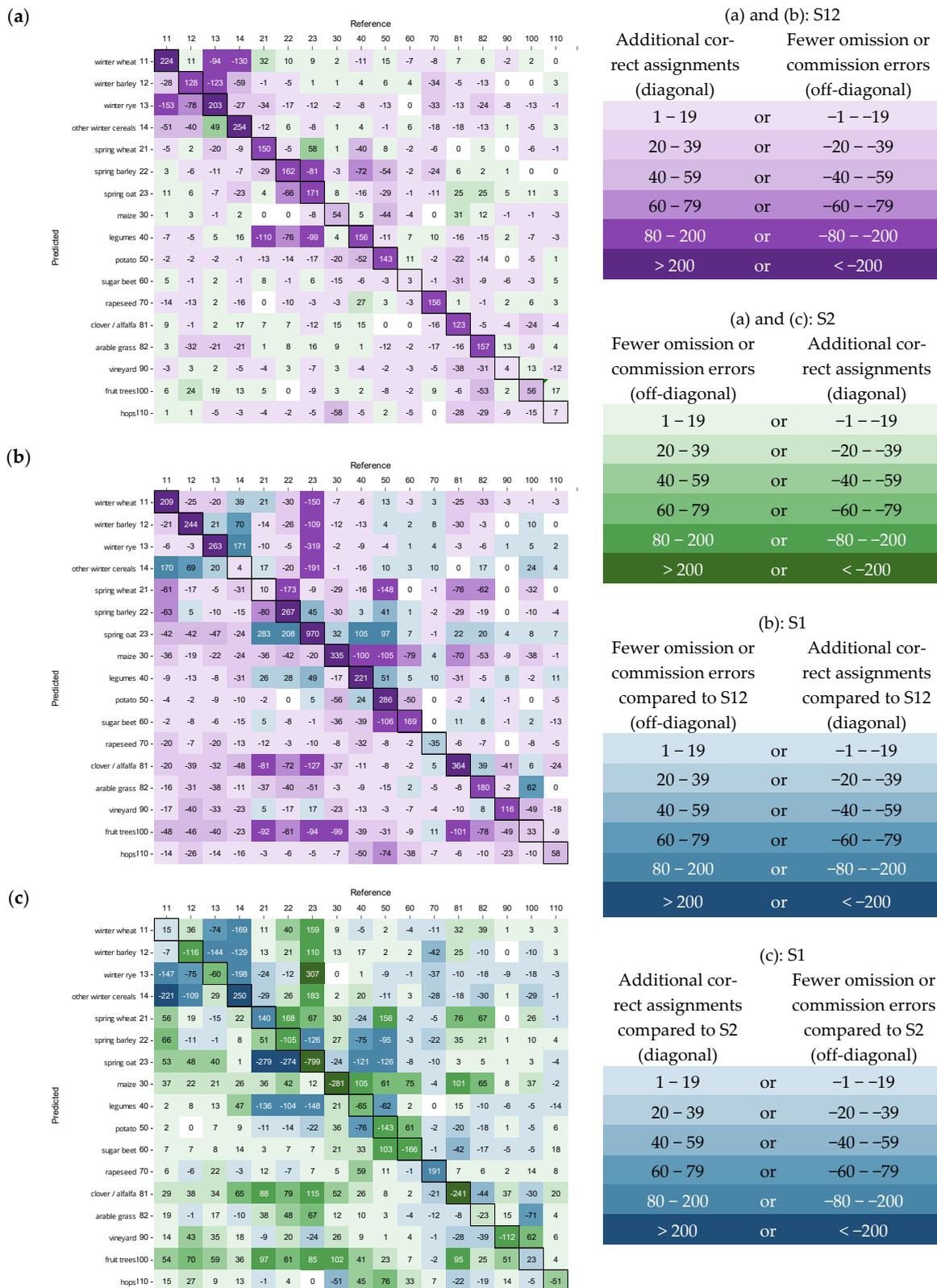


Figure 6. Sample count Confusion Matrices (CM) generated by subtracting (a) the S2 CM from the S12 CM, (b) the S1 CM from the S12 CM, and (c) the S2 CM from the S1 CM. The colors indicate if the S12 (violet) approach achieved more correct class assignments (diagonal)/less omission and commission errors (off-diagonal), or if S2 (green), or S1 (blue) features only achieved better results.

3.4. Country-Wide Crop Type Classification

The best classification model using the S12 RF (overall map accuracy of 75.5%) was used for the Germany-wide wall-to-wall crop type map 2018 (Figure 7). The map shows high spatial consistency and field parcels are overall well delineated and homogenous, i.e., most fields were classified as a single crop type (see subsets in Figure 7). This is a promising result for a pixel-based approach. The difference in size between agricultural parcels in the former East and West Germany states is well discernable (see Figure 7B). The map depicts the distribution of cropland in Germany within the fertile areas of the Northern lowland plains, the Central German Loess areas (at the northern foreland of the German Central Uplands, the so-called “Börden” characterized by high nutrient content and water holding capacity), and of the South-Western German Scarplands and the hilly landscape of Bavaria, as well as of the Loess areas and flood plains along the main rivers.

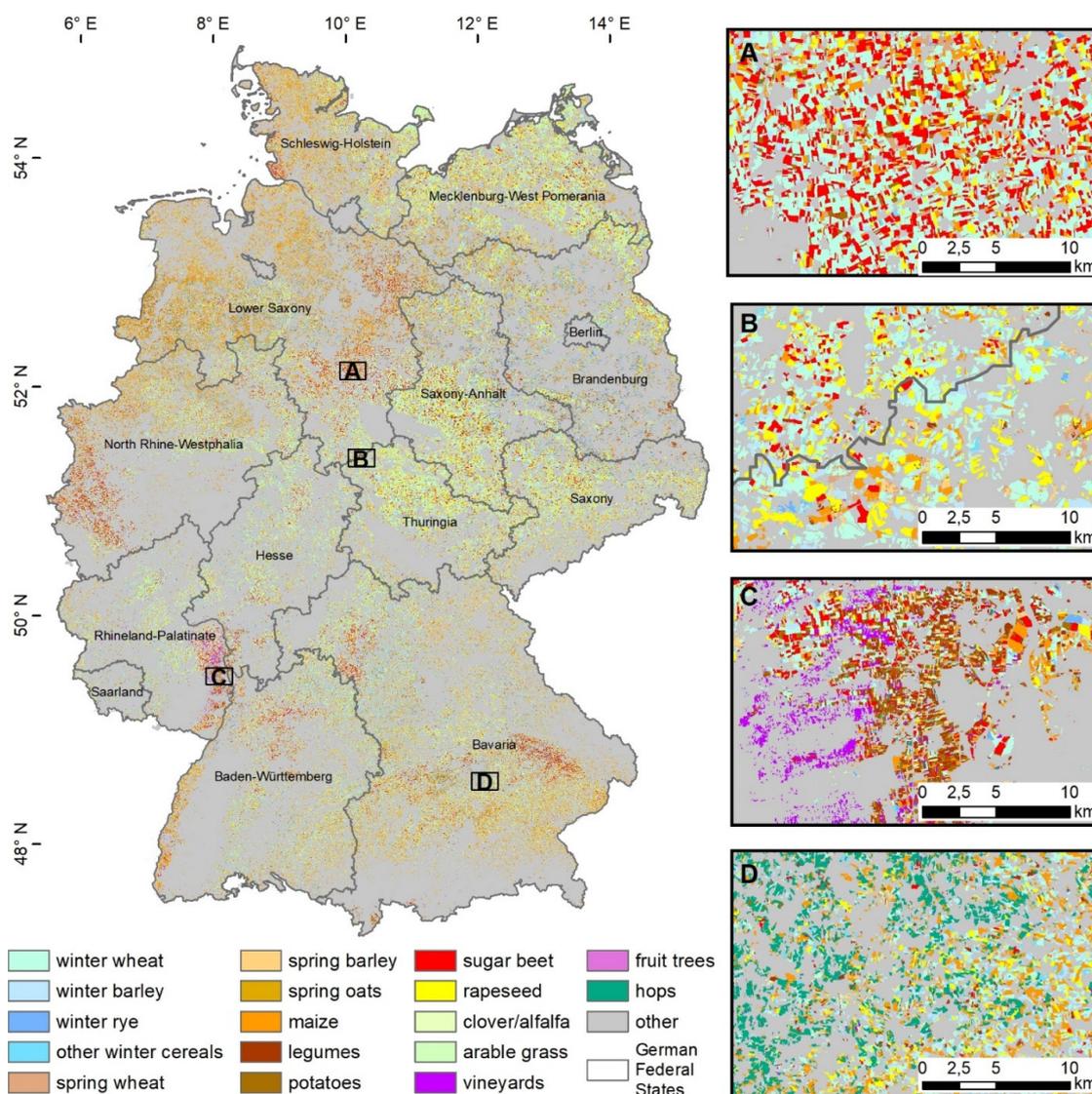


Figure 7. Crop type classification for Germany 2018 based on the S12 model. The class other comprises non-agricultural land use such as forest, urban, or water, as well as permanent grasslands. The subset maps show growing regions in (A) the very fertile and intensively used “Hildesheimer Börde” in Lower Saxony, (B) the “Eichsfeld” in the former border region of Eastern and Western Germany, (C) Rhineland-Palatinate west of Mannheim with a part of the “German Wine Route”, and (D) Bavaria between Ingolstadt and Landshut, the so-called “Hallertau” famous for hop production.

Areas that appear in shades of blue and brown indicate the prevalence of winter and spring cereals, e.g., winter cereals along the coast of the Baltic Sea and the Eastern North German Plains in Mecklenburg-Western Pomerania, in the Loess areas of Saxony, Saxony-Anhalt and Thuringia (see Figure 7B), and in parts of Hesse, Rhineland-Palatinate, and Franconia, and, e.g., spring cereals in parts of the Lower Saxony Loess area or Rhenish Hesse. Large areas of maize cultivation, marked in orange, are mapped in Bavaria, Lower Saxony, and Schleswig-Holstein. The main growing region of rapeseed is marked in yellow in the North-East and in the Central German Loess areas. Apart from these largest crop classes, growing regions of sugar beet are locally important in the Börde regions of Southern Lower Saxony, the Cologne-Aachen Lowlands, along the main rivers in Franconia (Main), Lower Bavaria (Danube), and the Northern Upper Rhine, as well as to a smaller extent in Württemberg, in Saxony-Anhalt in the Magdeburger Börde, and in Saxony around Leipzig (in red, see also subset A in Figure 7). Furthermore, high abundances of potato, e.g., in the Central North, the Emsland, the Lower Rhine region, the Palatinate region, or Lower Bavaria (dark brown color, see also subset C in Figure 7), are detected locally. Legumes are mainly grown in the Lower Rhine Bay and the Northern Upper Rhine region.

The spatial distribution of crops is further characterized by perennial crops, such as vineyards, which are mainly found in Rhineland-Palatinate (subset C in Figure 7) and to lesser extents in Baden-Württemberg and Franconia. Fruit trees are grown in the Lake Constance region of Baden-Württemberg and Bavaria and in Lower Saxony (e.g., the “Alte Land”), in the Upper Rhine region, and extensively in South-Western Rhineland-Palatinate, Saxony, and Franconia. Hops are almost exclusively cultivated in the Bavarian Hallertau region (subset D in Figure 7).

3.5. Comparison with National Agricultural Statistical Data

The classes with the largest extent in the 2018 S12 crop type map are winter wheat (24.0%), maize (22.8%), and rapeseed (11.7%). All other winter cereals make up for 19.2% of the area share, while all spring cereals together cover 9.4%. Legumes (3.1%), potatoes (2.8%), and sugar beet (4.4%) together cover about one tenth of the mapped arable land, while the fodder crops and the permanent cultures cover $\leq 1.0\%$ each.

The mapped areas align overall well with agricultural statistics at the national level, with on average 1% deviations in coverage (Figure 8). The individual classes differ from the official statistics by 0.1–2.7%. The largest percentage deviances occur for the classes winter barley (underestimated by 2.7% area share, i.e., 3444 km² in absolute area) and spring wheat (overestimated by 2.3% area share, i.e., 2389 km²). While the absolute area underestimation of winter wheat is 2699 km², this translates in only 1.2% of difference in area proportions. Further, the classes maize (0.6%), clover/alfalfa and arable grass (1.7% each), and all permanent crops (vineyards, 0.6%; fruit trees, 0.4%; and hops, 0.1%) are underestimated. The classes winter rye (0.8%), other winter cereals (1.3%), spring barley (0.3%), spring oat (0.4%), legumes (1.6%), potato (0.6%), sugar beet (0.7%) and rapeseed (0.7%) on the other hand, are overestimated in comparison to the official census data. It must be noted that the original aim of this crop type classification was not to estimate the crop area proportions, and that this comparison is solely thought to evaluate the plausibility of the map. To actually derive robust area estimates from the classification map, the accuracy could still be improved by integrating the class-wise user accuracy with map-based area estimates [83,92], which is, however, out of scope of this study.

3.6. Crop Type Classifications for the Individual Landscape Regions

Table 3 gives an overview on the crop type classification map accuracies (class-wise F1-scores and OA) achieved by the S12 RF models trained individually for the six landscape regions in Germany (see Figure 2), their class-wise averages, as well as the results of the Germany-wide classification for easier comparison. The results show that the OA is higher than the Germany-wide classification for some landscape regions (80.5% for the Alpine Foreland and 76.6% for the Northwestern Lowlands), but for some they are lower (the other

regions achieving 72.1–73.8% OA). Accordingly, the OA (74.7% averaged over all regions) could not be improved through stratification.

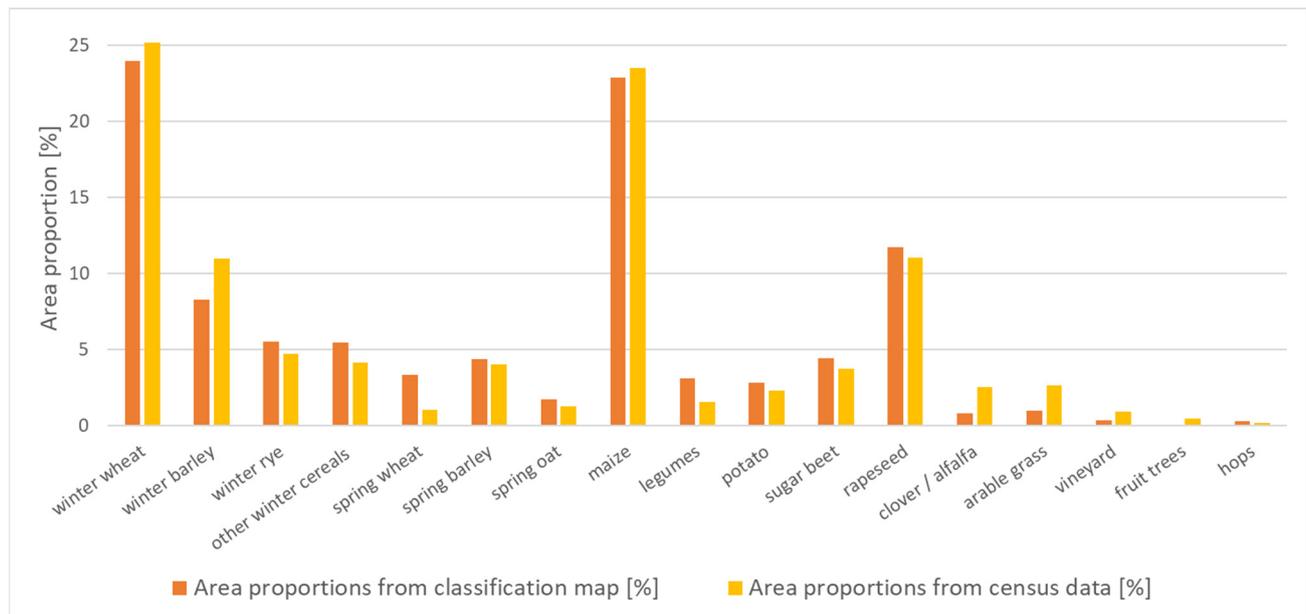


Figure 8. Comparison of area proportions estimated from the 2018 S12 crop type map for Germany to the area proportions as derived from official census data [79–82] (related to the total agricultural area in the respective data set).

Table 3. Class-wise F1-scores and overall map accuracies for the landscape regions in Germany, with Northwestern Lowlands—NWL, Northeastern Lowlands—NEL, Western Uplands—WUL, Eastern Uplands—EUL, Southwestern Uplands—SWUL, Alpine Foreland—AFL, and the accuracies of the Germany-wide classification (GER). Light colors indicate a minor deviation ($\leq 1\%$) from the Germany-wide classification, while strong colors indicate larger deviations.

Crop	Class	NWL	NEL	WUL	EUL	SWUL	AFL	Average Regions	GER
winter wheat	11	0.76	0.80	0.78	0.73	0.75	0.83	0.78	0.82
winter barley	12	0.75	0.75	0.77	0.70	0.71	0.86	0.76	0.76
winter rye	13	0.57	0.60	0.26	-	0.43	-	0.47	0.65
o. winter cereals	14	0.42	0.19	0.35	0.28	0.42	0.40	0.34	0.41
spring wheat	21	0.26	0.35	0.30	0.30	0.12	-	0.27	0.32
spring barley	22	0.54	-	0.65	0.72	0.72	0.71	0.67	0.65
spring oat	23	0.12	0.27	0.38	0.40	0.41	0.47	0.34	0.35
maize	30	0.93	0.80	0.89	0.91	0.87	0.91	0.89	0.90
legumes	40	0.54	0.49	0.51	0.70	0.53	0.56	0.56	0.48
potato	50	0.84	0.61	-	-	0.52	0.80	0.70	0.66
sugar beet	60	0.94	0.89	0.93	-	0.92	0.93	0.92	0.91
rapeseed	70	0.93	0.97	0.91	0.96	0.92	0.94	0.94	0.93
clover/alfalfa	81	-	-	-	0.69	0.69	0.62	0.67	0.54
arable grass	82	0.65	0.47	-	-	-	-	0.56	0.50
vineyard	90	-	-	0.75	-	0.81	-	0.78	0.62
fruit trees	100	0.30	0.15	0.21	0.16	0.37	0.72	0.32	0.23
hops	110	-	-	-	-	-	0.50	0.50	0.33
Overall accuracy		76.6	73.8	73.0	72.1	72.1	80.5	74.7	75.5

Averaged over all regions, class-wise metrics indicate a two-tier influence of stratification on the classification performance. The cereal classes in the regional classifications as well as on average show mostly lower F1-scores than in the Germany-wide classification, on average decreasing the class accuracies by 5%. Exceptions from this pattern are winter and spring barley as well as spring oat, which could be mapped either on average or in the majority of regions with an equal or slightly higher accuracy. The maize class gives an inconclusive picture, with very similar F1-scores on average and in the regions, compared to the national-scale classification.

All other classes, i.e., legumes, potato, sugar beet, rapeseed, the grassland classes, and the permanent classes, on the other hand, achieve higher accuracies when mapped per region. Thereby, class-wise F1-scores can vary strongly among the regions, with, e.g., fruit trees being mapped with an accuracy of 72% in the Alpine Foreland, but only 15% in the Northeastern Lowlands. Similarly, for winter rye, spring oat, or potato the regional class accuracies vary by more than 30%. Large classes, such as winter wheat, maize, or rapeseed, on the other hand, are more or less consistently mapped throughout the regions with accuracies varying only by up to 10%. Naturally, classes that only occur in certain regions, such as clover/alfalfa, vineyard, or hops, show even lower variations (up to 7%), as they are only mapped for 1–3 regions.

4. Discussion

4.1. Overall and Class-Specific Classification Accuracies

In this study we present a workflow for a national-scale crop type classification differentiating 17 crop type classes. A distinct characteristic of this study is the use of LPIS data from 15 out of 16 German Federal States, making this, to our knowledge, the first reported experiment with such a comprehensive coverage of ground-truth data for crop type classification at national level in Germany. Different from other national crop type mapping studies for Germany [22,26,56], permanent grasslands, which show a very distinct phenological development and are, thus, relatively easy to map with high accuracies (see Section 4.2), were not included as class in the map legend, while annual/arable grasslands as well as permanent crops were part of the classification.

As shown in Section 3.1, the generation of the cropland mask with an OA of 85% led to a slight underestimation of cropland area, i.e., some fields will be missing in the crop type map produced in this study. Improving the cropland mask is, however, not the primary goal of this study and should be adapted to the respective needs and facilities of the crop type map users. For example, State agencies could resort to the digital landscape model (DLM) from the German Federal Agency for Cartography and Geodesy (BKG) which is available for a fee only for external users, as suggested by [56].

An overall map accuracy of 75.5% could be reached, which compares well to other studies on crop type mapping in Germany using LPIS data, e.g., the OA of 79% [56], mapping 22 crops on a national scale; the OA of 88% [22], mapping 19 crops on a national scale; or the OA of 72% [57], mapping 15 crops in Brandenburg. However, all of these reference studies included permanent grasslands in their class legend, which was not the case in the study presented here. In two of these studies, this class, which covers approximately 37% of the agricultural area in Germany, could be mapped with high accuracies (UA \geq 84%/PA \geq 90% in [22], UA \geq 94%/PA \geq 82% in [56]), which also increases the OA, especially when investigating map accuracies. Moreover, the mapped areas of the crop classes in this study align, overall, well with official agricultural statistics at the national level, with on average 1% deviation in coverage.

Looking at the individual class accuracies, high map F1-scores were achieved for the largest classes winter wheat, winter barley, maize, sugar beet, and rapeseed. The sample-based confusion matrix (Figure 5) shows that greatest confusion occurred among morphologically and phenologically similar species within the winter and spring cereals groups, respectively. Further, legumes and potatoes are often confused with spring cereals, as their management and phenological stages are often in parallel. This mix-up was also

reported by [26]. Further, the fodder crop classes clover/alfalfa and arable grass, which achieve F1-scores around 0.5, show high rates of confusion among each other, which can be explained by the very similar herbaceous plant cover. The classes of fruit trees and vineyards of the S12 model results have high commission and omission errors with the forage crops. This was also observed by [26] and might be explained by a high share of herbaceous cover on many of these sites, but is probably also induced by the highly diverse canopy structure of fruit trees. In addition, these classes are very small classes and often concentrated in specific areas. A confusion of these classes with other crop types, although being theoretically well separable as can be seen in high model accuracies, strongly reduces the map accuracy.

4.2. Comparison of Classification Accuracies and Method Complexity to Reference Studies

Apart from the availability of reliable and comprehensive training and validation data, the most challenging task in remote sensing-based crop type classification is the distinction of phenologically similar crops (see e.g., Figure S1). In the last decades, characterizing plant development of different crop types at field scale was often hampered by insufficient revisiting times of high spatial resolution sensors, aggravated by missing observations due to cloud coverage during critical growth stages. However, even though with Sentinel-1 and Sentinel-2 the revisiting time has improved also for high spatial resolution optical and radar time series compared to previous sensors, similar phenologies of crop types are still a challenge for classification tasks. For instance [22,27] reported classification errors for wheat, spelt, and rye due to similar stages of development during the vegetation period for both crops.

For Germany-wide crop type mapping, different approaches have been suggested in the literature to further increase the information gain from remote sensing time series by testing and exploiting different preprocessing and compositing techniques (e.g., [22,26,56]). Overall, these studies achieved high OAs of 88% [22], 81% [26], and 79% [56], respectively, but these involved either very elaborated and computationally expensive preprocessing and interpolation [22,56] or complex input datasets by additionally including non-EO data [26,56], or mapped much fewer classes [26]. We, hence, wanted to assess the effects of a simpler and less data and processing intensive approach on the discrimination of crop types through choosing coarser temporal intervals of input features.

As mentioned above, considering a minimum length of 1 month for the temporal intervals of Sentinel-2 features reduced the occurrence of gaps in features even in periods of persistent cloud cover, enabling the use of simple linearly interpolation, and reduced the overall number of input data. Our hypothesis is that we, thereby, generate a feature data set that is easy to handle while retaining phenological information crucial for crop type discrimination. For example, [26] showed that using gap-filled 10-day instead of monthly features improved their OA only by 2%. It is, nevertheless, obvious that during one month, agricultural vegetation can undergo significant plant development and, thus, narrower temporal intervals potentially contain more detailed information on vegetation dynamics. Our approach to consider information on value distributions within monthly intervals by using mean, minimum, and maximum, aims to compensate for the longer interval. Furthermore, as opposed to [56], we included the Sentinel-2 red-edge bands in our approach, as we assumed that detailed spectral information can help distinguishing rather subtle differences among morphologically similar crops. In fact, the red edge features ranked among the most important optical features (see Figure 3).

We compared our results with other published studies using Sentinel data for crop type mapping in Germany, namely, the studies Preidl et al. [22] (PRE) and Blickensdörfer et al. [56] (BLI). It must be explicitly noted that such direct comparison is limited by a number of factors (see Table 4 for an overview on the differences in classification frameworks). Namely, the classifications partly cover other years (PRE), are mapped at lower spatial resolution (PRE), and the number and definition of crop classes deviate in many cases. Most prominently, both reference studies included permanent grasslands in their class

legend, which was not the case in the study presented here. To get the best comparability, the class-wise model accuracies of all three studies were investigated (taken from Table S1 of this study; from Figure S6 and Table S8 in BLI; and from Table 3 in PRE). Nevertheless, the comparison is still limited in this regard by the fact that in the presented study, as well as in BLI, an equal class sampling scheme was applied, while PRE used varying numbers of training and validation samples per class and highlight in their discussion the influence of the class imbalance. Finally, the legend used for class accuracy comparison had to be harmonized as far as information from the studies allowed for it. This means, the comparison was restricted to identical or very similar classes of the three studies (see Figure 9).

Table 4. Overview on crop type classification settings of the investigated studies Preidl et al. [22] (PRE), Blickensdörfer et al. [56] (BLI), and this study.

	PRE	BLI	This Study
year(s)	2016	2017–2019	2018
spatial resolution [m]	20	10	10
reference data	LPIS data from 7 Federal States + local patches	LPIS data from 4–5 Federal States	LPIS data from 15 Federal States
sampling scheme	proportional	equal	equal
number of crop classes	19	23	17
permanent grassland included	yes	yes	no
number of input features	54–126 (in 64–16,383 model runs)	483	336
input source types	optical	optical + radar + topography + climate + meteorology	optical + radar
regionalization	yes	no	no

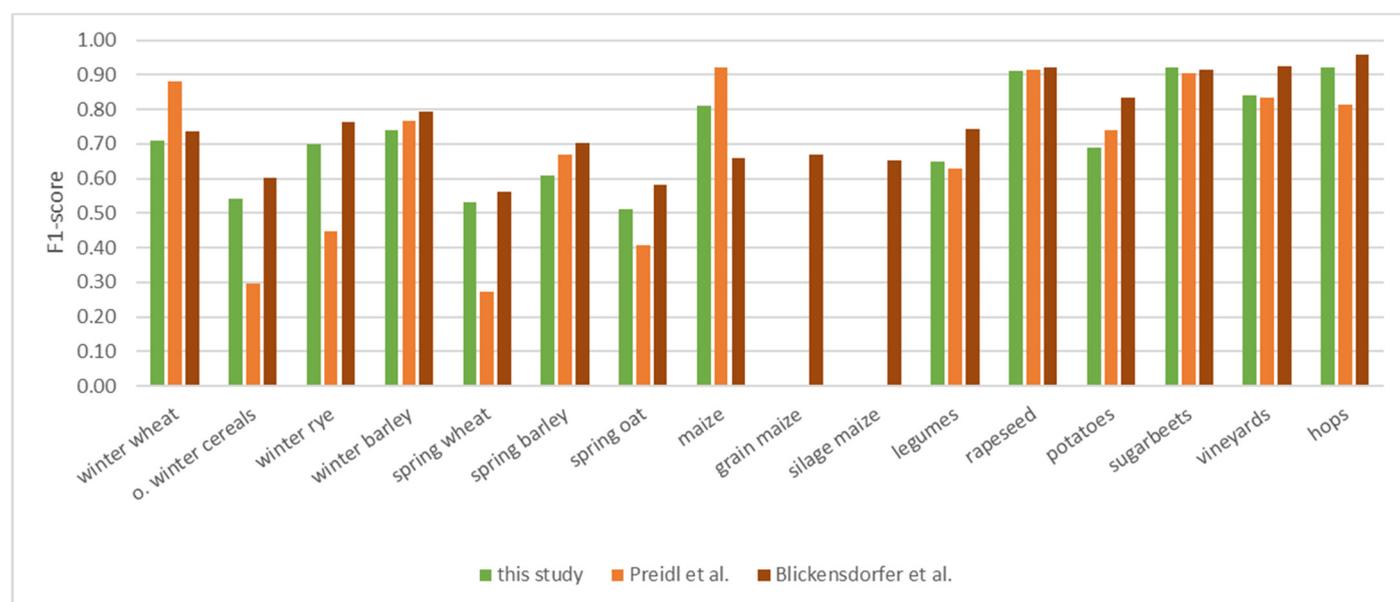


Figure 9. Comparison of class-wise model accuracies (F1-scores) achieved in the presented approach with the accuracy metrics from Preidl et al. [22] (PRE) and Blickensdörfer et al. [56] (BLI). The comparison is not perfect due to temporal and spatial mismatch, as well as partly different class definitions, different sampling schemes, and number of samples (see Table 4). It must be noted that the class “other winter cereals” comprises winter triticale and winter spelt in BLI and in this study, but only winter spelt in PRE, and that the class “spring wheat” also comprises spring triticale and spring rye in the case of BLI.

Figure 9 displays the class-wise model accuracies for the three selected Germany-wide crop type classifications. Overall, similar patterns are discernible. The most prominent classes (winter wheat, maize, rapeseed, and sugar beet) achieve high (≥ 0.8) F1-scores in almost all studies. The very high F1-scores reached by PRE in the two largest coverage classes winter wheat and maize could probably be attributed to the class-wise sample imbalance introduced through the proportional sampling procedure which favors the correct classification of large classes. BLI deviate from that pattern for the maize class, as they attempt to map two individual maize classes, which could each be mapped with accuracies of approximately 65%. Further, the permanent crop classes included in this comparison, i.e., the vineyards and hops classes, could be classified with high model accuracies in all three studies. Winter barley, spring barley, legumes, and potatoes are also mapped in all three studies with similar F1-scores ($\pm 5\%$), but the accuracy is mostly only between 60–80% and BLI achieve the highest scores for these classes. For the smaller crop classes, i.e., winter rye, other winter cereals, spring wheat, and spring oat, PRE achieves the lowest class accuracies, with decreases in accuracy compared to the other two studies by 10–25%. Based on this result, it seems that involving shorter or adapted temporal intervals does not generally benefit the differentiation of subtle differences between cereal classes.

From this class-wise accuracy metrics investigation, it can be concluded that the presented approach, based on a reduced temporal resolution of one month and employing relatively simple and efficient pre-processing steps, is a good trade-off between processing complexity and classification quality. For 10 classes, the presented approach achieved second best results (1–11% below the respective study with highest accuracies), in the case of sugar beet, even the best classification result could be reached. Lowest accuracies were obtained only for 4 out of 14 classes, with accuracies reduced by 3–5% in comparison to the next best classification. Especially, the most widely distributed crops, as well as the permanent crop classes, could be mapped with good F1-scores that compare well to the reference studies.

4.3. Combination of Optical and Radar Features

To assess the influence and importance of the Sentinel-1 as well as Sentinel-2 data on the map accuracy, three experiments have been conducted using the two sensors combined and each individually. Highest OA (75.5%) is reached using the S12 model, while the S1 and S2 models remain below 70% accuracy (66.6% and 69.7%, respectively). The same pattern, i.e., achieving higher OA using only optical data than using only radar, was found by [56]. The authors of [57], on the other hand, reached higher accuracies using SAR data alone than using solely optical data. Nevertheless, both studies found that the optical-SAR combination outperformed single sensor crop type predictions. This conclusion was also drawn by classifications conducted outside of Germany [48–55].

In Sections 3.2 and 3.3 we investigated this generally recognized finding in more detail and individually for each crop type class. The identified patterns of feature importance relate well to the principles of remote sensing-based vegetation mapping. In relative terms, the Sentinel-2 features importance outweighs Sentinel-1 feature importance by a factor of approximately three (73% Sentinel-2 vs. 27% Sentinel-1). This higher importance of optical features is plausible, as optical bands are known to be sensitive to photosynthetic activity (i.e., the red, red edge, and NIR bands as well as thereof derived vegetation indices), and, hence, different species phenologies and growth forms are generally well reflected in multispectral data. Furthermore, temporal features from the critical phases of vegetation growth, namely, the spring months April and May, as well as the ripening and harvesting period in July and August, showed high importance scores. Comparing the feature importance to other studies is limited, as, usually, the same individual features are not used, and investigating feature importance over groups of features might be influenced by the absolute number of features. The authors of [56], for example, also compared groups of features; however, in their study the difference between the numbers of optical and radar features is even larger than in our study. Nevertheless, also in the setting of [56], a

higher importance of optical features was observed compared to the feature importance of radar features, and looking at individual bands/indices. In addition, in their study, the NDVI shows highest importance scores in comparison to the individual bands in the visible light (red, green, blue), which compares well to the results of this study. Furthermore, the relevance of temporal intervals of features generally match our findings.

Through assessing the differences in the sample count-based confusion matrices of the different experiment results (Figure 6), it could be shown that the combination of Sentinel-1 and -2 mainly helped in reducing the omission and commission errors within the groups of different winter and spring cereal crops, respectively. Moreover, the distinction of spring cereals from other phenologically similar crops, such as legumes, was improved by the inclusion of radar data. This might be caused by the information on varying canopy structure that is added through the radar signal. While legumes and spring cereals show similar NDVI curves throughout the growing period, the backscatter patterns deviate, with legumes undergoing a much higher increase in backscatter signal during June (not shown). Another example is the hops class, for which the confusion with almost all other classes could be reduced through the addition of Sentinel-1, but mostly with the maize class and with the spectrally very similar grassland classes, as hop fields themselves have high shares of herbaceous cover between the rows. Moreover, in this case, the radar signal influenced by the distinct vertical structure of hop gardens transports higher backscatter values than the herbaceous fodder crops. The optical signal, on the other hand, strongly improved the distinction of spring oat from all winter crops, as well as the distinction of all spring crops from classes with herbaceous cover (clover/alfalfa, arable grass, fruit trees). Unexpectedly, rapeseed was better detected using Sentinel-1 features only compared to using only optical data (Figures 4 and 6), even though a high sensitivity of optical measurements to the distinctive yellow flowering signal would be expected. We, therefore, expect promising results through including a yellowness index feature in future map versions.

4.4. Regional Stratification

Stratifying Germany in landscape regions aimed at generating climatologically, geomorphologically, and structurally more homogeneous classification units, in order to reduce the variance in growth conditions and management, and, therewith, to better pinpoint regional differences in class-specific vegetation development. However, the OA values for the different landscape regions varied, with highest OA of 80.5% achieved in the Alpine Foreland and lowest OA of 72.1% reported for the Eastern and Southwestern Uplands. Moreover, OA averaged over all regions was with 74.7% slightly lower than for the Germany-wide classification. These results do not suggest a general improvement of crop type classification in Germany through stratification using the landscape regions. Furthermore, partly large variances within one class for different regions could be observed, which reduces the reliability of such an approach when used unconditionally. For example, in the of class fruit trees, the average OA could be improved by almost 10% through stratification, but at the same time, the F1-scores of two regions dropped to 0.15 and 0.16, respectively.

However, when looking at the class-wise map accuracies, it is also striking that only the accuracy metrics of the cereal classes decreased, while all other classes, i.e., legumes, potato, sugar beet, rapeseed, clover/alfalfa, arable grass, vineyard, fruit trees, and hops, on average, achieve higher accuracies when mapped per region. Hence, the class-wise accuracies for all but the cereal classes could be overall improved on average by 7%. Especially the small and permanent classes vineyard (increase of 16%) and hops (increase of 17%) demonstrate that a regionally adapted approach is able to improve the mapping of such specific classes. It must, however, be assumed that the accuracies of these classes profit not only from a better model fit to region-specific growth forms, but also from the constraint, that they are not attempted to be mapped over areas in which they do not or rarely occur, hence automatically reducing commission errors. On the other hand, small

cultivation areas, e.g., the vineyards in Saxony-Anhalt, are inevitably lost due to the chosen stratification settings.

So far, only a few studies have employed stratification for large area crop type mapping. While, for example, the authors of [60] demonstrate that the eco-climatic stratification applied in France yields better overall results by about 4%, other studies (e.g., [7]) employ stratification without comparing their classification results against non-stratified results. Similarly, the authors of [22] have used the same specific regionalization scheme for Germany as used in this study, but did neither discuss the benefit of stratification for OA, nor for class-specific accuracies. An evaluation of the implemented approach with regard to comparable studies remains hence limited.

Overall, the stratification experiment illustrates that the user of such an approach would need to evaluate the specific intended purpose of the generated crop type map, and therewith user's priorities, e.g., if a high OA, high class-specific accuracies, or minimum overall or class accuracies needs to be reached in each or in individual regions. Nevertheless, the potential of the method is shown, which could potentially further be improved through adapted strata or regional classification models, or through legend settings.

5. Conclusions

A crop type classification workflow for Germany based on Sentinel-1, Sentinel-2, and LPIS data has been presented for the year 2018. In this approach, we made use of LPIS data from 15 out of 16 German Federal States, making the training robust and the results representative for the entire country. The RF algorithm making use of Sentinel-1 and Sentinel-2 together performed best and was able to classify crops such as winter wheat, maize, winter rapeseed, sugar beet with high accuracies (F1-score ≥ 0.8), while achieving an OA of 75.5%. Omission and commission errors occurred mostly among winter cereals, summer cereals, and grassland classes. In comparison to official agricultural census data, the crop class areas could be approximated well with on average only 1% of deviation in class-specific acreages.

Differences between the class-specific F1-scores and OA across different landscape regions in Germany revealed that stratification improved the classification for all but the cereal classes, but that average OA could not profit from that approach. In comparison to previous, more complex classification approaches, the class-wise model accuracies achieved with this approach are generally in the same range, but, for most classes, remain 1–11% below the best results of other studies. This indicates, on the one hand, the suitability of the suggested streamlined procedure for a national-scale crop type mapping when efficiency, robustness, and reproducibility of the classification are a priority, while maintaining high map accuracies. On the other hand, these results show that more complex procedures might be needed if the map producer strives for highest possible OA or class-specific accuracies, which come at the cost of increased processing complexity and time.

The generated crop type map can be used for different applications, ranging from the analysis of yield estimates, land use intensity, or agrobiodiversity, to adaptation strategies for agriculture to climate change induced challenges, and, hence, is certainly of high interest for a range of users. For future studies, the potential of additional indices for improving class-specific accuracies would be of interest. With regard to the envisaged class legend, it should be further evaluated at which level of detail similar crop types (e.g., different cultivars of maize or legumes) can still be reliably distinguished. Last, but not least, tests regarding the transferability of the classification to other years are of very high relevance for future research. This is of special interest for situations in which official LPIS data are not (yet) published by the Federal States of interest, but crop type area estimates are needed in a timely manner. A robust classification procedure, transferable in time, would be the solution to this problem, for which the presented approach could provide the basic prerequisite.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs14132981/s1>, Figure S1: Crop calendars of 2018 for major crops in Germany based on DWD (Deutscher Wetterdienst) phenology observation data (DWD-CDC (2001); own outlier analysis and visualization); Figure S2: Comparison of the crop type classifications for Germany 2018 based on a) the S1 model, b) the S2 model and c) the S12 model for a subset in the Magdeburger Börde region in Saxony-Anhalt (upper row), and the “Eichsfeld” agricultural region in Thuringia (bottom row); Table S1: Model and map accuracy metrics (precision, recall, and F1-score) for crop type classifications for the Germany-wide RF models using different input feature sets, namely all Sentinel-1 and -2 features (S12), only the Sentinel-1 features (S1) and only the Sentinel-2 features (S2).

Author Contributions: Conceptualization, S.A., U.G. and C.K.; Data curation, S.A. and U.G.; Formal analysis, R.A.G.; Funding acquisition, S.A., U.G. and C.K.; Methodology, S.A., U.G. and R.A.G.; Project administration, S.A. and U.G.; Software, R.A.G., M.W. and J.K.; Validation, S.A., U.G. and R.A.G.; Visualization, S.A., U.G., R.A.G. and M.W.; Writing—original draft, S.A., U.G. and R.A.G.; Writing—review and editing, S.A., U.G., R.A.G., M.W., J.K. and C.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Helmholtz Association of German Research Centers under the Helmholtz Climate Initiative (“HI-CAM: Helmholtz-Initiative Climate Adaptation and Mitigation”) project and by the Federal Ministry of Food and Agriculture through the “Einsatz von Fernerkundungstechnologien für die Digitalisierung im Pflanzenbau (AgriSens DEMMIN 4.0)” project (Förderkennzeichen: 28DE114C18).

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to data policy restrictions of the Federal States offices providing the LPIS data.

Acknowledgments: The LPIS data were provided for usage within the HI-CAM project by the Ministerium für Ländlichen Raum und Verbraucherschutz Baden-Württemberg, the Bayerischem Staatsministerium für Ernährung, Landwirtschaft und Forsten, the Landesvermessung und Geobasisinformation Brandenburg, the Wirtschafts- und Infrastrukturbank Hessen, the Ministerium für Landwirtschaft und Umwelt Mecklenburg-Vorpommern, the Niedersächsisches Ministerium für Ernährung, Landwirtschaft und Verbraucherschutz, the Ministerium für Wirtschaft, Verkehr, Landwirtschaft und Weinbau Rheinland-Pfalz, the Landesamt für Vermessung, Geoinformation und Landentwicklung Saarland, the Sächsisches Staatsministerium für Energie, Klimaschutz, Umwelt und Landwirtschaft, the Ministerium für Umwelt, Landwirtschaft und Energie Sachsen-Anhalt, the Ministerium für Energiewende, Landwirtschaft, Umwelt, Natur und Digitalisierung des Landes Schleswig-Holstein, and the Thüringer Landesamt für Landwirtschaft und Ländlichen Raum.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Statistisches Bundesamt (Destatis). Genesis-Online. Available online: <https://www-genesis.destatis.de/genesis/online> (accessed on 21 March 2022).
2. Ray, D.K.; West, P.C.; Clark, M.; Gerber, J.S.; Prishchepov, A.V.; Chatterjee, S. Climate change has likely already affected global food production. *PLoS ONE* **2019**, *14*, e0217148. [[CrossRef](#)] [[PubMed](#)]
3. Le Gouis, J.; Oury, F.-X.; Charmet, G. How changes in climate and agricultural practices influenced wheat production in Western Europe. *J. Cereal Sci.* **2020**, *93*, 102960. [[CrossRef](#)]
4. Eckstein, D.; Künzel, V.; Schäfer, L.; Wings, M. *Global Climate Risk Index 2020*; Germanwatch, e.V.: Bonn, Germany, 2019; p. 44.
5. Vitasse, Y.; Rebetez, M. Unprecedented risk of spring frost damage in Switzerland and Germany in 2017. *Clim. Chang.* **2018**, *149*, 233–246. [[CrossRef](#)]
6. Lippert, C.; Krimly, T.; Aurbacher, J. A Ricardian analysis of the impact of climate change on agriculture in Germany. *Clim. Chang.* **2009**, *97*, 593. [[CrossRef](#)]
7. Defourny, P.; Bontemps, S.; Bellemans, N.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Nicola, L.; Rabaute, T.; et al. 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 Sens. Environ.* **2019**, *221*, 551–568. [[CrossRef](#)]
8. Ibrahim, E.S.; Rufin, P.; Nill, L.; Kamali, B.; Nendel, C.; Hostert, P. Mapping Crop Types and Cropping Systems in Nigeria with Sentinel-2 Imagery. *Remote Sens.* **2021**, *13*, 3523. [[CrossRef](#)]

9. Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W.-T. How much does multi-temporal Sentinel-2 data improve crop type classification? *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 122–130. [[CrossRef](#)]
10. Immitzer, M.; Vuolo, F.; Atzberger, C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sens.* **2016**, *8*, 166. [[CrossRef](#)]
11. Hernandez, I.; Benevides, P.; Costa, H.; Caetano, M. Exploring Sentinel-2 for Land Cover and Crop Mapping in Portugal. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 83–89. [[CrossRef](#)]
12. Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [[CrossRef](#)]
13. Mazzia, V.; Khaliq, A.; Chiaberge, M. Improvement in Land Cover and Crop Classification based on Temporal Features Learning from Sentinel-2 Data Using Recurrent-Convolutional Neural Network (R-CNN). *Appl. Sci.* **2019**, *10*, 238. [[CrossRef](#)]
14. Paris, C.; Weikmann, G.; Bruzzone, L. Monitoring of Agricultural Areas by using Sentinel 2 Image Time Series and Deep Learning Techniques. In Proceedings of the SPIE 11533, Image and Signal Processing for Remote Sensing XXVI, Online, 21 September 2020.
15. Račić, M.; Oštir, K.; Peressutti, D.; Zupanc, A.; Čehovin Zajc, L. Application of Temporal Convolutional Neural Network for the Classification of Crops on Sentinel-2 Time Series. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 1337–1342. [[CrossRef](#)]
16. Rousi, M.; Sitokostantinou, V.; Meditskos, G.; Papoutsis, I.; Gialampoukidis, I.; Koukos, A.; Karathanassi, V.; Drivas, T.; Vrochidis, S.; Kontoes, C.; et al. Semantically Enriched Crop Type Classification and Linked Earth Observation Data to Support the Common Agricultural Policy Monitoring. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 529–552. [[CrossRef](#)]
17. Sitokostantinou, V.; Papoutsis, I.; Kontoes, C.; Arnal, A.; Andrés, A.P.; Zurbarano, J.A. Scalable Parcel-Based Crop Identification Scheme Using Sentinel-2 Data Time-Series for the Monitoring of the Common Agricultural Policy. *Remote Sens.* **2018**, *10*, 911. [[CrossRef](#)]
18. Turkoglu, M.O.; D’Aronco, S.; Perich, G.; Liebisch, F.; Streit, C.; Schindler, K.; Wegner, J.D. Crop mapping from image time series: Deep learning with multi-scale label hierarchies. *Remote Sens. Environ.* **2021**, *264*. [[CrossRef](#)]
19. Belgiu, M.; Bijker, W.; Csillik, O.; Stein, A. Phenology-based sample generation for supervised crop type classification. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *95*, 102264. [[CrossRef](#)]
20. Martini, M.; Mazzia, V.; Khaliq, A.; Chiaberge, M. Domain-Adversarial Training of Self-Attention-Based Networks for Land Cover Classification Using Multi-Temporal Sentinel-2 Satellite Imagery. *Remote Sens.* **2021**, *13*, 2564. [[CrossRef](#)]
21. Gallo, I.; La Grassa, R.; Landro, N.; Boschetti, M. Sentinel 2 Time Series Analysis with 3D Feature Pyramid Network and Time Domain Class Activation Intervals for Crop Mapping. *ISPRS Int. J. Geoinf.* **2021**, *10*, 483. [[CrossRef](#)]
22. Preidl, S.; Lange, M.; Doktor, D. Introducing APiC for regionalised land cover mapping on the national scale using Sentinel-2A imagery. *Remote Sens. Environ.* **2020**, *240*, 111673. [[CrossRef](#)]
23. Rußwurm, M.; Körner, M. Self-attention for raw optical Satellite Time Series Classification. *ISPRS J. Photogramm. Remote Sens.* **2020**, *169*, 421–435. [[CrossRef](#)]
24. Marszalek, M.; Lösch, M.; Körner, M.; Schmidhalter, U. Multi-temporal Crop Type and Field Boundary Classification with Google Earth Engine. *Preprints* **2020**, *2020*, 040316. [[CrossRef](#)]
25. Piedadlobo, L.; Hernández-López, D.; Ballesteros, R.; Chakhar, A.; Del Pozo, S.; González-Aguilera, D.; Moreno, M.A. Scalable pixel-based crop classification combining Sentinel-2 and Landsat-8 data time series: Case study of the Duero river basin. *Agric. Syst.* **2019**, *171*, 36–50. [[CrossRef](#)]
26. Griffiths, P.; Nendel, C.; Hostert, P. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sens. Environ.* **2019**, *220*, 135–151. [[CrossRef](#)]
27. Heupel, K.; Spengler, D.; Itzerott, S. A Progressive Crop-Type Classification Using Multitemporal Remote Sensing Data and Phenological Information. *PFG–J. Photogramm. Remote Sens. Geoinf. Sci.* **2018**, *86*, 53–69. [[CrossRef](#)]
28. Kyere, I.; Astor, T.; Graß, R.; Wachendorf, M. Agricultural crop discrimination in a heterogeneous low-mountain range region based on multi-temporal and multi-sensor satellite data. *Comput. Electron. Agric.* **2020**, *179*, 105864. [[CrossRef](#)]
29. Orynbaikyzy, A.; Gessner, U.; Conrad, C. Crop type classification using a combination of optical and radar remote sensing data: A review. *Int. J. Remote Sens.* **2019**, *40*, 6553–6595. [[CrossRef](#)]
30. Arias, M.; Campo-Bescós, M.Á.; Álvarez-Mozos, J. Crop Classification Based on Temporal Signatures of Sentinel-1 Observations over Navarre Province, Spain. *Remote Sens.* **2020**, *12*, 278. [[CrossRef](#)]
31. d’Andrimont, R.; Verhegghen, A.; Lemoine, G.; Kempeneers, P.; Meroni, M.; 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.* **2021**, *266*, 112708. [[CrossRef](#)]
32. Beriaux, E.; Jago, A.; Lucau-Danila, C.; Planchon, V.; Defourny, P. Sentinel-1 Time Series for Crop Identification in the Framework of the Future CAP Monitoring. *Remote Sens.* **2021**, *13*, 2785. [[CrossRef](#)]
33. Mestre-Quereda, A.; Lopez-Sanchez, J.M.; Vicente-Guijalba, F.; Jacob, A.W.; Engdahl, M.E. Time-Series of Sentinel-1 Interferometric Coherence and Backscatter for Crop-Type Mapping. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4070–4084. [[CrossRef](#)]
34. Nasirzadehdizaji, R.; Sanli, F.B.; Cakir, Z. Application of Sentinel-1 Multi-Temporal Data for Crop Monitoring and Mapping. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 803–807. [[CrossRef](#)]

35. Planque, C.; Lucas, R.; Punalekar, S.; Chognard, S.; Hurford, C.; Owers, C.; Horton, C.; Guest, P.; King, S.; Williams, S.; et al. National Crop Mapping Using Sentinel-1 Time Series: A Knowledge-Based Descriptive Algorithm. *Remote Sens.* **2021**, *13*, 846. [[CrossRef](#)]
36. Teimouri, N.; Dyrmann, M.; Jørgensen, R.N. A Novel Spatio-Temporal FCN-LSTM Network for Recognizing Various Crop Types Using Multi-Temporal Radar Images. *Remote Sens.* **2019**, *11*, 990. [[CrossRef](#)]
37. Denize, J.; Hubert-Moy, L.; Pottier, E. Polarimetric SAR Time-Series for Identification of Winter Land Use. *Sensors* **2019**, *19*, 5574. [[CrossRef](#)]
38. Gella, G.W.; Bijker, W.; Belgiu, M. Mapping crop types in complex farming areas using SAR imagery with dynamic time warping. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 171–183. [[CrossRef](#)]
39. Valcarce-Diñeiro, R.; Arias-Pérez, B.; Lopez-Sanchez, J.M.; Sánchez, N. Multi-Temporal Dual- and Quad-Polarimetric Synthetic Aperture Radar Data for Crop-Type Mapping. *Remote Sens.* **2019**, *11*, 1518. [[CrossRef](#)]
40. Woźniak, E.; Rybicki, M.; Kofman, W.; Aleksandrowicz, S.; Wojtkowski, C.; Lewiński, S.; Bojanowski, J.; Musiał, J.; Milewski, T.; Slesiński, P.; et al. Multi-temporal phenological indices derived from time series Sentinel-1 images to country-wide crop classification. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *107*, 102683. [[CrossRef](#)]
41. Reuß, F.; Greimeister-Pfeil, I.; Vreugdenhil, M.; Wagner, W. Comparison of Long Short-Term Memory Networks and Random Forest for Sentinel-1 Time Series Based Large Scale Crop Classification. *Remote Sens.* **2021**, *13*, 5000. [[CrossRef](#)]
42. Bargiel, D. A new method for crop classification combining time series of radar images and crop phenology information. *Remote Sens. Environ.* **2017**, *198*, 369–383. [[CrossRef](#)]
43. Hütt, C.; Waldhoff, G.; Bareth, G. Fusion of Sentinel-1 with Official Topographic and Cadastral Geodata for Crop-Type Enriched LULC Mapping Using FOSS and Open Data. *ISPRS Int. J. Geoinf.* **2020**, *9*, 120. [[CrossRef](#)]
44. Kenduiywo, B.K.; Bargiel, D.; Soergel, U. Crop-type mapping from a sequence of Sentinel 1 images. *Int. J. Remote Sens.* **2018**, *39*, 6383–6404. [[CrossRef](#)]
45. Sun, L.; Chen, J.; Guo, S.; Deng, X.; Han, Y. Integration of Time Series Sentinel-1 and Sentinel-2 Imagery for Crop Type Mapping over Oasis Agricultural Areas. *Remote Sens.* **2020**, *12*, 158. [[CrossRef](#)]
46. Kpianbaareh, D.; Sun, X.; Wang, J.; Luginaah, I.; Bezner Kerr, R.; Lupafya, E.; Dakishoni, L. Crop Type and Land Cover Mapping in Northern Malawi Using the Integration of Sentinel-1, Sentinel-2, and PlanetScope Satellite Data. *Remote Sens.* **2021**, *13*, 700. [[CrossRef](#)]
47. Moumni, A.; Lahrouni, A. Machine Learning-Based Classification for Crop-Type Mapping Using the Fusion of High-Resolution Satellite Imagery in a Semiarid Area. *Scientifica* **2021**, *2021*, 8810279. [[CrossRef](#)] [[PubMed](#)]
48. Chakhar, A.; Hernández-López, D.; Ballesteros, R.; Moreno, M.A. Improving the Accuracy of Multiple Algorithms for Crop Classification by Integrating Sentinel-1 Observations with Sentinel-2 Data. *Remote Sens.* **2021**, *13*, 243. [[CrossRef](#)]
49. Inglada, J.; Vincent, A.; Arias, M.; Marais-Sicre, C. Improved Early Crop Type Identification By Joint Use of High Temporal Resolution SAR And Optical Image Time Series. *Remote Sens.* **2016**, *8*, 362. [[CrossRef](#)]
50. Kussul, N.; Lemoine, G.; Gallego, F.J.; Skakun, S.V.; Lavreniuk, M.; Shelestov, A.Y. Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2500–2508. [[CrossRef](#)]
51. Van Tricht, K.; Gobin, A.; Gilliams, S.; Piccard, I. Synergistic Use of Radar Sentinel-1 and Optical Sentinel-2 Imagery for Crop Mapping: A Case Study for Belgium. *Remote Sens.* **2018**, *10*, 1642. [[CrossRef](#)]
52. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geosci. Remote. Sens. Lett.* **2017**, *14*, 778–782. [[CrossRef](#)]
53. Kussul, N.; Mykola, L.; Shelestov, A.; Skakun, S. Crop inventory at regional scale in Ukraine: Developing in season and end of season crop maps with multi-temporal optical and SAR satellite imagery. *Eur. J. Remote. Sens.* **2018**, *51*, 627–636. [[CrossRef](#)]
54. Giordano, S.; Bailly, S.; Landrieu, L.; Chehata, N. Improved Crop Classification with Rotation Knowledge using Sentinel-1 and -2 Time Series. *Photogramm. Eng. Remote Sens.* **2020**, *86*, 431–441. [[CrossRef](#)]
55. Ofori-Ampofo, S.; Pelletier, C.; Lang, S. Crop Type Mapping from Optical and Radar Time Series Using Attention-Based Deep Learning. *Remote Sens.* **2021**, *13*, 4668. [[CrossRef](#)]
56. Blickensdörfer, L.; Schwieder, M.; Pflugmacher, D.; Nendel, C.; Erasmi, S.; Hostert, P. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote Sens. Environ.* **2022**, *269*, 112831. [[CrossRef](#)]
57. Orynbaikyzy, A.; Gessner, U.; Mack, B.; Conrad, C. Crop Type Classification Using Fusion of Sentinel-1 and Sentinel-2 Data: Assessing the Impact of Feature Selection, Optical Data Availability, and Parcel Sizes on the Accuracies. *Remote Sens.* **2020**, *12*, 2779. [[CrossRef](#)]
58. Orynbaikyzy, A.; Gessner, U.; Conrad, C. Spatial Transferability of Random Forest Models for Crop Type Classification Using Sentinel-1 and Sentinel-2. *Remote Sens.* **2022**, *14*, 1493. [[CrossRef](#)]
59. Salehi, B.; Daneshfar, B.; Davidson, A.M. Accurate crop-type classification using multi-temporal optical and multi-polarization SAR data in an object-based image analysis framework. *Int. J. Remote Sens.* **2017**, *38*, 4130–4155. [[CrossRef](#)]
60. Inglada, J.; Vincent, A.; Arias, M.; Tardy, B.; Morin, D.; Rodes, I. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sens.* **2017**, *9*, 95. [[CrossRef](#)]
61. Foerster, S.; Kaden, K.; Foerster, M.; Itzerott, S. Crop type mapping using spectral-temporal profiles and phenological information. *Comput. Electron. Agric.* **2012**, *89*, 30–40. [[CrossRef](#)]

62. Kerner, H.; Sahajpal, R.; Skakun, S.; Becker-Reshef, I.; Barker, B.; Hosseini, M.; Puricelli, E.; Gray, P. Resilient in-season crop type classification in multispectral satellite observations using growth stage normalization. *arXiv* **2020**. [CrossRef]
63. Skakun, S.; Vermote, E.; Franch, B.; Roger, J.-C.; Kussul, N.; Ju, J.; Masek, J. Winter Wheat Yield Assessment from Landsat 8 and Sentinel-2 Data: Incorporating Surface Reflectance, Through Phenological Fitting, into Regression Yield Models. *Remote Sens.* **2019**, *11*, 1768. [CrossRef]
64. DESTATIS. Landwirtschaftliche Betriebe, Ausgewählte Merkmale im Zeitvergleich. Available online: <https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Landwirtschaft-Forstwirtschaft-Fischerei/Landwirtschaftliche-Betriebe/Tabellen/ausgewaehlte-merkmale-zv.html> (accessed on 21 April 2022).
65. BMEL. Daten und Fakten-Land-, Forst- und Ernährungswirtschaft mit Fischerei und Wein- und Gartenbau. Available online: <https://www.bmel.de/SharedDocs/Downloads/DE/Broschueren/Daten-und-Fakten-Landwirtschaft.pdf> (accessed on 21 April 2022).
66. German Environment Agency. Environment and Agriculture 2018. 2018. Available online: https://www.umweltbundesamt.de/sites/default/files/medien/421/publikationen/180608_uba_fl_umwelt_und_landwirtschaft_engl_bf_neu.pdf (accessed on 21 April 2022).
67. ESA. *Sentinel-2 User Handbook*; ESA: Paris, France, 2015; p. 64. Available online: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook (accessed on 21 April 2022).
68. de Los Reyes, R.; Langheinrich, M.; Schwind, P.; Richter, R.; Pflug, B.; Bachmann, M.; Muller, R.; Carmona, E.; Zekoll, V.; Reinartz, P. PACO: Python-Based Atmospheric Correction. *Sensors* **2020**, *20*, 1428. [CrossRef] [PubMed]
69. Veci, L.; Prats-Iraola, P.; Scheiber, R.; Collard, F.; Fomferra, N.; Engdahl, M. The sentinel-1 toolbox. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Québec, QC, Canada, 14–18 July 2012; pp. 1–3.
70. Dobrinić, D.; Gašparović, M.; Medak, D. Sentinel-1 and 2 Time-Series for Vegetation Mapping Using Random Forest Classification: A Case Study of Northern Croatia. *Remote Sens.* **2021**, *13*, 2321. [CrossRef]
71. Ye, Y.; Yang, C.; Zhu, B.; Zhou, L.; He, Y.; Jia, H. Improving Co-Registration for Sentinel-1 SAR and Sentinel-2 Optical Images. *Remote Sens.* **2021**, *13*, 928. [CrossRef]
72. Yan, X.; Zhang, Y.; Zhang, D.; Hou, N. Multimodal image registration using histogram of oriented gradient distance and data-driven grey wolf optimizer. *Neurocomputing* **2020**, *392*, 108–120. [CrossRef]
73. EUROSTAT. Land Use and Coverage Area frame Survey-LUCAS. 2018. Available online: <https://ec.europa.eu/eurostat/de/web/lucas> (accessed on 21 April 2022).
74. European Union. Copernicus Land Monitoring Service 2018. 2018. Available online: <https://land.copernicus.eu/> (accessed on 21 April 2022).
75. mundialis GmbH & Co. KG. Landcover Classification Map of Germany 2019 Based on Sentinel-2 Data. 2020. Available online: <https://www.mundialis.de/en/deutschland-2019-landbedeckung-auf-basis-von-sentinel-2-daten/> (accessed on 21 April 2022).
76. OpenStreetMap-Contributors. OpenStreetMap. 2021. Available online: <https://planet.osm.org>. (accessed on 21 April 2022).
77. Marconcini, M.; Metz-Marconcini, A.; Üreyen, S.; Palacios-Lopez, D.; Hanke, W.; Bachofer, F.; Zeidler, J.; Esch, T.; Gorelick, N.; Kakarla, A.; et al. Outlining where humans live, the World Settlement Footprint 2015. *Sci. Data* **2020**, *7*, 242. [CrossRef]
78. Meynen, E.; Schmithüsen, J.; Gellert, J.; Neef, E.; Müller-Miny, H.; Schultze, H.J. *Handbuch der Naturräumlichen Gliederung Deutschlands*; Bundesanstalt für Landeskunde und Raumforschung: Bad Godesberg, Germany, 1953–1962.
79. DESTATIS. Land- und Forstwirtschaft, Fischerei; Wachstum und Ernte-Feldfrüchte 2018. Available online: https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00096325 (accessed on 21 April 2022).
80. DESTATIS. Land- und Forstwirtschaft, Fischerei; Wachstum und Ernte-Baumobst 2018. Available online: https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00095516 (accessed on 21 April 2022).
81. DESTATIS. Land- und Forstwirtschaft, Fischerei; Wachstum und Ernte-Weinmost 2018. Available online: https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00104653 (accessed on 21 April 2022).
82. Bayerische Landesanstalt für Landwirtschaft. Hopfen des Jahresheftes Agrarmärkte 2020. Available online: https://www.stmelf.bayern.de/mam/cms07/iem/dateien/16_hopfen_by_2020.pdf (accessed on 21 April 2022).
83. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [CrossRef]
84. Steinhausen, M.J.; Wagner, P.D.; Narasimhan, B.; Waske, B. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 595–604. [CrossRef]
85. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
86. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
87. Hüttich, C.; Gessner, U.; Herold, M.; Strohbach, B.; Schmidt, M.; Keil, M.; Dech, S. On the Suitability of MODIS Time Series Metrics to Map Vegetation Types in Dry Savanna Ecosystems: A Case Study in the Kalahari of NE Namibia. *Remote Sens.* **2009**, *1*, 620–643. [CrossRef]
88. Löw, F.; Michel, U.; Dech, S.; Conrad, C. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using Support Vector Machines. *ISPRS J. Photogramm. Remote Sens.* **2013**, *85*, 102–119. [CrossRef]
89. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [CrossRef]

90. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Processing Manag.* **2009**, *45*, 427–437. [[CrossRef](#)]
91. Stehman, S.V. Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *Int. J. Remote Sens.* **2014**, *35*, 4923–4939. [[CrossRef](#)]
92. Kontgis, C.; Schneider, A.; Ozdogan, M. Mapping rice paddy extent and intensification in the Vietnamese Mekong River Delta with dense time stacks of Landsat data. *Remote Sens. Environ.* **2015**, *169*, 255–269. [[CrossRef](#)]