Crop type classification using the combination of Sentinel-1 and Sentinel-2 data

Dissertation
zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät III
Agrar- und Ernährungswissenschaften,
Geowissenschaften und Informatik

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Frau Orynbaikyzy, Aiym

Gutachter:
1. Prof. Dr. Christopher Conrad
2. Prof. Dr. Dirk Tiede

Datum der Verteidigung: 28. Juni 2023

CROP TYPE CLASSIFICATION USING THE COMBINATION OF SENTINEL-1 AND SENTINEL-2 DATA

Main supervisor:
Prof. Dr. Christopher Conrad

Mentor:
Dr. Ursula Gessner

January, 2023

ABSTRACT

Agricultural systems are expected to face enormous challenges worldwide. On the one hand, the demand for food is increasing due to ongoing population growth and change in eating habits. At the same time, global environmental changes such as climate change and biodiversity loss are putting additional pressure on agricultural production.

High-quality and spatially explicit data on the extent and diversity of cultivated crops, i.e. crop type maps, are essential for planning and managing various agricultural activities such as yield prediction, water supply management, crop damage assessment, etc.

Earth Observation data has been one of the primary information sources for crop type mapping for decades. Regularly acquired remote sensing imagery over large geographic areas enables the mapping of crop types and their spatial-temporal extent at local, regional and even global scales.

In the past, optical or Synthetic Aperture Radar (SAR) data were the most commonly used remote sensing datasets for mapping crop types. In recent decades, an increasing number of studies have examined the benefits of using the combination of these two datasets. Research on this topic has intensified with the launch of Sentinel-1 and Sentinel-2 from the European Space Agency's Copernicus programme. Free access to global optical and SAR datasets with the high spatial and temporal resolution has boosted research on the optical-SAR combination.

This thesis researched the potential of the synergetic use of Sentinel-1 SAR data and Sentinel-2 optical data for accurate, large-scale crop type mapping. The systematic review of existing research publications (Chapter II) on the optical-SAR combination for crop type mapping helped uncover research gaps that were addressed further in two research studies.

The first study (Chapter III) focused on large-scale crop type mapping using bi-weekly dense time-series features from both sensors to map 16 plant classes in the German federal state of Brandenburg. In addition, auxiliary variables such as parcel sizes, pixel location within a parcel, and optical data availability were used to explain the classification results.

The second study (Chapter IV) addressed the issue of spatial transferability of Random Forest models based on only optical, only SAR and their combination using two feature selection methods. In the example of seven study sites distributed across Germany, the study investigates the effect of different datasets and feature selection methods on the spatial transferability of Random Forest models.

The research outcomes of the thesis present the advantages and limitations of using one or the combination of both, Sentinel-1 and Sentinel-2 data sources for large-scale, detailed crop type mapping in Germany.

ZUSAMMENFASSUNG

Es wird erwartet, dass landwirtschaftlichen Systeme weltweit vor zunehmenden Herausvorderungen stehen. Einerseits steigt die Nachfrage nach Nahrungsmitteln aufgrund des anhaltenden Bevölkerungswachstums und sich ändernder Ernährungsgewohnheiten. Gleichzeitig wird die Produktion durch globale Umweltveränderungen wie Klimaveränderungen und Biodiversitätsverlust zusätzlich unter druck gesetzt.

Qualitativ hochwertige und räumlich explizite Daten über den Umfang und die Vielfalt angebauter Nutzpflanzen, d. h. Karten von Nutzpflanzenarten, sind für die Planung und das Management einer Vielzahl landwirtschaftlicher Aktivitäten wie Ertragsvorhersage, Wasserversorgungsmanagement, Beurteilung von Ernteschäden usw. von entscheidender Bedeutung.

Seit Jahrzehnten sind Erdbeobachtungsdaten eine der wichtigsten Informationsquellen für die Kartierung von landwirtschaftliche Kulturarten. Regelmäßig erfasste Bilddaten der Fernerkundung über weite geografische Gebiete ermöglichen die Kartierung von Kulturarten und ihrer räumlich-zeitlichen Ausdehnung auf lokaler, regionaler und sogar globaler Ebene.

In der Vergangenheit waren optische oder Synthetic Aperture Radar (SAR)-Daten die am häufigsten verwendeten Fernerkundungsdatensätze für die Kartierung von Pflanzenarten. In den letzten Jahrzehnten haben immer mehr Studien die Vorteile der Verwendung der Kombination dieser beiden Datensätze untersucht. Die Forschung zu diesem Thema wurde mit dem Start von Sentinel-1 und Sentinel-2 aus dem Copernicus-Programm der Europäischen Weltraumorganisation intensiviert. Der freie Zugang zu globalen optischen und SAR-Datensätzen mit hoher räumlicher und zeitlicher Auflösung hat die Forschung zur optischen SAR-Kombination mit einem beispiellosen Detaillierungsgrad vorangetrieben.

In dieser Arbeit wurde das Potenzial der synergetischen Nutzung von Sentinel-1 SAR-Daten und Sentinel-2 optischen Daten für eine genaue, großflächige Kartierung von Pflanzenarten untersucht. Ein systematisches Review der existierenden Forschung (Kapitel II) zur optischen SAR-Kombination für die Kartierung von Kulturarten trug dazu bei, aktuelle Forschungslücken aufzudecken, die in zwei Forschungsstudien im Detail addressiert wurden.

Die erste Studie (Kapitel III) konzentrierte sich auf die großflächige Kartierung von Kulturarten unter Verwendung von zweiwöchentlichen dichten Zeitreihenmerkmalen von beiden Sensoren zur Kartierung von 16 Pflanzenklassen im deutschen Bundesland Brandenburg. Zusätzlich wurden Hilfsvariablen wie Parzellengröße, Lage des Pixels innerhalb einer Parzelle und optische

Datenverfügbarkeit verwendet, um die Klassifikationsergebnisse zu erklären.

Die zweite Studie (Kapitel IV) befasste sich mit der Frage der räumlichen Übertragbarkeit von Random-Forest-Modellen basierend auf nur optischen, nur SAR und der Kombination von optischen und SAT Daten unter Verwendung von zwei Merkmalsauswahlverfahren. Am Beispiel von sieben über Deutschland verteilten Standorten untersucht die Studie den Einfluss unterschiedlicher Datensatz- und Merkmalsauswahlverfahren auf die räumliche Übertragbarkeit von Random-Forest-Modellen.

Die Forschungsergebnisse der Dissertation zeigen die Potentiale und Grenzen der Nutzung von Sentinel-1 bzw. Sentinel-2 Daten und deren Kombination für flächige, detaillierte Kulturartenkartierung in Deutschland auf.

CONTENT

x

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ARD	Analysis Ready Data
BAP	Best-Available-Pixel
BB	Brandenburg
BGR	The Federal Institute for Geosciences and Natural Resources (Bundesanstalt für Geowissenschaften und Rohstoffe)
BT	Brovey Transform
BW	Baden-Wüttenberg
BY	Bavaria
CAP	Common Agricultural Policy
CNN	Convolutional Neural Network
CV	Cross-Validation
DIAS	Data and Information Access Services
DWD	German Weather Service
DWT	Discrete Wavelet Transform
EF	Ehlers Fusion
EO	Earth Observation
ESA	European Space Agency
EU	European Union
EVI	Enhanced Vegetation Index
F1	F1-score
FAO	Food and Agriculture Organization
FD	Freeman-Durden Decomposition
FFS	Forward Feature Selection
FN	False Negative
FP	False Positive

GEE	Google Earth Engine
GIS	Geographic Information System
GRD	Ground Range Detected
HE	Hessen
HLS	Harmonized Landsat-8 Sentinel-2
HPF	High Pass Filter
IACS	Integrated Administration and Control System
IGARSS	International Geoscience and Remote Sensing Symposium
IHS	Intensity, Hue and Saturation
ISO	International Organization for Standardization
IWS	Interferometric Wide Swath
KELM	Kernel-based Extreme Learning Machine
L2A	Level-2A
LAI	Leaf Area Index
LPIS	Land Parcel Identification System
LUCAS	Land Use/Cover Area Frame Survey
LULC	Land Use Land Cover
MAJA	MACCS-ATCOR Joint Algorithm
MIR	Mid-Infrared
ML	Machine Learning
MLC	Maximum Likelihood Classifier
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	Multi-Spectral Instrument
MSMD	Maximum Separability and Minimum Dependency
MV	Mecklenburg-Western Pomerania
NASA	National Aeronautics and Space Administration

NDFI	Normalized Difference Flood Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NDYI	Normalized Difference Yellow Index
NIR	Near-Infrared
NN	Neural Network
OLI	Operational Land Imager
PA	Producer's Accuracy
PCA	Principal Component Analysis
PMI	Paddy Rice Mapping Index
PSRI	Plant Senescence Reflectance Index
RADAR	Radio Detection And Ranging
RF	Random Forest
RGRI	Red-Green Ratio Index
S1	Sentinel-1
S1TBX	Sentinel-1 Toolbox
S2	Sentinel-2
SAR	Synthetic Aperture Radar
SAVI	Soil-Adjusted Vegetation Index
SCL	Scene Classification Layer
SITS	Satellite Image Time Series
SNAP	Sentinel Application Platform
SR	Simple Ratio
SRTM	Shuttle Radar Topography Mission
SVI	Simple Vegetation Index
SVM	Support Vector Machine
SWIR	Short-wave infrared

TH	Thuringia
TM	Thematic Mapper
TP	True Positive
UA	User's Accuracy
UN	United Nations
URL	Uniform Resource Locator
USGS	United States Geological Survey
ZALF	Leibniz Centre for Agricultural Landscape Research

I

CHAPTER I: INTRODUCTION

1. Motivation

With the human population reaching the eight billion mark, food security and food demand have become top priority issues worldwide (Foley et al., 2011). The actions towards "sustainable intensification" of agriculture (H. Charles J. Godfray et al., 2010), which implies increasing the yields from the same area of land while reducing the negative environmental impact, require detailed spatial-temporal information on croplands at different scales. Such information is vital, for instance, to understand the regional cropping patterns, their impact on the various environmental factor (e.g. groundwater use, agrobiodiversity, soil erosion) and the prediction of expected yields. Due to the dynamic nature of croplands, information on their state has to be regularly updated (Atzberger, 2013). Earth Observation (EO) data, as a significant resource of reliable land surface information from a global to local scale, plays a fundamental role in building sound information products for decision-makers (Weiss et al., 2020).

The generation of crop type maps, which are one of the Essential Agricultural Variables (EAV) for the Group on Earth Observations Global Agricultural Monitoring Initiative (GEOGLAM) (Whitcraft et al., 2019), using EO data, has been a subject of numerous studies (Atzberger, 2013). Research studies have shown the successful applicability of optical (e.g., Conrad et al., 2017), radar (e.g., Bargiel, 2017), hyperspectral (e.g., Aneece and Thenkabail, 2021) and lidar (e.g., Prins and Van Niekerk, 2020) remote sensing data for crop type mapping. Optical data provides information such as plant pigment and water content within tissues and is the most widely used remote sensing dataset for crop type classification. The research on optical remote sensing data was particularly boosted with the opening of the Landsat archive in 2008 (Wulder et al., 2012). However, the quality of the optical data can be substantially reduced due to atmospheric effects, clouds, and cloud shadows. Such issues could be eliminated with weather-independent radar data. Early studies have shown that the quality of radar-based maps is lower than those based on optical data (Orynbaikyzy et al., 2019). However, the earlier studies were mainly based on mono-temporal scenes.

The launch of optical - Sentinel-2 A/B and Synthetic Aperture Radar (SAR) – Sentinel-1 A/B satellites as a part of the European Space Agency's (ESA) Copernicus Program gave new possibilities for studies on crop type mapping by providing high spatial-temporal resolution data. The free availability of both datasets and their global coverage triggered a strong interest in exploring the potential of their combined use in various application domains (Joshi et al., 2016), including agriculture (Dusseux et al., 2014). New three red-edge bands from Sentinel-2 were expected to improve further the monitoring and mapping of cropping practices (Weiss et al., 2020). The research investigating the significance of the new remote sensing datasets, particularly their combination, was needed for the crop type mapping domain.

2. Scientific Background

This chapter outlines the scientific background of the thesis. It presents state-of-the-art research in crop type mapping using remote sensing data, the application of machine learning techniques and the issues of spatial transferability.

2.1. Crop Type Classification using Remote Sensing Data

Within the remote sensing domain, an action of assigning the crop type label to a pixel or object based on the associated descriptive features is called crop type classification. Since the output of crop type classification is commonly a map with spatial and temporal context, the term crop type mapping and crop type classification are interchangeably used in remote sensing research.

Before the rise of remote sensing technologies, crop type mapping was done manually by inspecting the fields or collecting census data at the farmland level. With technological advancement and subsequent increased use of remote sensing techniques in agricultural applications, such time-consuming and costly data collection methods are expected to be replaced by (semi-)automatic classification of agricultural fields using remote sensing images. For instance, to reduce the subsidy control costs in the context of the European Union's Common Agricultural Policy (CAP), much research is being done towards incorporating remote sensing-based solutions for large-scale crop type labelling (López-Andreu et al., 2022; Schmedtmann and Campagnolo, 2015). In this respect, remote sensing data allows for gathering and analysing data over large geographic spaces without physically visiting each field or farmland.

The main objective of crop type classification using remote sensing data is to correctly label the crop class of a given pixel or object using a set of input variables. Input variables, depending on the remote

sensing sensor type, contain information on the physical or biophysical properties of the target on the ground. For example, the absorbed and reflected portions of the light at visible and infrared spectral regions depend on the plants' pigmentation. Likewise, plants' canopy structure affects the amount of backscattered signal to the active radar sensor.

Up to now, most of the studies on crop type mapping was based on optical remote sensing data (Belgiu et al., 2021; Conrad et al., 2010; Defourny et al., 2019). By collecting data on reflected solar radiation from the target on the ground at visible, near-infrared and short-wave infrared portions of the electromagnetic spectrum, optical data capture vital information on vegetation health, leaf pigments, water content and other biophysical parameters. Besides, dense time series of optical data may reproduce the phenological development stages of crops (Veloso et al., 2017), which helps to differentiate crop classes with varying phenological patterns.

Nonetheless, the main limitations of optical data are the gaps due to cloud cover, cloud shadows and other atmospheric effects, which hide or degrade the information of interest. This is particularly an issue for large-area mapping, where the derivation of high-quality large-area features is challenging due to strongly varying data availability (Gómez et al., 2016).

For areas with persistent cloud cover, using Synthetic Aperture Radar (SAR) data for crop type mapping is particularly attractive. The longer wavelength (1cm – 1m) at which SAR sensors operate allows penetration through the clouds that make the system weather-independent. The SAR backscatter delivers information on plants' canopy structure (McNairn and Brisco, 2004) and can capture the changes in plants' phenological phases when its time-series is used (Veloso et al., 2017). SAR-based crop types classification has been successfully implemented and tested in many studies (Bargiel, 2017; Clauss et al., 2018; Kenduiywo et al., 2018; McNairn et al., 2014).

2.2. Combination of Sentinel-1 and Sentinel-2

The complementary nature of optical and SAR datasets allows for the simultaneous utilization of information on plants' structural and bio-physical parameters. The first investigative studies on the potential of optical-SAR combination for crop type mapping date back to the 80s (Li et al., 1980; Ulaby et al., 1982). Starting from 2000, the number of studies on this topic started to increase due to the launch of several satellites with radar (e.g., ERS, RADARSAT, ENVISAR ASAR) and optical (e.g., Landsat, SPOT, IRS, MODIS, QuickBird) data acquisition sensors, and the advancements in computing capabilities (Orynbaikyzy et al., 2019). Notably, the launch of Sentinel-1 A/B (in 2014/2016) and Sentinel-2 A/B satellites (in 2015/2017) from ESA's Copernicus Program added more pace to research on the combined use of optical

and SAR data. The first global products such as WorldCover (https://esa-worldcover.org) and WorldCereal (https://esa-worldcereal.org), aim to map global land cover and monitoring of global croplands at the field level accordingly, were initiated using the combination of Sentinel-1 and Sentinel-2 data.

Table I-1. Description of Sentinel-2 A/B spectral bands. Source: ESA.

| Spectral Band | Sentinel-2 A | | Sentinel-2 B | | Spatial Resolution |
	Central Wavelength (nm)	Bandwidth (nm)	Central Wavelength (nm)	Bandwidth (nm)	
B01	442.7	21	442.2	21	60
B02	492.4	66	492.1	66	10
B03	559.8	36	559.0	36	10
B04	664.6	31	664.9	31	10
B05	704.1	15	703.8	16	20
B06	740.5	15	739.1	15	20
B07	782.8	20	779.7	20	20
B08	832.8	106	832.9	106	10
B8A	864.7	21	864.0	22	20
B09	945.1	20	943.2	21	60
B10	1373.5	31	1376.9	30	60
B11	1613.7	91	1610.4	94	20
B12	2202.4	175	2185.7	185	20

The MultiSpectral Instrument (MSI) onboard Sentinel-2 satellites sense spectral reflectance at three visible, two near-infrared, four red-edge and two short-wave-infrared regions of the electromagnetic spectrum (Table I-1). New bands at the red-edge spectral region, sensitive to the plants' nutritional status (Filella and Peñuelas, 1994), were reported to be helpful for successfully classifying various crops (Immitzer et al., 2016). After satellite data were available to the public, numerous research studies have shown the added values of new red-edge bands and improved mapping capabilities that were now possible with Sentinel-2 data (Belgiu and Csillik, 2018; Griffiths et al., 2019; Veloso et al., 2017). Notably, with the minimum mapping unit of 1ha, Sentinel-2 datasets allowed analysis at a sub-field scale and facilitated the research on mapping minor crops that often have very small field sizes. With short revisit frequency, the data can capture the temporal growth dynamics of crops.

Active C-band SAR sensors onboard Sentinel-1 A/B satellites record backscattered radar signals regardless of daytime and weather conditions. SAR-based crop type classification has been successfully implemented in numerous studies (e.g., Bargiel, 2017; Clauss et al., 2018; Kenduiywo et al., 2018; McNairn et al., 2014). Based on the frequency and polarization, backscattering information delivers the complex representation not only of crop plants but also soil moisture, surface roughness, canopy structure and topography (McNairn and

Brisco, 2004). Thus, the synergetic use of SAR and optical data is particularly interesting for crop type mapping.

Table I-2. Description of Sentinel-1A/B satellites. Source: ESA.

Band type	C-band
Centre frequency, GHz	5.405
Orbit:	Sun-synchronous
- height, km	- 693
- inclination, degree	- 98.1
Swath width, km	20-400
Image resolution, m	5-40
Polarization	Dual. (VH, VV)
Incidence angle, degree	18.3 to 47
Repeat rate, days	12 (combined – 6)

With freely accessible Sentinel-1 and Sentinel-2 data, combining optical and SAR data became much more straightforward. While using data from each sensor separately for crop type mapping is possible with reasonable accuracy (Preidl et al., 2020), the early research showed that combining these datasets results in much more improved accuracy than single sensor cases (Sonobe et al., 2017). The increased number of studies on the combination of Sentinel-1 and Sentinel-2 also enables a more consistent comparison of the methodologies (e.g., machine learning models, generated features) based on the same input data.

2.3. Mapping Crops with Machine Learning

As part of the agricultural sector's digital transformation, machine learning techniques have gained increasing attention due to their ability to automate complex tasks such as e.g. yield prediction, crop type mapping, and weed detection (Benos et al., 2021). The core idea of machine learning is to 'learn' from the data without explicitly being instructed by a human. Among core machine learning methods (unsupervised, supervised, semi-supervised and reinforcement learning), supervised machine learning that requires labelled data for the 'learning' process is the most widely used in crop type classification research.

Supervised machine learning for crop type classification typically consists of four main stages: input data generation, model building, model performance evaluation and crop type map generation (Figure I-1). Each stage consists of a multistep workflow that varies depending on the study design, data availability and spatial-temporal extent of the study.

The input feature generation includes several pre-processing steps that transform raw data into analysis ready data (ARD). As the literature shows, the type of applied pre-processing steps affects the quality of output map products (Inglada et al., 2016; Sun et al., 2020). Apart from core remote sensing pre-processing steps (e.g.,

atmospheric correction for optical data or, e.g., radiometric correction for SAR data), such techniques as gap-filling and monthly and seasonal composite calculation could be required for generating consistent features over large geographic areas.

Another essential data pre-processing step is splitting reference samples into training, validation and test sets (Figure I-1). As the name suggests, training data is used for training a machine learning classifier. It has been shown that the quality (Foody et al., 2016) and quantity (Foody et al., 1995) of training samples substantially impact the model performance and, in some cases, even more than the selection of the classification algorithm itself (Maxwell et al., 2018). The validation data is typically used for getting the model performance estimates while tuning its hyperparameters. Since validation data is used for model adjustments, it cannot be further used for deriving unbiased estimates of model performance. The independent test set is used exclusively for final classification accuracy estimation. The issue of acquiring reliable quality estimates of map products is widely discussed in the remote sensing community (Stehman and Foody, 2019).
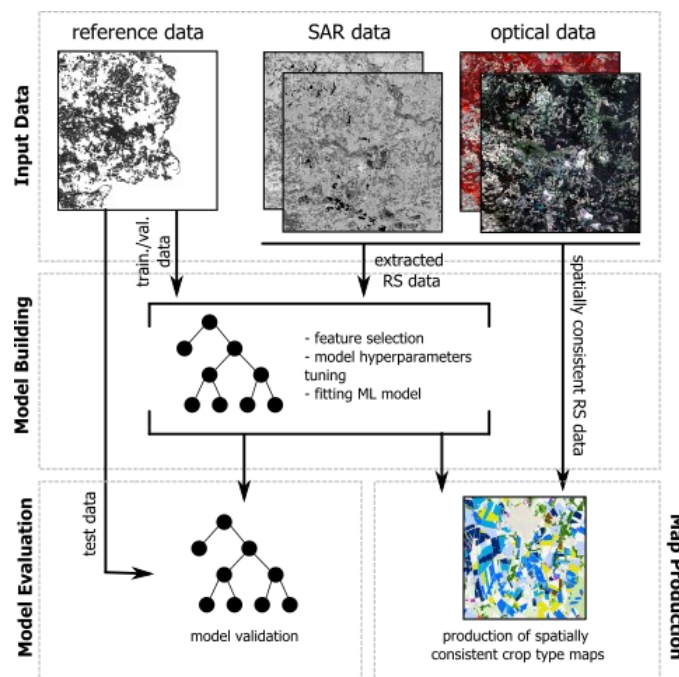
Figure I-1. The supervised machine learning workflow for crop type mapping with optical-SAR data combination.

The most common strategy of splitting reference data to training-validation and test sets is splitting based on a split ratio (e.g., 70/30, 80/20), where the larger share is used for model training, and the smaller share is used for model evaluation. However, Lyons et al. (2018) argue that a multiple partitioning (e.g. cross-validation) strategy should be preferred over a single hold-out split to acquire

unbiased accuracy estimates. The sampling design and data partitioning are widely discussed topics within the remote sensing community (Stehman and Foody, 2019).

While the quality and quantity of input data significantly impact the construction of an accurate machine-learning classifier, the type of classifier and the way it was built is another crucial aspect to consider. The algorithm that assigns class labels to unlabelled instances is called a classifier (Bishop, 2006). The numerous comparative studies on commonly used machine learning classifiers have shown the Random Forest (Breiman, 2001b) algorithm's superiority over others due to its ability to handle large feature spaces and being less sensitive to the noise in the training data (Belgiu and Drăgu, 2016). Random Forest is one of the most used algorithms for crop type classification (Asam et al., 2022; Griffiths et al., 2019; Preidl et al., 2020). A comprehensive review of the application of the classifier within the remote sensing domain is given by Belgiu and Dragu (2016).

The generation of crop type maps is the last stage in the workflow for research studies whose objective is the production of crop type maps (e.g., Inglada et al., 2017). The map production stage could be omitted in some investigative studies that only use sample pixels or objects to test the raised hypothesis (e.g., Waldner et al., 2019). The map production stage's complexity increases with the study site's spatial extent and the cropping pattern's complexity that should be mapped. Generating consistent features across whole study sites typically requires large processing and storage capacities. But, with the recent advancements of cloud computing in remote sensing, producing seamless maps across large regions or even globally has become more feasible (Azzari et al., 2017).

2.4.   *Spatial Transferability of Crop Classification Models*

The first law of geography states that 'near things are more related than distant things' (Tobler, 1970), which emphasizes the main property of spatial data: spatial dependence. Spatial dependence is inherent to geospatial information, such as remote sensing data. The presence of spatial dependence in the data violates the assumption of the independent and identical distribution of samples, which is one of the central machine learning concepts (Nikparvar and Thill, 2021). The effect of spatial autocorrelation, a measure of the similarity of nearby observations, in remotely sensed data is well-known and widely discussed in the research community (Congalton, 1991).

Large-scale crop type mapping using spatially limited reference data assumes the spatial transfer of the trained models to previously unseen geographical spaces. Such transfer over space often significantly reduces the model's performance. The main reason for the spatial non-transferability of supervised machine learning models is

the absence of representative crop type samples, samples that adequately reflect the characteristics of the entire crop class population, in the region where the model is being spatially transferred, i.e. target region. Solving this issue by acquiring the appropriate datasets from target regions is often impossible due to the data acquisition costs (time, finance and general accessibility of the region).

The spatial transferability of statistical models is an old and certainly everlasting topic (Geary, 1954). Since the 2000s, this issue has gotten more attention with the rise of the species distribution modelling domain (Wenger and Olden, 2012). Studies on the spatial transferability of machine learning models emphasized the pitfalls of ignoring the spatial dependencies present in the data (Ferraciolli et al., 2019; Overmars et al., 2003; Randin et al., 2006; Rocha et al., 2018). When feeding a machine learning classifier with remote sensing time-series features, we desire the model to learn distinct phenological phase differences among various crop types and be able to assign correct labels to unknown pixels or objects. While this could be a valid strategy for predicting under the environmental conditions of the training sites, it is not always suitable to predict distant geographical spaces where e.g. a shift of phenological phases may occur for several days or even weeks. Predictions outside the 'known' to model environments and geographical spaces can result in substantial accuracy losses, as illustrated in recent studies (Karasiak et al., 2019; Ploton et al., 2020).

While the absence of reference data from the target region is one of the main reasons for increased misclassification, the quality, quantity and type of remote sensing data can also affect the spatial transferability of the classifier. For instance, Inglada et al. (2016) confirmed that errors in gap-filled optical features significantly impact classification accuracy, especially when these errors coincide with the training samples. In large-scale mapping, due to altering atmospheric conditions during image acquisitions across different sensor swaths, the gap-filling is unavoidable to get consistent optical features across the entire study site. When using optical features to map large-scale study sites, we continuously face spatially clustered noise due to undetected clouds or cloud shadows. Such issues are not present in SAR data due to its weather-independent data acquisition capabilities. This suggests that models based on SAR features may show more robust performance in spatial transfer than those based on optical data.

3. Objectives of the Thesis

The overarching objective of this thesis is to explore the potential of the combined use of optical data from Sentinel-2 and SAR data from Sentinel-1 for crop type mapping. Together, these satellites deliver

unprecedented high spatial-temporal resolution remotely sensed information and enable global agricultural land monitoring. The combined use of these datasets allows simultaneously utilising the spectral reflectance and the SAR backscatter to build more accurate machine learning models for crop type mapping. The following research topics guided the setup of the present thesis:

- Review the currently available research on combining optical and SAR data for crop type mapping and outline common findings and possible research gaps;

- Large-scale mapping of a wide variety of crop classes using dense time-series optical and SAR features;

- Investigating the spatial transferability of machine learning models using optical-SAR combination over single sensor data.

First, an in-depth review of the published studies was conducted (Chapter II). Since the 2000s, the number of research papers focusing on the optical-SAR combination for crop type mapping has rapidly increased (Orynbaikyzy et al., 2019). The studies investigated diverse aspects of the crop type classification, such as early crop type mapping (e.g. Inglada et al., 2016), selection of best-suited features (e.g. Inglada et al., 2016), large-scale mapping (e.g. Griffiths et al., 2019), etc. The review work presented in section II focuses on summarizing existing knowledge based on outcomes of the reviewed research publications. The following objectives were set for the review study:

- provide a general overview of research studies on crop type mapping using optical and SAR data published from 1972 to late 2018;

- identify similarities and trends in the fusion methods and classification approaches;

- outline common and conflicting findings;

- define the research gaps and possible future development directions.

Based on the outcomes of the review study, two research directions were delineated: mapping a wide variety of crops with dense time-series data at a large scale (here, an area covering more than 10.000 km^2) and spatial transferability of the machine learning models for crop type mapping. The first research objective was addressed in Chapter III. The predictive performance of Random Forest models using bi-weekly time series of optical-SAR features was compared to single sensor features to classify 16 crop classes in Brandenburg, Germany. The effect of feature selection, parcel sizes, location of the pixels within the field and optical data availability on the classification

were investigated. More precisely following research questions were addressed:

- How does the optical-SAR combination perform (here, accuracy) compared to single sensor data for crop type mapping?

- Would the reduction of feature space improve the performance of Random Forest models? Which features show high relevance for the classifier?

- How do parcel sizes, pixel location within the field and optical data availability affect the classification performance?

The spatial transferability of crop type classification models was another research gap identified during the review. Unless representative reference data is disturbed across whole study sites, as in Chapter III, large-scale mapping assumes the spatial transfer of the pre-trained model to unseen geographic spaces. Previous studies did not address the spatial transferability of models based on the optical-SAR combination. In Chapter IV, the following research questions were addressed:

- How much does the classification accuracy decrease in the target system compared to the reference system?

- Which dataset (SAR, optical or optical-SAR combination) would perform better in terms of spatial transferability?

- How does the application of spatial feature selection affect the spatial transfer ability of Random Forest models?

The following three chapters address the above-described research objectives (II-IV). Each of these chapters is a self-contain manuscript that was published in peer-reviewed international journals:

- Orynbaikyzy, A., Gessner, U., Conrad, C., 2019. Crop type classification using a combination of optical and radar remote sensing data: a review. Int. J. Remote Sens. 40, 6553–6595.
https://doi.org/10.1080/01431161.2019.1569791

- Orynbaikyzy, A., Gessner, U., Mack, B., Conrad, C., 2020. 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. 12. https://doi.org/10.3390/RS12172779

- Orynbaikyzy, A., Gessner, U., Conrad, C., 2022. Spatial Transferability of Random Forest Models for Crop Type

Classification Using Sentinel-1 and Sentinel-2. Remote Sens. 14, 1493. https://doi.org/10.3390/rs14061493

4. Structure of the Thesis

The thesis is based on three research papers published in peer-review international journals in 2019, 2020 and 2022. First, the overall motivation, objectives and scientific background of the thesis and its organization are presented in Chapter I. The following three chapters correspond to three research publications (Chapters II, III, and IV). The last chapter (Chapter V) presents a comprehensive discussion of all papers and the main findings of the thesis.

II

CHAPTER II: CROP TYPE CLASSIFICATION USING A COMBINATION OF OPTICAL AND RADAR REMOTE SENSING DATA: A REVIEW

III

CHAPTER III: 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

Abstract

Crop type classification using Earth Observation (EO) data is challenging, particularly for crop types with similar phenological growth stages. In this regard, the synergy of optical and Synthetic-Aperture Radar (SAR) data enables a broad representation of biophysical and structural information on target objects, enhancing crop type mapping. However, the fusion of multi-sensor dense time-series data often comes with the challenge of high dimensional feature space. In this study, we (1) evaluate how the usage of only optical, only SAR, and their fusion affect the classification accuracy; (2) identify the combination of which time-steps and feature-sets lead to peak accuracy; (3) analyze misclassifications based on the parcel size, optical data availability, and crops' temporal profiles. Two fusion approaches were considered and compared in this study: feature stacking and decision fusion. To distinguish the most relevant feature subsets time- and variable-wise, grouped forward feature selection (gFFS) was used. gFFS allows focusing analysis and interpretation on feature sets of interest like spectral bands, vegetation indices (VIs), or data sensing time rather than on single features. This feature selection strategy leads to better interpretability of results while substantially reducing computational expenses. The results showed that, in contrast to most other studies, SAR datasets outperform optical datasets. Similar to most other studies, the optical-SAR combination outperformed single sensor predictions. No significant difference was recorded between feature stacking and decision fusion. Random Forest (RF) appears to be robust to high feature space dimensionality. The feature selection did not improve the accuracies even for the optical-SAR feature stack with 320 features. Nevertheless, the combination of RF feature importance and time- and variable-wise gFFS rankings in one visualization enhances interpretability and understanding of the features' relevance for specific classification tasks. For example, by enabling the identification of features that have high RF feature importance values but are, in their information content, correlated with other features. This study contributes to the growing domain of interpretable machine learning

1. Introduction

Crop type maps deliver essential information for agricultural monitoring and are likewise relevant for other fields such as environmental assessments. Respective classification approaches using Earth Observation (EO) stand to benefit from the availability of high-resolution Sentinel-1 and Sentinel-2 dense time series.

56

Particularly, the synergistic use of these optical and Synthetic-Aperture Radar (SAR) datasets bears high potential. In previous research, single-sensor approaches based on optical and SAR data were used the most to map crop types at different scales. A larger share of research was based on optical data (Belgiu and Csillik, 2018; Conrad et al., 2010; Defourny et al., 2019a; Griffiths et al., 2019; Inglada et al., 2015). The classification of crop types using only SAR data has also been successfully implemented in several studies (Bargiel, 2017; Clauss et al., 2018; Kenduiywo et al., 2018; McNairn et al., 2014). Optical data that makes use of the visible, near-infrared, and short-wave-infrared portion of the electromagnetic spectrum, provides valuable information about leaf pigments, water content, and plants' overall health condition.; whereas SAR data, dependent on the frequency and polarization, delivers a complex representation of canopy structure, surface roughness, soil moisture, and topography (H and B, 2004).

The first studies that focused on optical-SAR data fusion date back to the 80s (Li et al., 1980; Ulaby et al., 1982). However, starting from 2000, the number of studies on this topic gradually increased, which can be partially explained by the launch of new space-borne radar (ERS, RADARSAT, ENVISAR, ASAR, etc.) and optical (Landsat, SPOT, IRS, MODIS, QuickBird, etc.) satellites (Orynbaikyzy et al., 2019). The launch of Sentinel-1 (2014/2016) and Sentinel-2 (2015/2017), operated by the European Space Agency (ESA), boosted the interest in using these freely available SAR and optical datasets for crop type mapping in a synergistic way. Optical and SAR data's complementary nature gives the possibility to simultaneously utilize information on plants' structural and bio-physical conditions. This explains that most previous fusion studies reported improvements in crop classification accuracy when optical and SAR data were combined compared to single sensor experiments (Denize et al., 2018; Forkuor et al., 2014; Torbick et al., 2017). Commonly, optical-SAR fusion studies used several cloud-free optical and available SAR scenes for their studies (Orynbaikyzy et al., 2019). The use of dense time-series of both became more common in recent studies (Demarez et al., 2019; Van Tricht et al., 2018). However, for large-scale studies, it is still a challenge to combine these two datasets since vast volumes of diverse datasets demand higher processing power and resources.

There are three primary data fusion levels: pixel-level, feature-level (further, feature stacking), and decision-level (Pohl and Van Genderen, 1998) (for details see Section 2.4.1). Among crop type classification studies, Gibril et al. (Gibril et al., 2017) compared pixel-level fusion techniques (Brovey Transform, Wavelet Transform, Ehlers) with feature stacking results. To the best of our knowledge, the comparison between optical-SAR feature stacking and optical-SAR fusion at decision-level was not subject to any study even though decision fusion was successfully applied (Waske and Van Der Linden,

2008). Multi-sensor feature stacking may result in high dimensional feature space, which may negatively affect the classification accuracy. While decision fusion, based on classification confidences derived from single sensor predictions, might profit from less complex models. It could be valid to expect that optical-SAR fusion at the decision-level could be more performant than simple feature stacking. To test this hypothesis, in this study, we compare the classification results derived from optical-SAR feature stacking and optical-SAR fusion at decision-level (further, decision fusion).

Combining dense optical and SAR time series features can quickly result in a high-dimensional feature space that can pose challenges for pattern recognition and machine learning. In terms of classification accuracy, earlier methods such as the parametric maximum likelihood classifier are more susceptible to high-dimensional feature spaces than to state-of-the-art classifiers, such as Support Vector Machines and Random Forest (RF). Nevertheless, studies have shown that these classifiers' accuracy can also be increased by feature selection (Löw et al., 2013), particularly when the amount of training samples is limited (Jain, 1997). Elaborated pre-processing, feature generation, and selection of suitable features are needed for many classification algorithms. Commonly derived features as spectral-temporal variability metrics (Müller et al., 2015), time-weighted interpolation (Inglada et al., 2015), or computationally more expensive approaches such as time-weighted dynamic time warping (Maus et al., 2019) are being used less often for large-scale applications. This happens probably due to, among other reasons, high computational complexity, particularly over very large areas.

Feature selection methods have a long tradition in pattern recognition analysis of remote sensing data (Goodenough et al., 1978; Richards, 2005), alleviating the challenges mentioned earlier. It helps to train more accurate models, decrease computational complexity, and improve the understanding of the used features (Roscher et al., 2020; Yu et al., 2016). Data and model understanding is essential in the development of operational products (Inglada et al., 2017) and in scientific research where feature importance and rankings are often essential elements (Immitzer et al., 2019; Roscher et al., 2020; Sitokonstantinou et al., 2018). The main groups of feature selection are filters and wrappers methods (Guyon and Elisseeff, 2011). Filter methods use various statistical measures (e.g., Chi-square test, Pearson's correlation) to score the relevance of features. Whereas, wrapper methods perform feature selection based on the chosen classifiers' performance, which enables to select most performant features with low correlation. The main drawback of the wrapper approaches is the high computational costs. In crop type classification studies, embedded methods such as classifier specific RF feature importance is often used to select the most important features (Belgiu

and Drăgu, 2016). It is a fast approach, but a drawback is that features with correlating information content can show similarly high importance scores.

To alleviate the limitations of wrapper methods concerning high computing efforts, in this study, we perform a group-wise forward feature selection (gFFS). gFFS allows focusing the analysis and interpretation on useful feature sets like spectral bands, vegetation indices (VIs), or data acquisition dates (for details see Section 2.4.3). A similar approach was followed by Defourny et al. (2019b). The grouping strategies of features in the gFFS have two important benefits: 1) gFFS allows to tailor the feature selection towards better interpretability and supports a more efficient feature selection process. For example, the time-steps with most discriminative power can be better analyzed by considering all information available at a particular time-step as a group and not only as single features. 2) gFFS allows the substantially reduced computational time of the feature selection step while considering all features at hand.

In addition to an appropriate feature selection, the quality of the satellite data features concerning data availability and gaps are crucial aspects that can influence classification accuracy. Croplands are typically characterized by management activities such as tilling, sowing, and harvesting or cutting and are, therefore, amongst the most dynamic land cover elements. Their successful identification via remote sensing data often requires dense time-series information capable of capturing critical plant development phases and the mentioned land management activities (Veloso et al., 2017). The comparison of seasonal, monthly composites, and gap-filled Harmonized Landsat and Sentinel-2 (Claverie et al., 2018) time-series (10-days interval) data by Griffiths et al. (Griffiths et al., 2019) showed that the highest classification performance was achieved with gap-filled 10-day time-series data. However, the quality of the gap-filled data depends on the duration and number of gaps. Gaps originate mainly from cloud cover, cloud shadows, and other atmospheric effects from which optical data are often suffering. In this sense, combining dense time series of SAR features and gap-filled optical data could result in more accurate classifications.

In summary, despite the large number of studies focusing on the combination of optical and SAR features for crop type mapping, the following aspects were not sufficiently investigated: (1) combined performance of a comparably high number of relevant Sentinel-2 and Sentinel-1 dense time-series features for larger areas covering more than one sensor swath, where the derivation of spatially homogeneous features across all the study area is required; (2) impact of optical-SAR feature stacking and decision fusion on classification accuracies; (3) analysis and interpretation of feature importance and ranking based on dates, bands, and VIs instead of single features and their effect on

classification accuracies; (4) the performance of respective classification approaches to differentiate 16 crop types; (5) analysis of the influence of the optical data availability, parcel size, and pixel location within the parcel on the classification results.

This study aims to contribute to these open issues by evaluating how the synergetic use of dense optical and SAR time-series data derived from Sentinel-1 and Sentinel-2 improves the classification accuracy of typical crop types of Central Europe. The study site covers the Brandenburg state, located in northern Germany. In this context, we investigate how a large number of optical and SAR features affect the performance of RF models and evaluate the relevance of various feature-sets and time-steps to the classifier. It is also analyzed how agricultural parcel size, mixed pixels occurrence at parcel borders, and the non-availability of optical satellite data at specific dates of the year affect the classification accuracy. More explicitly, the following questions were addressed:

- How do the usage of single-sensor dense time-series data and the fusion of Sentinel-1 and Sentinel-2 data affect the classification accuracy? Do classification accuracies based on optical-SAR feature stacking differ from decision fusion?
- How does a high dimensionality of optical-SAR feature stack impact on the performance of RF? Which features are most relevant, and which dates and bands or VIs lead to the highest accuracies?
- What is the influence of cloud-related gaps in optical data, how do parcel sizes, and the pixel location within a parcel affect classification accuracy?

2. Materials and Methods

2.1. Study Area

The study site (Figure 1) is located in Northern Germany and covers the territory of Brandenburg state with an area of 29,654 km², where 45% is agricultural land (Gutzler et al., 2015). Large-scale farms dominate the state croplands with 238 ha of average field size (Gutzler et al., 2015). The area has low topographic complexity with the highest peak in the state being at 201 m.a.s.l. (Kutschenberg hill) and the lowest point is Rüdersdorfer opencast mining area with −46.5 m below sea level. The average annual precipitation is 719 mm, and the average annual temperature is 9.9 °C. Winter cereals are sown at the beginning of September and harvested at the end of July and beginning of August. Summer crops are sown at the end of March—beginning of April and harvested at the end of July—beginning of August (German National Weather Service, 2019).
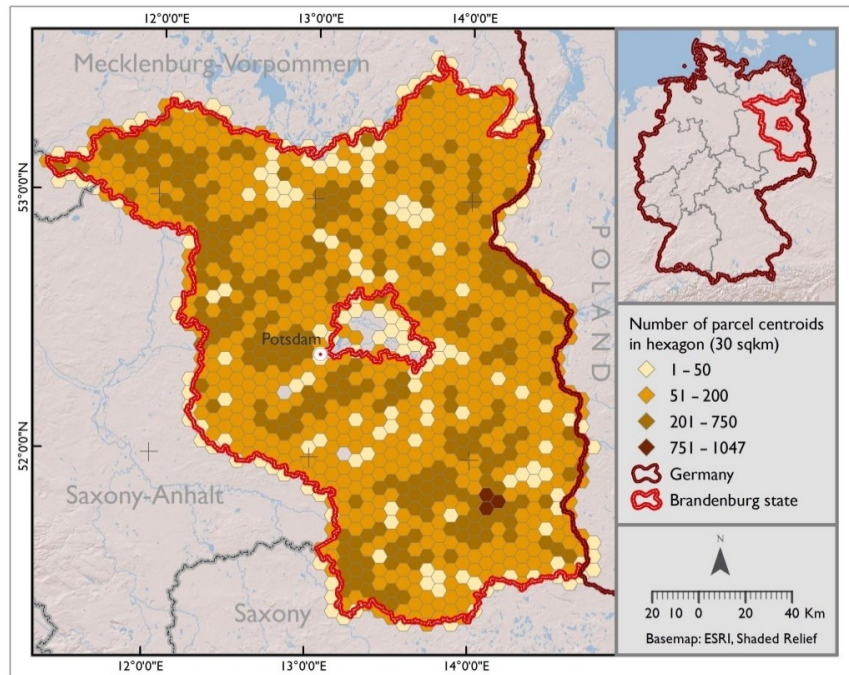
Figure III-1. The study area Brandenburg with the density of agricultural parcels according to Land Parcel Identification System (LPIS) data for the year 2017.

2.2. Reference Data

As ground truth data on crop types for the year 2017, we used reference data from the Brandenburg Surveying and Geospatial Information Office (Landesvermessung und Geobasisinformation Brandenburg) web-portal (Brandenburg Surveying and Geospatial Information Office, 2017). This reference parcel data, managed by Land Parcel Identification System (LPIS), is based on the reports of farmers who applied for agricultural subsidies in the frame of the European Union's (EU) Common Agricultural Policy (CAP). Further, we refer to this dataset as LPIS. The data contains parcel boundaries and crop types harvested in the year 2017.

The original dataset is available in a geospatial vector format and contains 161,503 parcels. Overlapping parcels and parcels with an area less than 1000 m² were excluded from the reference data (1364 parcels, 3017.516 ha). Out of the 158 original crop types, several crop classes were merged into grouped classes based on their biological plant family membership and phenological similarity. For example, class maize includes silo maize, maize for biogas, maize with flowering path, etc., and class potatoes comprise starch potatoes and potatoes for food. Supplementary material A gives an overview of the original and grouped classes. Only those grouped crop classes were selected from the original reference data, which accounted for at least 0.5% of the full LPIS area. The only exception was made for the class sugar beets,

which areal cover was close to the threshold (0.46%). This resulted in 16 crop classes (Table 1) for which a total of 134,379 parcels (1,220,160.86 ha) was available in the LPIS data. Since crop groups such as winter cereals (winter wheat, winter rye, winter rape, winter barley, winter triticale), summer cereals (summer barley, summer oat), and legumes (legume mixture, peas-beans, lupins) are expected to show high intra-class confusion, we additionally report the accuracy results when these classes are grouped into one. The present study focuses only on crop type classification. No other land cover land use classes were considered.

Table III-1. Overview information for the crop types considered in this study.

Crop Type	Number of Parcels	Average Parcel Size [ha]
Permanent grasslands	59,182	4.94
Temporal grasslands	12,092	3.77
Maize	14,449	14.27
Sunflowers	834	12.26
Potatoes	1015	9.17
Sugar beets	240	24.86
Winter wheat	9758	17.59
Winter rye	14117	11.45
Winter rape	6299	20.09
Winter barley	5189	17.36
Winter triticale	3289	11.01
Summer barley	934	7.49
Summer oat	2394	5.91
Legume mixture	2297	8.98
Peas-Beans	897	10.89
Lupins	1393	8.70

2.3. Remote Sensing Data Pre-Processing and Features Generation

The data sensed by the Multi-Spectral Instrument (MSI) onboard Sentinel-2A/B, and by the C-band synthetic aperture radar (SAR) instrument onboard Sentinel-1A/B were analyzed in this study. The data were accessed via The Copernicus Open Access Hub (https://scihub.copernicus.eu/). The tiling grid of Sentinel-2 data was used as a base grid (Figure 2.). Brandenburg's entire territory is covered by eight Sentinel-2 tiles. Overall, 494 optical scenes and 473 SAR scenes (in ascending mode) were utilized with temporal coverage from the beginning of January until the end of September 2017. Optical data acquired in October were entirely excluded from the analysis due to the lack of any scene with cloud cover below 80%.
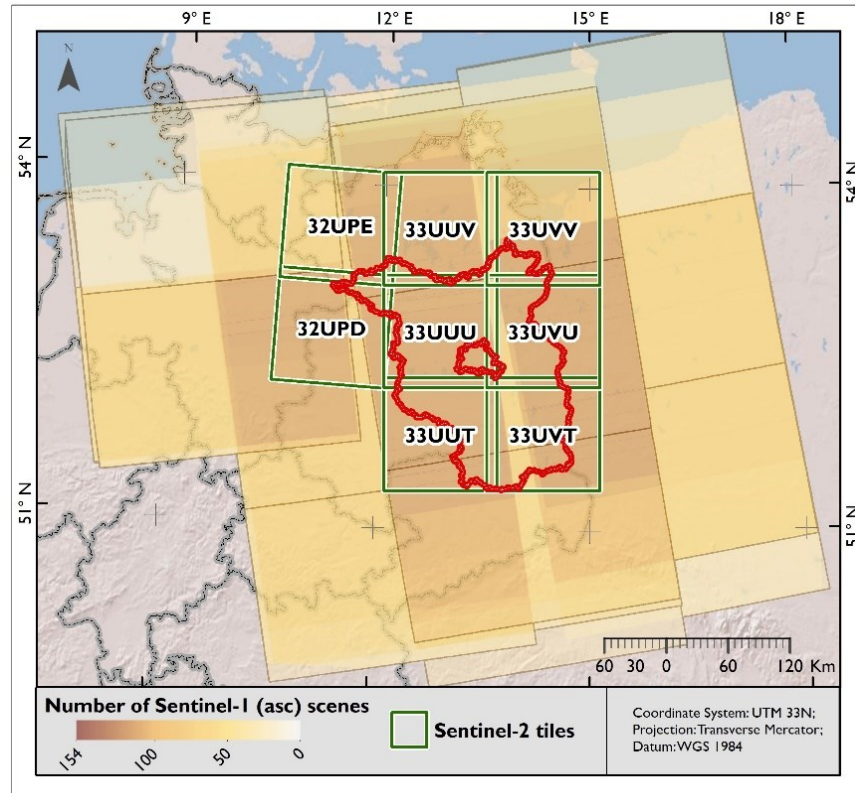
Figure III-2. Overview of Sentinel-2 tiles covering Brandenburg and available Sentinel-1 SAR data (ascending mode) in the study region.

2.3.1. Optical Data Pre-Processing and Gap-Filling

The individual pre-processing steps that were applied to Sentinel-2 data are shown in Figure 3. Sentinel-2 data at Level-1C (top of atmosphere) were processed to Level-2A (bottom of atmosphere) using sen2cor v2.4.0 (Müller-Wilm, 2016). Ten Sentinel-2 bands were used for further analysis. Bands 1 (coastal aerosol), 9 (water vapor), and 10 (Short-Wave-Infrared (SWIR)-cirrus) were excluded from the analysis because of their irrelevance for crop type mapping. The data of the red-edge (5, 6, 7), near-infrared narrow (8A), and SWIR bands (11, 12) were resampled from 20 to 10 m spatial resolution using the nearest neighbor algorithm. In addition to the original Sentinel-2 bands, four well-known VIs were generated: Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Normalized Difference Yellow Index (NDYI), and Plant Senescence Reflectance Index (PSRI) (Hatfield and Prueger, 2010; Sulik and Long, 2015).

Cloud masks were produced in two steps. First, cloud masks were calculated using the fmask (Zhu et al., 2015) extended for Sentinel-2 data, according to Frantz (2018). Second, we combined information from the Scene Classification Layer (SCL) generated by sen2cor (Müller-Wilm, 2016) and the output of the fmask into a single binary

invalid pixel mask. We flagged a pixel as invalid if at least one of these two input layers detected cloud, cloud shadow, snow, defect, saturated pixels.

Different acquisition times, clouds and cloud shadows lead to irregular time series of valid observations over the study area. Instead of using generic ways of handling missing data within RF (Tang and Ishwaran, 2017), we build a consistent gap-free time series over the whole study site with time-weighted linear interpolation. It allows accounting better for the temporal information contained within the original satellite time series data. Dense time series were created based on invalid pixel masks, four VIs, and the 10 bands of Level-2A data at 10 m spatial resolution. Following the approach of Inglada et al. (Inglada et al., 2015), gap-filling started by defining bi-weekly target dates from January to September 2017. In total, we defined 20 target dates. After applying the invalid pixel mask, band wise time-weighted linear interpolation was performed considering only valid pixels to fill the defined target dates.
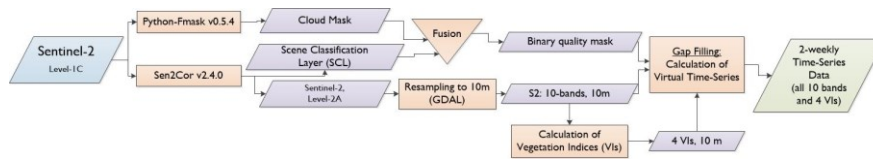
Figure III-3. Sentinel-2 data pre-processing and feature generation.

2.3.2. SAR Data Pre-Processing

C-band Level-1 Ground Range Detected (GRD) Sentinel-1 products, acquired in Interferometric Wide (IW) swath mode, were accessed via the Google Earth Engine (GEE) platform (Gorelick et al., 2017). The data available in GEE were pre-processed with the Sentinel-1 toolbox from ESA, which involved updating orbit metadata, thermal noise removal, radiometric calibration, and terrain correction. For the scenes in ascending mode, we filtered extreme incidence angles so that only observations with incidence angles of 32° to 42° were used. The average incidence angle in the study area is equal to 37°. Lee speckle filtering and square cosine correction were applied as it was outlined by Tricht et al. (Van Tricht et al., 2018). After the incidence angle correction, we calculated bi-weekly medians and matched the time steps to the 20-target time-series dates of the optical features (Figure 4).
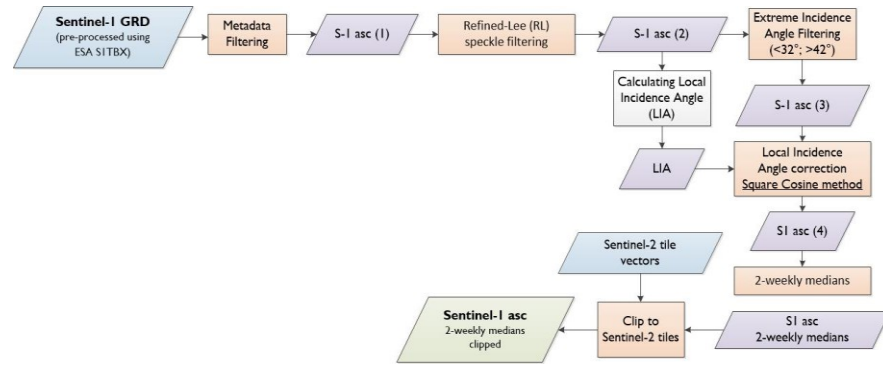
Figure III-4. Sentinel-1 data pre-processing and feature generation.

3. Methodology

The methodological workflow of this study consists of the following steps (Figure 5): (1) extraction of optical and SAR time-series features at the pixel level (see Section 2.3 for more details); (2) sampling of training and testing pixels; (3) performing group-wise forward feature selection (gFFS), where individual features are grouped by time or variable respectively; (4) building RF models using all existing features and the best-performing feature subsets identified in (3); (5) predicting test-sets; (6) extracting accuracy metrics; (7) analyzing results using the information of RF feature importance and feature group ranking; (8) analyzing auxiliary data on the parcel size, optical data availability, temporal profiles of correctly and misclassified pixels.
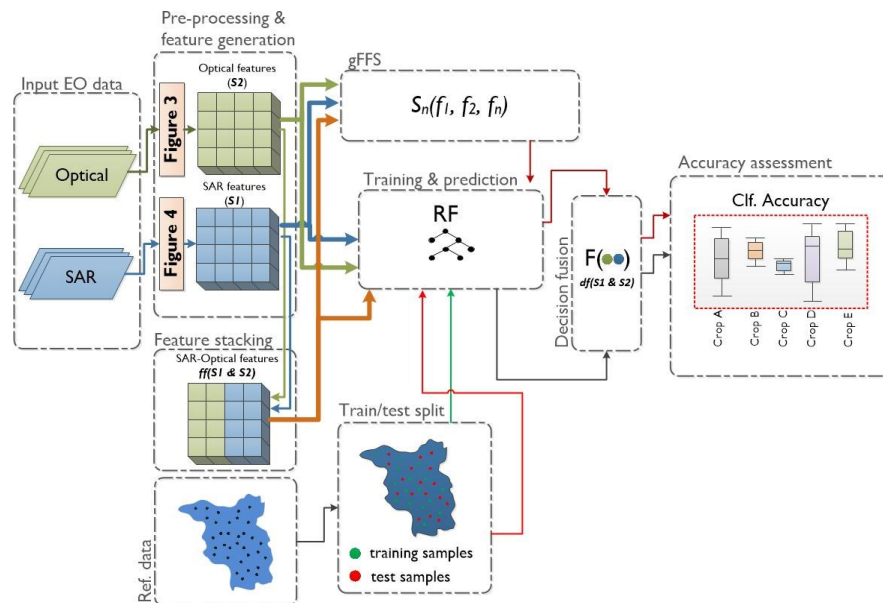
Figure III-5. Schematic overview of the methodological workflow.

3.1. Single Sensor Features Versus SAR-Optical Combination

To compare and evaluate the accuracy achieved with single sensor features and their fusion, we separately classified only optical features, only SAR features, and a fusion of optical and SAR features. As optical features (in plots shown as **S2**), we selected 10 Sentinel-2 spectral bands and four VIs, as described in Section 2.3.1. In total, 20 time-steps of 10 spectral bands and four VIs summed up to 280 optical features. As for SAR features (in plots shown as **S1**), bi-weekly medians of VV and VH bands were used (pre-processing steps in Section 2.3.2.). They summed up to 40 SAR features.

According to Pohl and van Genderen (Pohl and Van Genderen, 1998), there are three levels of data fusion: pixel-level, feature-level, and decision-level. At the pixel-level fusion, multi-sensor input data are fused into a new dataset using various compression, dimensionality reduction methods (e.g., Principal Component Analysis, Wavelet-based approaches, Brovey Transform) and then used for prediction or analysis. The feature-level fusion implies combining extracted features from different sensors to form a new multi-source feature stack. In this study, we will refer to this method as a feature stacking. Such multi-sensor feature stacking is more commonly used in crop type classification studies (Orynbaikyzy et al., 2019). Joshi et al., (Joshi et al., 2016) refers to these two data fusion levels as 'pre-classification or -modeling fusion'. Whereas 'post-classification or -modeling fusion' would be a decision-level fusion, which is performed by fusing the classification results of single sensor features based on pre-delineated rules or decisions. In this study, we perform optical-SAR feature stacking and decision fusion.

The feature stacking was performed by stacking the abovementioned optical and SAR features into one optical-SAR feature stack (in plots shown as **ff(S1&S2)**). The decision fusion was done by fusing classifications derived from single-sensor features. We used the class probabilities of the RF to derive confidence values for **S1** and **S2,** respectively. In the scikit-learn implementation of RF (Pedregosa et al., 2011), the class probabilities were calculated as the mean predicted class probabilities of the trees in the forest. For a single tree, the class probability is defined by the fraction of samples of the same class in a leaf. The confidence level was calculated by the difference between the highest and second-highest class pseudo-probabilities. During decision fusion (in plots shown as **df(S1&S2)**), the prediction with the highest confidence level was selected as a final prediction.

3.2. Sampling Strategy

From the refined LPIS reference data (Section 2.2.), 50% of the parcels were selected for training, and 50% of the parcels for testing. Train-test split was done at the parcel level, to ensure that none of the

test-set pixels come from a parcel which already was chosen for training. Training and testing pixels were not chosen within a buffer of 1 (10 m) pixel distance from parcel borders to avoid mixed spectral signatures. Training and testing pixels were sampled using equally stratified random sampling with 3000 samples size per crop type. This sample size was chosen to competently represent the spectral and phenological variability of the classes under investigation. It was ensured that training and testing pixels were equally sampled from small and large parcels to avoid underrepresentation of the small parcels. For each training and testing sample, information about the parcel size, the number of valid and invalid optical observations per month, and the sample's distance to the parcel border was stored as auxiliary data. These datasets were further used for the analysis of the results (Section 3.3.).

3.3. Group-Wise forward Feature Selection

To evaluate the significance of the features and the effect of the high feature amounts on the accuracy of the RF classifier, we used a modified sequential forward feature selection (FFS) approach. FFS is one of the variations of the sequential feature selection (SFS) approach, which belongs to wrapper methods. The procedure starts with building several RF, each using only one of the available features (1st sequence). Based on the accuracies of these RFs, the best (e.g., in terms of accuracy) feature is selected and combined to sets of two features using all remaining features. Again, several RFs are generated, based on sets of two features (2nd sequence). This process is repeated, and with every iteration, a new feature is added until only one RF, including all features, is constructed. This is often done to get a full feature ranking and investigate if the accuracy decreases at a certain point while increasing dimensionality. Early stopping is also possible, e.g., by defining a desired number of features, by stopping when the peak performance is reached, or if the accuracy of a new iteration does not increase significantly compared to the previous iteration (Saeys et al., 2007).

Wrapper methods such as FFS are computationally expensive and often impractical to perform when large amounts of features are used. It is particularly true for studies using dense time-series like the presented study. For example, to run a complete FFS with the 320 features considered in this study, it would need 51,360 model evaluation runs (Equation (1)), where one model evaluation comprises to train an RF, predict validation data, and calculate the accuracy of the model. Moreover, with 5-fold cross-validation, it sums up to 256,800 model evaluation runs (Equation (2)).

$$\sum_{i=1}^{320} i = 51{,}360$$

$$\sum_{i=1}^{320} 5i = 256{,}800$$

Since the focus of our study is not to evaluate the significance of individual features but rather to understand which time steps or spectral bands, VIs, and backscattering coefficients contribute the most to the accuracy of the classification model, we modified FFS in a way that also reduced computational efforts (Figure 6). The modification was done by performing group-wise FFS (gFFS), where features are grouped based on a time-step (further, time-wise gFFS) and on a variable (further, variable-wise gFFS). For example, in time-wise gFFS, the group '07-May' consists of 2 (VV, VH), 14 (all bands and VIs), and 16 (a combination of those) features in case of **S1**, **S2**, and **ff(S1&S2)**, respectively. Whereas the variable-wise gFFS considers the full time-series of a particular band or index as one entity (e.g., NDVI full time-series, VH full time-series). As a base of gFFS, we used the core implementation of the sequential feature selection (SFS) available in the MLxtend python package (Raschka, 2018). In total, time-wise gFFS required 1050 model evaluations for only optical, only SAR, and a combination of optical and SAR features. Variable-wise gFFS required 15 runs for SAR, 525 runs for optical, and 680 runs for optical-SAR features. These numbers already include the five-fold cross-validation runs, which were applied for receiving more robust estimates.
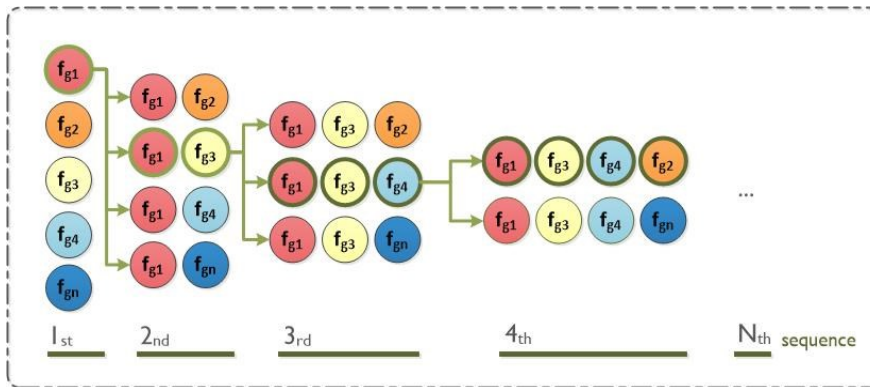
Figure III-6. Group-wise forward feature selection scheme, where f_{gi} is a group of features, e.g., f_{g1} = all features on date 1, f_{g2} = all features on date 2, etc. In the first sequence, f_{g1} is selected as the best (highest accuracy) of all feature groups. In the second sequence, each of the remaining feature groups is evaluated pair-wise together with already selected ones. The sequences are continued until the final set of features is evaluated.

3.4.   Classification Approach

The classification process was performed using state-of-the-art remote sensing image classifier-RF algorithms. It is a non-parametric machine learning algorithm consisting of an ensemble of randomized decision trees (Breiman, 2001a). Each decision tree predicts a target class for each training sample, and the class with the highest number of votes within the forest is selected as the ensemble's final decision. Previous studies focusing on crop type mapping (Conrad et al., 2014; Inglada et al., 2015; Onojeghuo et al., 2018) show that RF produces generally accurate results. Due to its robustness to class label noise and high dimensional input data (Belgiu and Drăgu, 2016), it is extensively used in crop type classifications (Forkuor et al., 2014; Inglada et al., 2017; Zhou et al., 2017).

In our study, we used the scikit-learn Python implementation of the RF algorithm. Based on the results of a randomized search using a five-fold cross-validation (Pedregosa et al., 2011), the following parameters were applied: (1) number of trees—700; (2) maximum depth—30; (3) maximum number of features used to split the node—square root of the sum of features; (4) minimum sample number to split a node—25.

The final RF models were built for optical, SAR, optical-SAR features using all features, and feature subsets selected based on the results of the time-wise and variable-wise gFFS (Section 2.4.3.). For all final RF models, the Gini importance score (also known as Mean Decrease in Impurity (MDI)) was used to evaluate the significance of the single features, further referred to as the RF feature importance. These feature importance scores were compared to the outcomes of the gFFS.

To assess and compare the classification accuracies, we computed class-specific metrics such as precision (i.e., user's accuracy), recall (i.e., producer's accuracy) and class-specific f1-score. The f1-score (Equation (3)) is a weighted average measure of precision and recall, where f1-score reaches the best values at 1 and worst values at 0. The average over the class-specific f1-scores has been computed to get a single accuracy metrics over all classes (Pedregosa et al., 2011). In addition, we calculated confusion matrices.

$$f1 = 2 * (precision * recall)/(precision + recall)$$

Accuracies were calculated for all 16 classes. To understand how accuracy is impacted by expected confusion among cereal and legume crops, we also calculated grouped classification accuracy while treating cereals and legumes as one crop class.

3.5.   Analysis of the Impact of Parcel Size, Pixel's Location within a Parcel, Optical Data Availability on Classification Accuracy

As was mentioned in Section 2.4.2., for each training and testing sample, we stored the information about the parcel size from which it was sampled. This data was then used to plot parcel size distributions for mis- and correctly-classified test samples.

The pixel's distance to the parcel border was calculated using the eo-box python package (Mack, 2019). Further, testing samples from individual crop types were grouped based on the distance to parcel borders, and then for each group, f1-score was calculated. Because of the parcel size variations, the number of samples in each group varies. Especially for more considerable distances such as 30−40 pixels away from parcel border, the underlining number of pixels used to calculate f1-score could be minimal. This results in expected significant variations on large distances. In the plot, the red vertical line indicates the distance at which 80% of samples have already been used to calculate f1-scores. Depending on the parcel size distributions of each crop type, the red vertical line switches along the x-axis.

Using the invalid pixel mask (see Section 2.3.1.), we calculated the number of valid optical observations per month for each sample. The numbers varied from 0 to 5, where 0 means that no valid optical observation was available for considered samples at the specific month.

The temporal profiles were built using NDVI and VH values for mis- and correctly classified samples.

4. Results

1.1.   Classification Accuracies (Overall and Class-Specific)

The f1-scores obtained from the predictions based on all features and the best-performing feature subsets selected using time-wise gFFS and variable-wise gFFS are shown in Table 2. The gFFS did not have any effect on the accuracy values for the experiments based on only SAR features (**S1**), optical features (S2), and decision fusion (**df(S1&S2)**). The decrease of f1-score by only 0.01 was recorded with time-wise gFFS for experiments based on optical-SAR features stacks (**ff(S1&S2)**). Also, in the class-specific accuracies, no significant changes were recorded (Supplementary Materials B). Based on these outcomes, we continue reporting classification accuracies based on the results when all the existing features were used.

The f1-score of the predictions based on **S2** and **S1** was equal to 0.61 and 0.67, respectively, without class-grouping. After grouping cereal and legume classes, f1-score derived from **S2**, and **S1** classifications increased to 0.73 and 0.76, respectively. There was no significant difference in the mean precision and recall values for single

sensor experiments (**S2**: precision—0.62, recall—0.61; **S1**: precision—0.68, recall—0.67).

Table III-2. Classification accuracies (f1-score) based on all features and the features subsets selected based on time-wise and variable-wise gFFS after grouping legume and cereal classes.

	All Features	Subset (Variable-Wise gFFS)	Subset (Time-Wise gFFS)
S1	0.76	0.76	0.76
S2	0.73	0.73	0.73
ff(S1&S2)	0.81	0.81	0.80
df(S1&S2)	0.80	0.80	0.80

The classifications based on ***ff(S1&S2)*** resulted in an f1-score of 0.72 with a precision of 0.73 and a recall of 0.72 without class-grouping. Decision fusion showed similar results (f1-score—0.71, precision—0.72, recall—0.72). Considering the grouped cereal and legume classes, f1-scores of both fusion approaches increased to 0.81, 0.80 for ***ff(S1&S2)*** and ***df(S1&S2)*** accordingly. Thus, both considered fusion approaches outperformed the single sensor accuracies.

Figure 7 gives an overview of the class-specific accuracies for a single sensor and fused feature results. In general, class-specific accuracies showed higher diversity compared to overall accuracies. Winter rape and sugar beet classes showed the best accuracy results (f1-score > 0.90), with only minor differences between the used sensors or fusion types.
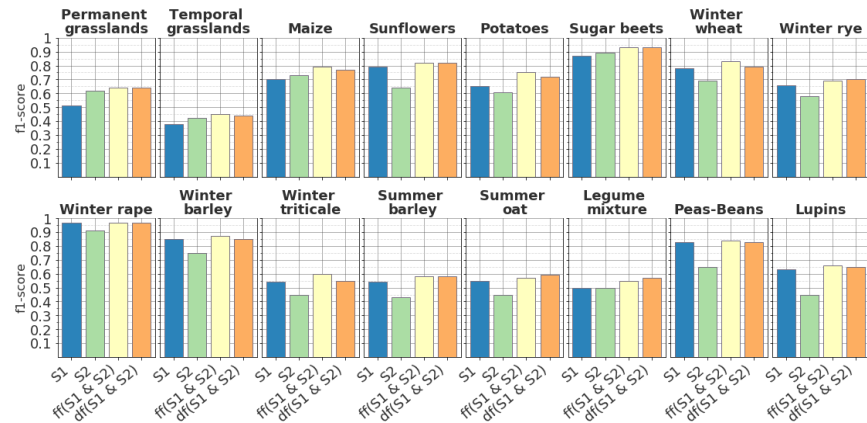
Figure III-7. Crop-specific accuracies derived from the classifications based on only SAR (***S1***), only optical (***S2***) features, and optical-SAR feature stacks (***ff(S1&S2)***) and decision fusion (***df(S1&S2)***).

Most crop types showed the highest accuracies when using ***ff(S1&S2)***, and were better classified with **S1** compared to **S2**. For example, this applies to all winter cereals (winter rape, winter wheat, winter rye, winter barley, and winter triticale) with f1-scores > 0.70. The same pattern could be seen for summer cereals (summer barley and summer

oat) and potatoes. Nevertheless, f1-scores remained below 0.6 for both summer cereals. Potatoes and maize had the highest f1-score with *ff(S1&S2)*, which were 0.75 and 0.79. For classes such as sunflowers, lupins, and peas-beans, the performance of models built using *S1* was quite close to the performance of the models built using *ff(S1&S2)* (f1-scores: $\Delta = 0.03$, $\Delta = 0.03$, $\Delta = 0.01$, accordingly) while using *S2* resulted in slightly lower accuracies (f1-scores: $\Delta = 0.18$, $\Delta = 0.21$, $\Delta = 0.19$, accordingly). In contrast, the two considered grassland classes had higher classification accuracies when using *S2* (f1-scores: temporal grasslands—0.42, permanent grasslands—0.62) compared to *S1* (f1-scores: temporal grasslands—0.38, permanent grasslands—0.51). This can also be seen from the map presented in Figure 8. Nonetheless, for these classes, the maximum accuracy score is reached with fused datasets (f1-scores *ff(S1&S2)*: temporal grasslands—0.45, permanent grasslands—0.64).
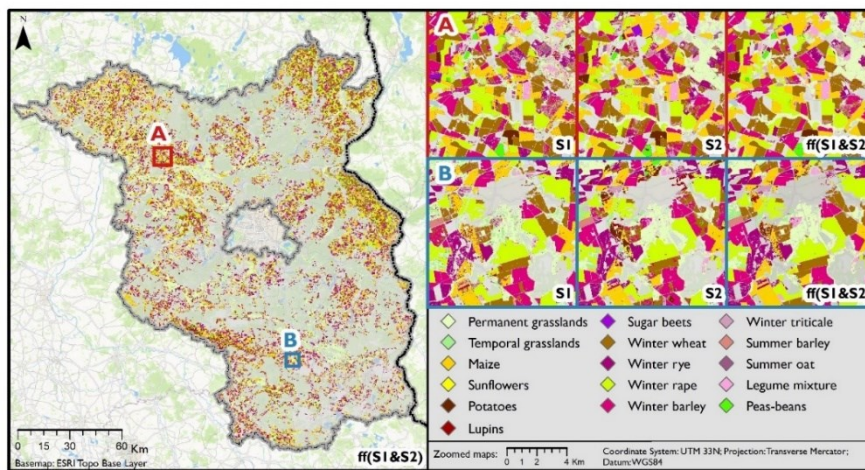
Figure III-8. Classification maps based on only SAR (S1) features, only optical (S2) features, their combination (ff(S1&S2)).

Permanent and temporal grasslands had a high within-group confusion rate (Figure 9). The test samples of legume mixture and temporal grasslands were often predicted as permanent grasslands, reflecting low precision and high recall values. Maize samples were often predicted correctly (recall = 86), but false predictions of sunflower, potato, and lupin samples as maize affected the precision, which was equal to 73. Summer and winter cereals formed two groups with high intra-group confusion. For winter cereals, higher confusion was present among classes such as winter wheat, winter rye, and winter triticale. Whereas, summer cereals were not only confused within the group but also often were classified as one of the legume classes. The confusion between summer cereals and legume classes are higher when using only optical feature compared to SAR only features. Supplementary Materials C includes all confusion matrices for all experiments.

72

After combing summer barley and summer oat (which were heavily confused, see Figure 10) into a single class of summer cereals, the f1-score increased to 0.78 with *ff(S1&S2)* (Figure 8). The classes lupins, peas-beans, and legume mixture, when grouped to one legumes class, got the highest f1-score of 0.77 with *ff(S1&S2).* The f1-score of class winter cereals raised above 0.90 with *ff(S1&S2)* after grouping. However, when merging classes, the general pattern, that optical-SAR feature combination and decision fusion were outperforming single sensor information, remained unchanged.

True labels

Predicted labels	Permanent grasslands	Temporal grasslands	Maize	Sunflowers	Potatoes	Sugar beets	Winter wheat	Winter rye	Winter rape	Winter barley	Winter triticale	Summer barley	Summer oat	Legume mixture	Peas-beans	Lupins	Prec.
Permanent grasslands	2,209	886	19	17	13	1	24	34	7	19	31	60	85	411	25	22	57
Temporal grasslands	528	1,376	22	13	18	0	32	85	0	25	65	131	166	593	15	45	44
Maize	14	50	2,578	125	139	61	11	9	8	4	14	123	91	86	46	153	73
Sunflowers	2	9	61	2,453	264	144	0	3	0	0	0	7	12	3	12	44	81
Potatoes	1	11	113	221	2,200	53	1	6	1	2	5	35	31	28	60	124	76
Sugar beets	0	2	11	37	49	2,717	1	0	0	0	0	0	3	3	9	25	95
Winter wheat	3	7	2	0	6	0	2,421	37	2	24	278	18	51	4	5	0	84
Winter rye	18	42	25	8	33	8	91	2,262	14	182	652	77	37	52	16	38	63
Winter rape	1	3	0	6	0	0	3	5	2,909	8	1	3	6	2	32	5	97
Winter barley	14	14	3	0	2	4	12	60	4	2,513	135	17	9	10	6	7	89
Winter triticale	11	24	12	1	9	0	235	297	1	136	1,635	26	38	22	0	9	66
Summer barley	9	32	12	9	17	0	60	54	4	21	66	1,550	423	30	16	69	65
Summer oat	20	97	24	22	45	0	77	46	4	8	47	744	1,759	102	27	137	55
Legume mixture	157	373	40	15	14	6	18	48	5	25	26	78	81	1,513	16	52	61
Peas-beans	3	4	4	3	34	2	3	2	33	4	8	12	16	4	2,365	150	89
Lupins	10	70	74	70	157	4	11	52	8	29	37	119	192	137	350	2,120	61
Rec.	74	46	86	82	73	91	81	75	97	84	55	52	59	50	79	71	72

Figure III-9. Confusion matrix derived from the classification results using a combination of optical and SAR features (*ff(S1&S2))*.
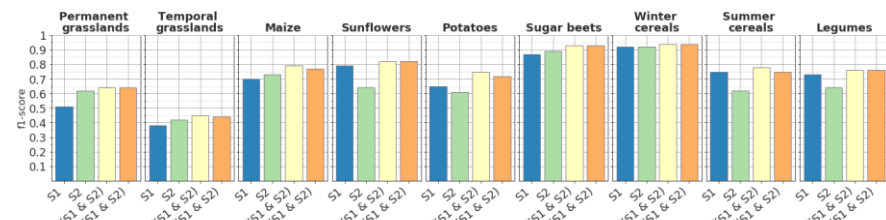
Figure III-10. Grouped crop-specific accuracies derived from the classifications based on only SAR (*S1*) features, only optical (*S2*) features, their combination (*ff(S1&S2)*), and decision fusion *df(S1&S2)*.

4.2. gFFS Rankings and Feature Importance

The RF model with the highest accuracy (f1-score: 0.7) was achieved using a variable-wise gFFS at the 5th sequence with 100 features with *ff(S1&S2)* (Figure 11). By variable-wise gFFS, the full time-series of VH, VV, red-edge (B06), green (B03), and SWIR (B11) were identified as the most performant feature-sets (Figure 12, border axis). The f1-

score difference at the point with the maximum performance (100 features) and the last sequence (320 features) was only 0.01.

The time-wise gFFS showed the maximum performance (f1-score: 0.69) with 112 features at the 7th sequence with **ff(S1&S2)**. The dates chosen as being most crucial by time-wise gFFS are shown in the border axis of Figure 12. No significant difference was observed between the maximum performance value and the f1-score at the last sequence with 320 features ($\Delta = 0.001$).
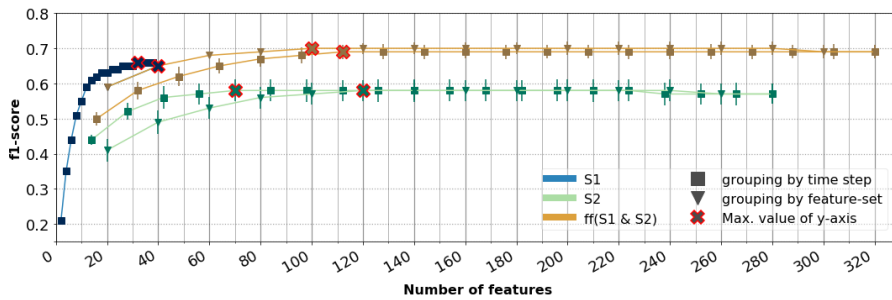
Figure III-11. Feature learning curves of time-wise and variable-wise gFFS based on only optical (**S1**), only, SAR (**S2**), and optical-SAR feature stacking (**ff(S1&S2)**).

When applied to only optical features, time-wise and variable-wise gFFS showed identical maximum accuracy results with f1-score of 0.58. The difference was in the number of features, where time-wise gFFS peaked at the 5th sequence with 70 features, and variable-wise gFFS peaked at the 6th sequence with 120 features.
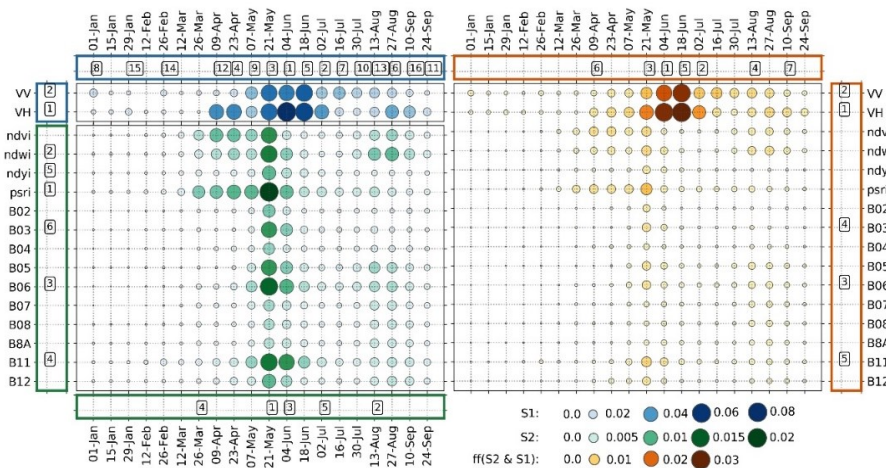
Figure III-12. Feature importance derived from RF models built using only optical, only SAR, and optical-SAR features. Circles illustrate the RF importance scores, while border axes illustrate the number of sequences at which variable-wise and time-wise gFFS have been picked.

The use of both, VH and VV feature-sets (40 features), showed the highest performance (f1-score: 0.65) of the variable-wise gFFS when applied to only SAR features. The maximum performance (f1-score:

74

0.66) of the time-wise gFFS for VV and VH features was achieved at the 16th sequence.

The RF importance scores derived from the RF models trained using optical (greens), SAR (blues), and optical-SAR feature stack (orange-brown) are illustrated in Figure 12. The experiments using only SAR features showed that data from April until mid-September were the most valuable for the classifier. The highest importance scores were given to the features acquired in June when the majority of crops are close to their full development stage. The results of those experiments that were based on only optical features showed that information obtained from three VIs (psri, ndwi, ndvi), red-edge (B06), and SWIR (B11) bands had higher importance scores compared to other features. The features acquired at the end of May had a distinct significance to the classifier. When optical and SAR data were used simultaneously, SAR features received higher ranking compared to optical features. Notably, none of the VIs were selected by variable-wise gFFS when using the stacked features.

4.3. Potential Influences of Parcel Size, Optical Data Availability, and Pixel Location within the Parcel on the Classification Accuracy

For mis- and correctly-classified pixels, we analyzed the information on parcel sizes, the distance of pixels to parcel borders, optical data availability, their NDVI, and VH temporal profiles to assess their influence classification outcomes.

Except for winter triticale and winter rye, all crops with f1-score below 0.70 at *ff(S1&S2)* have average field sizes below 9 ha (see, Table 1). Among winter cereals, winter triticale and winter rye have the smallest average parcel sizes. Figure 13 shows, for each class, the distribution of field sizes of correctly and incorrectly classified pixels. For most classes, the parcel size distribution of correctly classified records is much broader than from misclassified records. Also, differences in medians suggest that pixels coming from large parcels were more often correctly classified. The medians' differences derived from the parcel size distributions of correctly and misclassified records are smaller for the classes with the smallest average parcel sizes. For example, classes such as permanent grasslands (0.29 ha), temporal grasslands (0.33 ha), winter rye (0.85 ha), summer oat (1.03 ha), lupins (1.71 ha) show differences in medians of less than 2 ha. Whereas the highest differences in median parcel sizes were recorded for the following classes: winter wheat (7.44 ha), winter rape (6.15 ha), winter barley (5.80 ha), potatoes (5.14 ha), and peas-beans (5.12 ha).
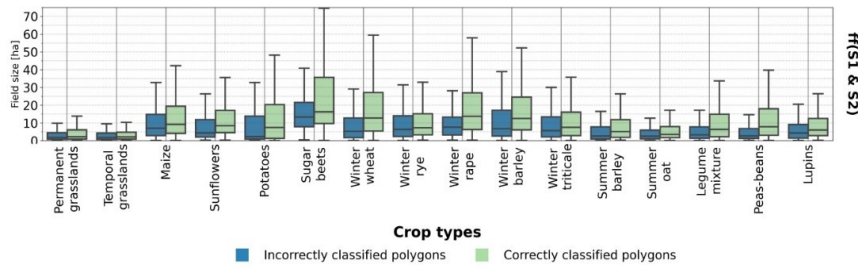
Figure III-13. The parcel size distribution of misclassified and correctly classified pixels for **ff(S1&S2)**. The plots for **S1**, **S2**, and **df(S1&S2)** can be found in Supplementary Materials D.

The variations of f1-scores depending on a pixel's distance to parcel border are illustrated in Figure 14. Plots for all remaining crop types can be found in Supplementary Materials E.

As can be seen from Figure 14, border pixels have lower accuracy than those located 6–8 pixels away from the parcel borders. Predictions based on only optical features seem to have a higher variation of f1-scores even for classes with overall high accuracy values (e.g., winter rape, peas-beans, sunflowers) except for maize and sugar beets classes (see Supplementary Materials E). The classifications based on optical-SAR stacked features and decision fusion, in most of the cases, showed the highest accuracy, even for mixed pixels close to parcel border, compared to classifications based on one sensor only.

Figure III-14. Variations in accuracy, depending on the distance of pixels to the parcel borders (y-axis). Vertical red lines indicate that 80% of the data lie on the left side of this axis.

As various atmospheric conditions can considerably influence optical data quality, the effect of optical data availability on the classification accuracy was assessed based on the invalid pixel mask (see Section 2.3.1.). We analyzed monthly optical data availability for correctly classified and misclassified pixels predicted by **S2** (Figure 15). Further plots showing all crop types are available as Supplementary Materials F.

Correctly predicted pixels have, on average, more valid optical observations compared to misclassified pixels. For example, the class maize had the highest difference in optical data availability for correctly and incorrectly classified pixels in May and August. In May, 94.4% of the correctly predicted pixels had one or more valid optical observations, whereas for misclassified pixels, this number was equal to 65.2%. In August, only 20.1% of all misclassified maize test pixels had one or more valid optical observations, while for the correctly classified pixels, it was 70%. For winter cereals, the maximum differences in the percentages occurred in July for winter barley and winter triticale; in May for winter rape and winter rye, in August for winter wheat. No significant differences were observed for summer cereals, temporal grasslands, and legume classes. In August, the difference for correctly- and misclassified pixels was the highest for classes such as sugar beets (77.1% vs. 37.6%), potatoes (58.9% vs. 38.2%), and sunflowers (41.2% vs. 21.3%). Consequently, such optical data scarcity influenced the temporal profile curves.
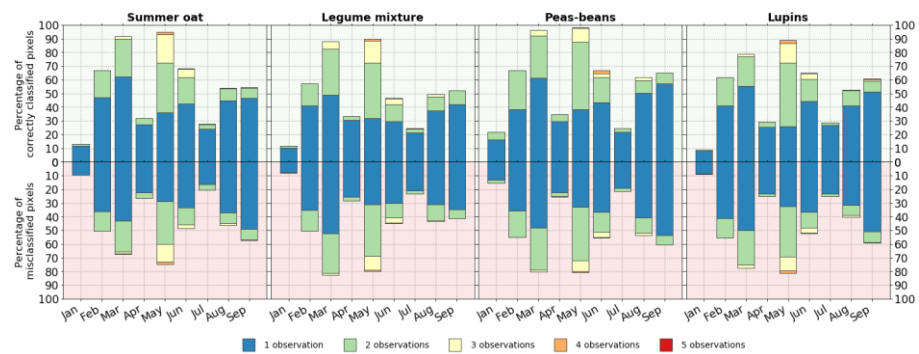
Figure III-15. Incidence of 1–5 monthly observations for correctly- and misclassified pixel.

NDVI temporal profiles of mis- and correctly classified samples deviate considerably in the month, where we also observed large differences in optical data availability reported in the previous paragraph (Figure 15). Optical data scarcity can also be seen from the NDVI temporal profiles of misclassified and correctly classified pixels (Figure 16). Supplementary Materials G gives an overview of NDVI temporal profiles for all crop types. The months when misclassified pixels had the lowest amount of valid optical observations, apparent discrepancies in the NDVI curves can be identified. For example, the lowest percentage of optical data availability for misclassified pixels in August (Figure 15) is reflected in the lower NDVI values in August for the class maize. Surely, climatological, meteorological, and biological factors should also be taken into account when interpreting the alterations seen in these NDVI profiles. Smaller differences were present between the VH temporal profiles of correctly- and

misclassified pixels (Figure 16). These profiles are presented in Supplementary Materials H.
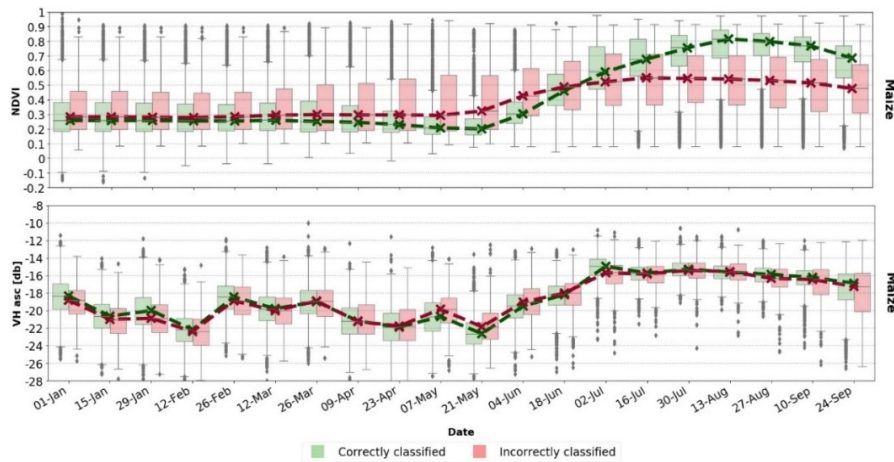
Figure III-16. NDVI and VH temporal profiles of correctly classified and misclassified pixels of class maize.

5. Discussion

We evaluated the classification accuracies of RF models built using dense time-series features of only optical data (Sentinel-2), only SAR data (Sentinel-1), their combination, and decision fusion. Time-wise and variable-wise gFFS was employed to investigate the relevance of specific bands, Vis, and time-steps. To understand further influences on misclassifications, we additionally analyzed parcel sizes, optical data availability, and pixel distance to polygon borders.

The results acquired for the example of Brandenburg (year 2017) showed that the combination of optical and SAR features and decision fusion leads to better overall classification accuracy (f1 score: 0.81) compared to single-sensor approaches. These findings are in line with several previous studies (Inglada et al., 2016; Salehi et al., 2017; Sonobe et al., 2017). In the presented study, no significant differences in accuracy were found between optical-SAR feature-stacking and decision fusion. When considering class-specific accuracies, several crop types were classified more accurately with only SAR features (e.g., lupins, peas-beans), while fewer crop types showed higher accuracies when using only optical features (e.g., permanent grasslands, sugar beets). However, in most cases, crop-specific accuracy based on a combination of sensors was insignificantly higher than the better-performing single sensor accuracy. Respectively, the increase of overall accuracy in the fusion approaches is an expected consequence as each target class could achieve higher accuracies using best performing optical or SAR features. These results suggest that if the remote sensing data availability for the study region is not a subject of concern, the

decision to use optical-SAR fusion or single sensor data should be considered depending on the crop types being investigated. Data fusion is an attractive option for the classification of a broad range of diverse crop classes, but it also comes with more significant computational expenses.

For the majority of investigated classes, classification accuracies derived from only SAR features showed higher values compared to only optical features. These results are contrasting to those where the performance of optical features was higher than SAR (Demarez et al., 2019; Denize et al., 2018). However, several reasons can serve as a possible explanation for these outcomes. First, a large number of crop classes show similar phenological patterns. Existing crop type mapping studies often focus on few and partly merged classes such as maize, grasslands, winter cereals (one class), summer cereals (one class), and sugar beets. These classes have distinct phenological development profiles, which are well reflected in optical data. As our grouped class-specific accuracy showed (Figure 10), the accuracy values based on only optical features increases when both cereals and legumes are grouped into single classes. Nonetheless, when predicting individual cereal classes, the SAR data were more successful in capturing the important short term differences in phenological development phases (see Supplementary Materials H, for VH temporal profiles) for those crops which optical data could not separate (see Supplementary Materials G, for NDVI temporal profiles). We assume that this is most likely due to the data gaps in optical data. It surely depends on the quality and amount of optical and SAR data used for feature generation, as seen in (Van Tricht et al., 2018). Second, in the presented study, we focus purely on the crop type classification, but no other land cover class was considered. The majority of crop type mapping studies often include land cover and land use classes such as urban areas, water bodies, forest, etc. [e.g., 4,17,21]. When differentiating, e.g., maize pixels from water, built-up, or forest pixels, it could be the case that optical features are more relevant for the classifier rather than SAR features.

When looking at Figure 11, we may notice that the feature learning curve of **S2** has a higher starting point than **S1**. However, with the increasing number of grouped features, the accuracy for **S1** increases much stronger than the accuracy of **S2**. It could, thus be concluded that single optical features are more informative than single SAR features; nevertheless, dense time-series of SAR data are able to reach higher accuracies compared to dense time-series of optical features.

No significant accuracy differences were recorded among classifications based on full time-series data (320 features) and their subsets selected using gFFS. This supports that RF algorithm is robust to a large number of input features as stated also by other authors (Belgiu and Drăgu, 2016). However, it is also shown that

dimensionality reduction of input features improves the stability of RF classification accuracies (Millard and Richardson, 2015). Only in the last sequences of gFFS, we recorded an insignificant decline in accuracies (f1-score: Δ = 0.01). But at the same time, the gFFS proved that based on a selection of the most relevant feature subsets (here: 120 features), it was possible to reach the peak accuracy.

Feature selection is used not only to improve the classification accuracy but also to better understand and interpret the input dataset (Immitzer et al., 2019; Sitokonstantinou et al., 2018; Yu et al., 2016), leading to explainable classification outcomes (Roscher et al., 2020). In this context, combining the information of RF feature importance and the rankings of the gFFS allows more profound insights into the relevance of individual features and groups of features for mapping typical crop types of Germany. It enables the identification of features that have high RF feature importance values but are correlated with other features in their information content. For example, many *S2* based features of 21 May show very high RF feature importance values. However, many of these features are not needed to reach maximum accuracy with the gFFS (see Figure 12). Thus, their information content for the classification must be highly correlated with one of the already-selected features. Similarly, the RF feature importance of the psri from 9 April, 23 April, and 7 May are high, but they are not needed to reach maximum accuracy in case of time-wise gFFS on *S2*. Instead, the date 2 July is selected where the Gini feature importance is relatively low for all single features. However, concerning the information already selected in previous sequences, it seems to include some unique additional details. The combination of this information shows a certain limitation of the RF feature importance when the goal is to select a minimum of features with the maximum amount of information for the classification tasks. It also shows that the combination of both feature relevance information sources (Gini feature importance and gFFS ranking) increases the understanding and insight in feature relevance.

Group-wise FFS allowed to substantially reduce the computational costs of a feature selection step while receiving meaningful outcomes. Usage of such feature grouping strategies could be an alternative choice for studies where computational expenses of feature selection were considered as one of the main challenges (Liang et al., 2020; Löw et al., 2013). For studies investigating the spatial-temporal transferability of machine learning models (Meyer et al., 2019, 2018), where the feature selection methods as FFS were extensively used, the application of gFFS might make this process more performant in large input feature-set cases.

The feature learning curve of time-wise gFFS applied on optical-SAR feature combination (Figure 11) shows high accuracy increases in the first five sequences, which correspond to 4 June, 2 July, 21 May, 13 August, and 18 July (Figure 12). These time frames cover critical

phenological phases such as full stem development of most of the cereals, flowering, and land management events like harvest and hay cut. The following two time-steps, 9 April and 10 September improved the f1-score only by 0.01 each. According to the German National Weather Service data (German National Weather Service, 2019), these two time-steps are associated with significant plant height development for winter cereals. As for summer crops, it is the phase of the first plant emergence above ground. The beginning of September is a time of harvest for classes such as maize, sunflowers, and pre-harvest phase for sugar beets. Thus, all time steps selected by time-wise gFFS reflect significant phenological developments or management actions on the ground. The results we acquired from time-wise gFFS and RF feature importance scores enable us to infer that the temporal coverage from the beginning of April until the end of September is sufficient to classify the classes under consideration, including winter cereals in our study region.

The results of variable-wise gFFS (Figure 12) applied on the optical-SAR combination, together with our results on the comparison of single-sensor performances (Section 3.1.), showed the high importance of SAR features. These outcomes were reaffirmed by the highest RF importance scores of VH and VV. This result was different from the findings of some studies, where optical features turned out to be more relevant for crop type classification (Demarez et al., 2019; Denize et al., 2018). In our variable-wise gFFS results, VH and VV were followed by the red-edge band (B06) as the next most relevant feature set. A recent study by Griffiths et al. 2019 (Griffiths et al., 2019) also pointed out that adding the red-edge band to other optical bands improved crop-specific accuracy while it had a lower impact on the accuracy of non-cropland cover classes. The last feature sets that finally raised the feature learning curve to maximum were the green and SWIR bands, which improved the f1-score by only 0.01 each (Figure 11). This supports the finding of previous studies stating that green and SWIR bands also contain important information for crop type mapping tasks (Immitzer et al., 2016).

The crop-specific accuracies were influenced by the parcel sizes (Figure 13). The results showed that classes with small average parcel sizes (Table 1) were classified with lower f1-scores. This issue was similarly discussed in recently published studies (Arias et al., 2020; Defourny et al., 2019a). One of the reasons for this effect is the influence of mixed pixels at parcel borders. Often, small parcels have elongated shapes, which would result in an increased number of mixed pixels. As it was seen from Figure 14, in the majority of cases (all except sugar beets and winter rape), border pixels were classified with lower accuracies almost until the 4th pixel away from parcel border. However, some studies successfully employ information from mixed pixels in their classification tasks (Foody and Mathur, 2006).

Apart from the effect of border pixels, differences in the field management practices (e.g., tillage practices, fertilization, time, and frequency of weeding, water management) for small and large parcels certainly influence the spectral response. Based on this, we explicitly built our sampling strategy in a way that all training and testing pixels were equally distributed among all polygons. We assumed that the differences in classification accuracy for large and small parcels could have been bigger if we had not adjusted our sampling strategy.

It was also shown how optical data availability affected the performance of the classifier. As we saw from Figure 15 (for all crop types, see Supplementary Materials F), predictability of all crop types suffered from a scarcity of optical observations. Nonetheless, some crops appeared to be more sensitive to the lack of optical data at specific months. For example, for the winter rape, it was May, which was associated with the flowering phase. The high RF importance scores (Figure 12) of feature '21 May' in the experiments based on only optical features could also be explained by the optical data availability, as the number of optical observations was much higher in May compared to all other months (Figure 15). Therefore, it is suggested to consider the aspect of data availability when concluding the importance of specific features or time-steps.

Further steps could include testing spectral-temporal variability metrics such as medians, percentiles of optical and SAR features. The spatial transferability of machine learning models would also be considered in future research.

6. Conclusions

The present study investigated the advantages of using the fusion of optical (Sentinel-2) and SAR (Sentinel-1) dense time-series data over the single sensor features. The importance of the features was evaluated using variable- and time-wise grouped forward feature selection (gFFS). Additionally, the effects of optical satellite data gaps and parcel size were analyzed to understand the reasons for misclassifications.

The classification accuracy based on only SAR features outperformed those based on optical features alone. Optical-SAR feature stacking showed the highest accuracies, while no significant difference was found between feature stacking and decision fusion.

The combined assessment of feature ranking based on gFFS and RF feature importance enabled a better interpretation of the results and selecting the most relevant features from both data sources. The question of selecting time-steps with most discriminative power for a classification task is better analyzed by considering all the information available at a particular time step compared to single features.

With optical-SAR feature combination, the peak accuracy of the RF model was achieved when using the full time-series of VH, VV, red-edge (B06), green (B03), and short-wave infrared (B11) bands. As for temporal information, the classifier's performance was the highest when the full feature-sets acquired on 9 April, 21 May, 4 June, 2–18 July, 13 August, and 10 September were used.

The analysis of the parcel sizes showed that these had a high impact on classification accuracies. Crop classes with a large number of small parcels are harder to classify than large parcels. One reason for this is that border pixels had lower classification accuracy than those in the center of the agricultural parcels.

Also, it was shown that for most of the crop classes, the classification accuracy drops when a lower amount of valid optical observations is available at specific months.

$$IV$$

CHAPTER IV: SPATIAL TRANSFERABILITY OF RANDOM FOREST MODELS FOR CROP TYPE CLASSIFICATION USING SENTINEL-1 AND SENTINEL-2

Abstract

Large-scale crop type mapping often requires prediction beyond the environmental settings of the training sites. Shifts in crop phenology, field characteristics, or ecological site conditions in the previously unseen area, may reduce the classification performance of machine learning classifiers that often overfit to the training sites. This study aims to assess the spatial transferability of Random Forest models for crop type classification across Germany. The effects of different input datasets, i.e., only optical, only Synthetic Aperture Radar (SAR), and optical-SAR data combination, and the impact of spatial feature selection were systematically tested to identify the optimal approach that shows the highest accuracy in the transfer region. The spatial feature selection, a feature selection approach combined with spatial cross-validation, should remove features that carry site-specific information in the training data, which in turn can reduce the accuracy of the classification model in previously unseen areas. Seven study sites distributed over Germany were analyzed using reference data for the major 11 crops grown in the year 2018. Sentinel-1 and Sentinel-2 data from October 2017 to October 2018 were used as input. The accuracy estimation was performed using the spatially independent sample sets. The results of the optical-SAR combination outperformed those of single sensors in the training sites (maximum F1-score–0.85), and likewise in the areas not covered by training data (maximum F1-score– 0.79). Random forest models based on only SAR features showed the lowest accuracy losses when transferred to unseen regions (average F1loss–0.04). In contrast to using the entire feature set, spatial feature selection substantially reduces the number of input features while preserving good predictive performance on unseen sites. Altogether, applying spatial feature selection to a combination of optical-SAR features or using SAR-only features is beneficial for large-scale crop type classification where training data is not evenly distributed over the complete study region.

1. Introduction

Supervised machine learning methods are widely used for large-scale crop type classification (Griffiths et al., 2019; Inglada et al., 2017; Preidl et al., 2020). Due to the limited availability of field data (e.g., because of location inaccessibility or reference data collection costs), large-scale crop type mapping often implies model predictions in geographical spaces far beyond the training locations. Due to the presence of spatial autocorrelation in the geo-referenced datasets, the predictor variables in **reference systems** (i.e., the training sites)

86

might significantly differ from those in ***transfer systems*** (i.e., unseen by model transfer sites). Spatially transferring the model outside the 'known' to a model environment could substantially reduce its performance. In the context of crop type mapping, the good spatial transferability of a machine learning classifier indicates its ability to predict crop classes in unseen environments with minimal accuracy losses compared with classification accuracies achieved in training areas.

In recent years, the spatial transferability of machine learning models has been rigorously studied in various geo-spatial application fields (e.g., land cover classification (Lucas et al., 2020), species distribution modelling (Schratz et al., 2018)). Many crop type classification studies have illustrated the successful use of transfer learning and domain adaptation techniques (Ajadi et al., 2021; Gadiraju and Vatsavai, 2020; Nowakowski et al., 2021). For example, Bazzi et al. (2020) (Bazzi et al., 2020) applied 'distil and refine' approach, where the Convolutional Neural Network (CNN) trained with large reference system samples is first distilled into a smaller 'student' model and then refined using the limited target system samples for mapping irrigated areas. Lucas et al. (2021) (Lucas et al., 2021) presented a semi-supervised domain adaptation technique with a novel regularisation method for CNN for mapping a wide variety of crops with the limited number of samples available in the target system. Gilcher and Udelhoven (2021) (Gilcher and Udelhoven, 2021) compared the spatial and temporal transferability of pixel-based and convolution-based classifiers for binary maize and non-maize classification using Synthetic Aperture Radar (SAR) data. Hao et al. (2020) (Hao et al., 2020) researched how the length of time-series of Normalized Difference Vegetation Index (NDVI) features affects the predictive performance of Random Forest models in target systems. Most of such studies were investigating the classifier adaptation techniques to the target domain using semi-supervised or unsupervised learning. In comparison, much less research is available on the influence of input remote-sensing datasets on the classifier's performance in the target systems as performed by, e.g., Hao et al. (2020) (Hao et al., 2020).

Besides the lack of representative samples, overfitting of a classification model to reference samples is a major reason for poor spatial transferability and hence poor generality. Spatial overfitting can occur when machine learning algorithms such as Random Forest are optimized, e.g., for the training data acquired from certain localities (Wenger and Olden, 2012). Recent studies have illustrated that a reduction in spatial overfitting, i.e., fitting the model to samples of one location exclusively, is possible by performing spatial cross-validation (CV) based feature selection (Meyer et al., 2018; Roberts et al., 2017), also known as spatial feature selection. Spatial feature selection allows

detecting and removing problematic predictor variables that carry information about specific training sites but negatively affect the accuracy of predictions in a new geo-location (Meyer et al., 2019). Such approaches to feature selection fall into the 'invariant feature selection' category of domain adaptation techniques (Tuia et al., 2016). While spatial feature selection showed improvements in model transferability in other research fields (Meyer et al., 2018), the effect of spatial feature selection on improving the spatial transferability of crop type mapping has not yet been tested.

The type of remote-sensing datasets used for crop type classification has a substantial effect on crop type accuracies (Orynbaikyzy et al., 2019). Many studies underpin higher classification accuracy based on optical-SAR combinations than single sensor datasets (Inglada et al., 2016; Orynbaikyzy et al., 2020; Van Tricht et al., 2018). Joint use of sensors provides complementary information, such as plant pigment information and canopy structure, and allows improved discrimination of crop types (Veloso et al., 2017). However, to the best of our knowledge, it is unknown if Random Forest models for crop type mapping based on the combination of optical-SAR data show superior results when spatially transferred to the previously unseen environment compared with single-sensor models. Moreover, no comparative studies were found investigating the spatial transferability of models based on only optical and only-SAR datasets. Operational SAR sensors such as Sentinel-1 observe the Earth's surface through clouds at regular intervals over large spaces. Whereas acquisitions from optical data are less regular due to the clouds, which in turn affect the generation of regular time series. It can be hypothesised that SAR-based models with more regular data acquisitions would perform better concerning spatial transferability to distant geographical spaces than optical datasets. This hypothesis is relevant to the areas where the persistent presence of clouds could substantially affect the quality of optical features.

Against this background, this study aims to quantify, reduce, and assess the accuracy losses introduced through the spatial transfer of Random Forest models for crop type mapping in the example of the diverse agricultural landscapes of Germany. First, we test the performance of single SAR or optical data in comparison to a combination of both when predicting crop type classes in a target system, i.e., the transfer region. Second, we attempt to improve the spatial transferability of the widely used machine learning classifier Random Forest using spatial feature selection with a modified feature selection approach—three-step group-wise forward feature selection. Moreover, we analyse auxiliary information such as surface elevation, parcel sizes, soil quality rating, and phenological observation data to understand their possible influences on spatial transferability.

2. Study Sites and Data

2.1. Study Sites

Seven study sites across Germany in the shape of Sentinel-2 tiles (109.8 km × 109.8 km) were chosen based on the reference data availability, the reference data quantity, the distance between study sites, and their regional dissimilarities (Figure 1). The acronyms of the study sites correspond to the second part of the ISO 3166-2 codes of the German federal states where the study sites are mainly located. (Here, BW-Baden-Württemberg, BY–Bavaria, BB–Brandenburg, HE–Hesse, MV–Mecklenburg-Western Pomerania, NI-Lower Saxony, TH–Thuringia). Three sites are located in the Northern German Lowlands (MV, BB, NI), one site in the Central Uplands (TH), two sites in the South German Scarplands (BW, HE), and another one in the Alpine Foreland (BY). The elevation gradually increases from the German Lowlands in the north to the Alps in the country's south. Furthermore, we will use the codes of the German federal states to refer to the specific study site.
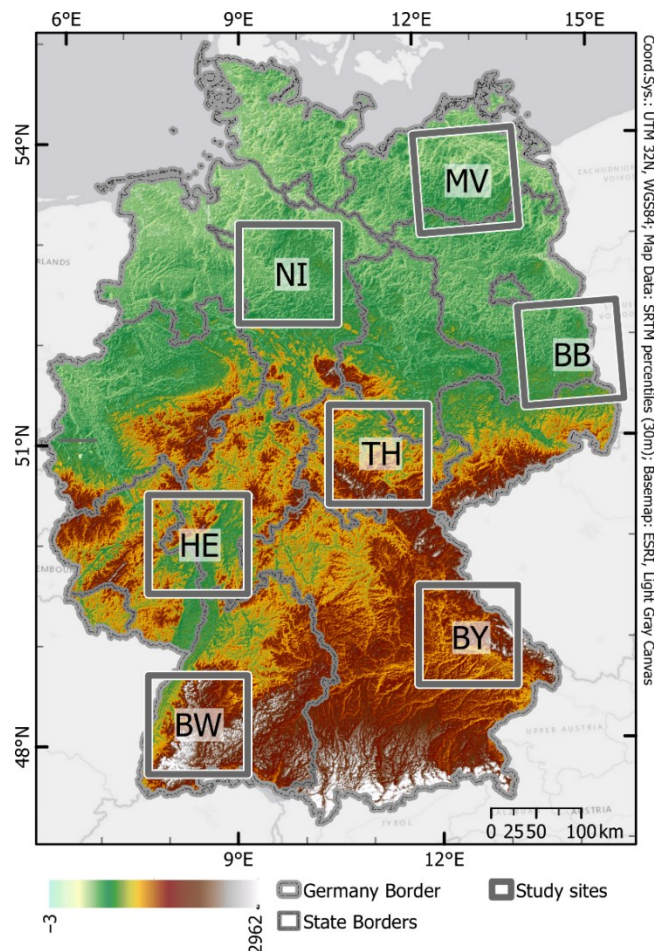
Figure IV-1. Location of the seven study sites in Germany.

According to the present Köeppen-Geiger climate classification (Beck et al., 2018), the western three tiles (NI, HE, and BW) are located in class Cfb, which is characterised by a temperate oceanic climate with warm summers and no dry season. A warm-summer continental climate defines the eastern four tiles (TH, BY, MV, and BB) with no dry seasons (class Dfb). During the summer months of 2018, the lowest and the highest monthly mean air temperatures were recorded in tiles BY and BB (DWD, 2018a). The precipitation pattern varied over the year in all tiles (DWD, 2018b). The outstanding peaks of the monthly total precipitation occurred in tile BW (Figure 2). In general, the year 2018 was recorded as the warmest and sunniest year in Germany since at least 1881 (DWD, 2018c), with the longest heat periods in July and August. This led to substantial negative anomalies in remotely sensed vegetation activity on agricultural land (Reinermann et al., 2019) and substantial yield losses (Klages et al., 2020). However, the spatial patterns of anomalies recorded in 2018 were different across the country. It was a good study case for the assessment of the spatial transferability of Random Forest models under varying environmental and climatic conditions at the country scale.
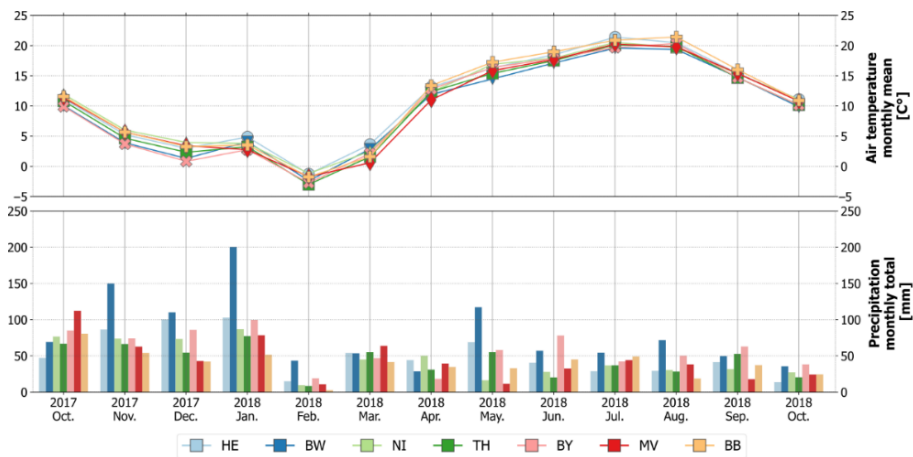
Figure IV-2. Mean monthly air temperature (**top**), monthly total precipitation (**bottom**) across the seven study sites from October 2017 to October 2018. Data source: German Weather Service (DWD).

An agricultural season in Germany typically lasts from March to September for the majority of summer crops and from September to August of the following year for winter cereals. However, due to the differences in natural landscapes and abiotic factors across the country, regional variation of a few days or even weeks can occur in phenological crop growth stages (Gerstmann et al., 2016).

2.2. Reference Data

The reference datasets were acquired from seven German federal states in the form of vector files containing agricultural parcels and crop types

for the year 2018. These datasets rely on farmers' crop declarations, which are part of a subsidy payment scheme within the European Union's Common Agricultural Policy. Agricultural parcels were recorded in the context of the Integrated Administration and Control System (IACS) that is executed by national administrations (in Germany, at the federal state level) and uses the Land Parcel Identification System (LPIS) as a basis. We will refer to the reference datasets as 'LPIS data.'

The declarations by farmers involve manual digitalization of parcel borders. In many cases, such datasets contain geometry overlaps. Parcels overlapping adjacent parcels by more than 500 m² and parcels with a parcel size of less than 0.1 ha were filtered out from the original dataset. These thresholds were selected empirically. The following spectrally inseparable classes were combined: Maize with flowering path, silo maize, and maize for biogas classes were merged into one 'maize' class; Starch potatoes and potatoes for food were merged into one 'potatoes' class; Temporal and permanent grasslands were merged into one 'grasslands' class. The 'summer oat' and 'summer barley' classes were merged into a general 'summer cereals' class. We selected all crop types that were present in all tiles and that had at least 20 parcels after filtering. The threshold of 20 polygons was set to limit the number of pixels sampled from one polygon. The resulting selection of crop types is shown in Table 1.

Table IV-1. Number of parcels per crop type in each study site.

	HE	BW	NI	TH	BY	MV	BB	Sum
Grasslands	2563	124,977	55,511	25,233	95,925	22,766	19,615	346,590
Maize	9837	28,756	26,486	2762	37,947	4165	2383	112,336
Alfalfa	486	942	46	1007	532	103	516	3632
Potatoes	1390	561	6023	159	4551	360	206	13,250
Sunflowers	43	69	20	31	32	83	406	684
Winter wheat	29,547	22,397	13,149	10,495	32,742	5585	1259	115,174
Winter barley	8741	7959	6356	3418	17,511	2513	1065	47,563
Winter rape	9392	5433	5273	5327	5694	3746	860	35,725
Winter triticale	1721	4585	3293	839	5137	386	717	16,678
Winter rye	2497	846	9944	468	2069	1902	4195	21,921
Summer cereals	7017	11,399	10,065	3272	6001	1519	1176	40,449
Sum	73,234	207,924	136,166	53,011	208,141	43,128	32,398	754,002

2.3. Remote Sensing Data and Pre-processing

Optical and SAR data sensed by the Multi-Spectral Instrument (MSI) onboard Sentinel-2 A/B and by the C-band SAR instrument onboard

Sentinel-1 A/B were downloaded from the Copernicus Open Access Hub covering the time frame from 1st of October 2017 to 31st of October 2018. In total, 679 Sentinel-2 scenes and 3709 Sentinel-1 scenes (1898 scenes in ascending and 1811 in descending modes) were processed. Due to its all-weather sensing capabilities, SAR data provides more consistent and valid observations over time. Whereas the availability of valid optical data highly depends on the weather conditions of the sensed locations.

For the pre-processing of optical data, we used the MACCS-ATCOR Joint Algorithm (MAJA) version 3.3 (Hagolle et al., 2017). From the available 12 Sentinel-2 bands, we selected three visible (B2, B3, B4), one near-infrared (B8), four red-edge (B5, B6, B7, and B8A), and two short-wave infrared (B11, B12) bands that were corrected for slope effects (so-called 'FRE products' from MAJA). The red-edge and short-wave infrared bands with a 20 m spatial resolution were resampled to 10 m using the nearest neighbour algorithm. Commonly used vegetation indices (d'Andrimont et al., 2020; Tardy et al., 2017; Van Tricht et al., 2018), namely, Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) and Normalized Difference Yellow Index (NDYI), were calculated from Sentinel-2 bands (Equations (1)–(3)).

$$NDVI = (B8 - B4)/(B8 + B4) \qquad (4)$$

$$NDWI = (B8 - B12)/(B8 + B12) \qquad (5)$$

$$NDYI = (B3 - B2)/(B3 - B2) \qquad (6)$$

We pre-processed Level-1 Ground Range Detected (GRD) and Interferometric Wide Swath (IWS) Sentinel-1 scenes using the S1TBX toolbox (v7.0.4) of the SNAP software. The following pre-processing steps were conducted: (1) applying orbit files; (2) removing GRD border noise; (3) thermal noise removal; (4) subset to the study site area; (5) radiometric calibration; (6) refined Lee speckle filtering (filter window size–5 × 5); (7) terrain flattening; (8) terrain correction; (9) conversion of data from digital numbers to decibels (dB). The output images were resampled with the nearest neighbour algorithm to 10 m spatial resolution using gdal's gdalwarp utility.

We have used pre-processed co-polarized VV and cross-polarized VH bands in ascending and descending data acquisition modes. Additionally, we have calculated the VH/VV ratio for each data acquisition mode.

2.4. Auxiliary Data

To explore potential factors influencing the quality of the spatial transferability, we gathered auxiliary information such as parcel sizes, phenological observation records, surface elevation, and soil quality rating values for each sampled pixel. The parcel sizes were calculated

based on the reference LPIS datasets. Surface elevation information was extracted from a digital elevation model of the Shuttle Radar Topography Mission SRTM (NASA-JPL, 2013). We downloaded the Müncheberger soil quality rating layer from the product centre of the German Federal Institute for Geosciences and Natural Resources (BGR, 2013). The Müncheberger soil quality rating, developed by the Leibniz-Centre for Agricultural Landscape Research (ZALF), comprises information on basic soil and soil hazard indicators (Mueller et al., 2014). For each sample pixel, we extracted scores that ranged from 0 to 100, where a higher score indicates better soil quality for cropping and grazing and higher crop yield potential. We further processed phenological observation records provided by German Weather Service (DWD) via the Climate Data Center (DWD, 2018d) for maize, summer barley, summer oat, winter wheat, and winter rape crops for the season 2017–18.

3. Methodology

3.1. Generation of Dense Time Series Features

The remote-sensing data acquired for seven study sites are located in different orbit tracks, resulting in variation of data acquisition times across sites. In the case of optical data, clouds and cloud shadows reduce the consistency of the time-series. To generate evenly distributed dense time-series features for all study sites, we first generated bi-weekly datetime arrays from the 1 October 2017 to the 31 October 2018. The resulting 29 time-steps were used as the anchor dates to which we interpolated nearest (on time dimension) observation values from optical data (Figure 3). For SAR data, we selected images recorded seven days before and six days after the anchor date and calculated the median of these images at the pixel level. Generation of dense time-series features was performed for all optical and SAR variables described in Section 2.3.
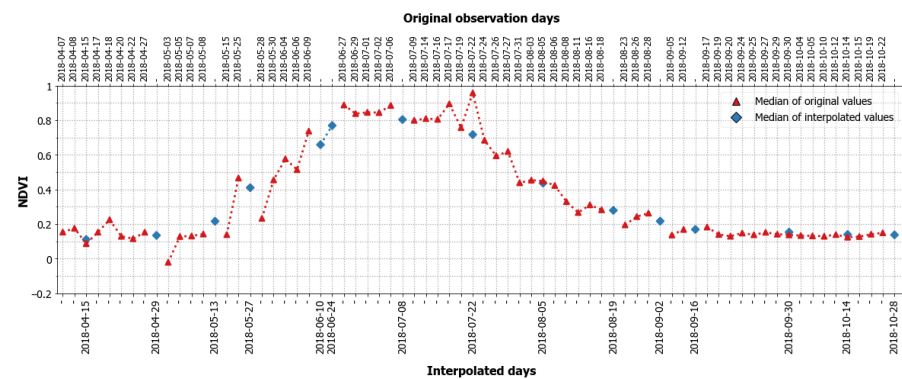
Figure IV-3. NDVI profile of original and interpolated values for the class potatoes (500 sample points) in tile HE.

3.2. Training and Testing Samples

The study was performed at the pixel-level to avoid the introduction of biases due to the segmentation quality across seven study sites (Tetteh et al., 2021). For each of the seven study sites, we sampled 500 pixels per crop type using stratified random sampling. From the resulting sample set, 60 percent (300 pixels) was used as a training-set and 40 percent (200 pixels) as a test-set for the classification model. It was ensured that no overlaps occurred between training and test samples at the parcel level. To avoid the underrepresentation of samples from small parcels, we adjusted our sampling scheme to consider the parcel size information by distributing samples more evenly among parcels of different sizes. A negative buffer of 10 m (one Sentinel-2 pixel) was applied to exclude the border pixels from sampling. For each sample, we kept information about the size of the parcel from which it was sampled.

3.3. Model Performance Estimation Using Spatial Cross-Validation

To evaluate the models' performance on a spatially independent test-set, we ran a 7-fold spatial CV where sample data from one study site was considered as one-fold (see Figure 4, 'Model Validation' part). In literature, spatial CV is also called leave-location-out CV (Meyer et al., 2018) and block CV (Roberts et al., 2017).

In each run of 7-fold spatial CV, the entire test data from one-fold is held out as an independent test-set representing the target system. The remaining six folds are used as a training site representing the reference system. After building the Random Forest model using training samples from the reference system, we spatially transfer the model to predict the test-set in the target system. Since we also want to evaluate the model's performance in the training sites, we additionally predict the crop types for the test-set samples in the reference system. This procedure runs seven times; each time, the hold-out fold changes so that each fold is once the spatially independent test-set (target system) and six (k – 1) times the training-set (reference system).
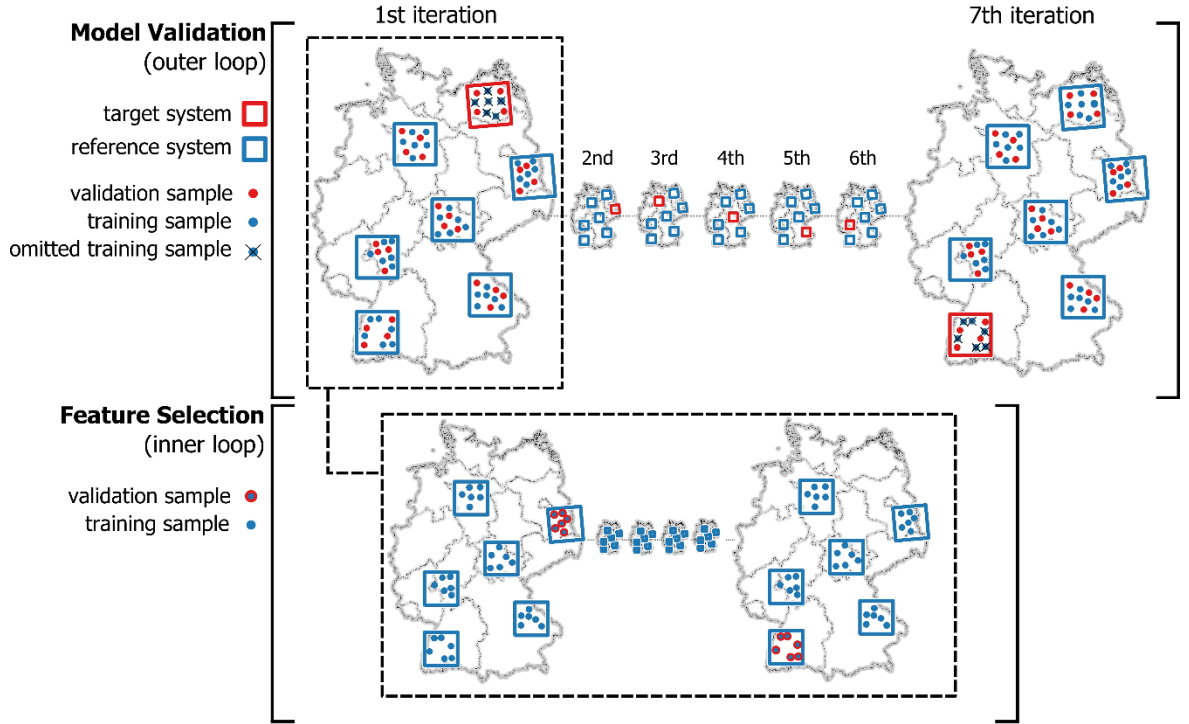
Figure IV-4. Graphical illustration of model validation and feature selection procedures. The pseudo-code of this workflow is provided in Supplementary Material A.

We call the accuracy scores received from the reference system test-set samples 'reference system accuracy'. The accuracy scores received from the target system test-set samples we call 'target system accuracy'. The average and class-specific F1-scores (F1) were used as accuracy measures (Equation (4)). To better assess the quality of the model transfer to the target systems, we calculated the 'accuracy loss' ($F1_{loss}$) by subtracting the F1-score acquired from the reference system from the F1-score acquired from the target systems (Equation (5)).

$$F1 = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \frac{1}{2}(\text{False Positive (FP)} + \text{False Negative (FN)})} \tag{7}$$

$$F1_{loss} = F1_{\text{target system}} - F1_{\text{reference system}} \tag{8}$$

3.4. Feature Selection and Model Building

In reference systems, Random Forest models were built with the following two strategies: first by using all available features ('All features') and second by applying spatial feature selection ('gFFS+sCV'). In total, we ran six experiments (see Table 2). We applied two feature selection approaches for each of only optical ('S2'), only SAR ('S1') and the combination of optical-SAR ('S1+S2') datasets. The combination of optical and SAR features was performed by

stacking features together. For each of conducted six experiments, we calculated reference and target system accuracy and classification accuracy losses by subtracting target system accuracy from reference system accuracy.

Table IV-2. Overview of conducted six model building approaches with three input datasets and two feature selection approaches. For each experiment, the classification accuracies from the reference system ('Ref.System') and target system ('Trg.System') were recorded.

		Input Dataset		
		S1 [1]	S2 [2]	S1+S2 [3]
Feature Selection Method	All features [4]	Ref. System [6] and Trg. System [7]	Ref. System [6] and Trg. System [7]	Ref. System [6] and Trg. System [7]
	gFFS+sCV [5]	Ref. System [6] and Trg. System [7]	Ref. System [6] and Trg. System [7]	Ref. System [6] and Trg. System [7]

1 **'S1'**—SAR data from Sentinel—1 satellite; 2 **'S2'**—optical data from Sentinel—2 satellite; 3 **'S1+S2'**—combination of S1 and S2 by feature stacking; 4 **'All features'**— all features of input datasets were used to build the final Random Forest model; 5 **'gFFS+sCV'**—a subset of features selected using three-step group-wise Forward Feature Selection (gFFS) with spatial cross-validation (sCV) was used to build the final Random Forest model; 6 **'Ref.System'**—reference system is the system from which the training samples were used to build a model; 7 **'Trg.System'**—target system is the system from which no training samples were used. The target systems are only used to evaluate the classification performance of the models.

In 'All features', we selected all input features and all training samples from the six reference system folds to build final Random Forest models. The spatial feature selection was performed using a 3-step group-wise Forward Feature Selection (gFFS, described below). In 'gFFS+sCV', all training samples of the reference systems were split into six folds (see Figure 4, bottom) based on their spatial allocation (one study site = one spatial fold). The feature selection is then performed using gFFS and 6-fold spatial CV.

The final Random Forest model is then built using all reference system training samples and the selected feature subset. The Random Forest algorithm (Breiman, 2001b) was selected based on numerous reports of its successful application in crop type classification tasks (Forkuor et al., 2014; Zhong et al., 2014), its ability to handle high dimensional feature spaces (Belgiu and Drăgu, 2016) and relatively low sensitivity to hyperparameter tuning (Schratz et al., 2018). The standard setting of the scikit-learn (version 0.22) implementation (Pedregosa et al., 2011) of the Random Forest algorithm was used to build the final prediction models, with the only change in the number of trees from 100 to 500 as recommended in (Belgiu and Drăgu, 2016). This hyperparameter setting is also commonly used among large-scale crop type mapping studies (Blickensdörfer et al., 2022; Preidl et al., 2020). The square root of the total number of features was used to split the nodes in the single trees of this ensemble classifier.

Group-wise Forward Feature Selection (FFS) is a variation of standard Forward Feature Selection (FFS) (Saeys et al., 2007) that begins by evaluating all single features individually. Here, by 'evaluation', we mean model building, predicting, and measurement of a performance score. After the first iteration, the input feature of the model with the highest performance score is selected as a fixed feature and passed to the next sequence. The procedure is reiterated by evaluating the set of fixed features from previous iterations together with one new feature from the remaining unselected features. The best-performing feature pair is fixed for the next iteration and subsequently again combined with each of the remaining features individually. The process runs until, e.g., no unselected features are left, the number of desired features is reached, or other custom stopping criteria are met.

One of the main limitations of FFS is its computational intensity. Intending to reduce the computational costs and still investigate all available features, we used group-wise FFS (gFFS) as presented in Orynbaikyzy et al. (2020), based on Defourny et al. (2019c). In gFFS, instead of considering single features within a given FFS iteration, groups of features were used.

The gFFS was conducted in three sequential steps (Figure 5). First, we run variable-wise gFFS, where features are grouped based on variables (e.g., complete time-series of S2 bands, vegetation indices, two S1 bands and their ratio). Each group of variables was considered as a single entity within gFFS. The feature groups selected by the variable-wise gFFS step then go to the second step—the time-wise gFFS. Here, the features are grouped based on time-steps (e.g., all features selected in the 1st step are from the 7th of June) and each group of time-steps was considered as a single entity within gFFS. The resulting selection of variable and time feature groups is then passed to the final third step, the standard FFS, where only single features are considered.
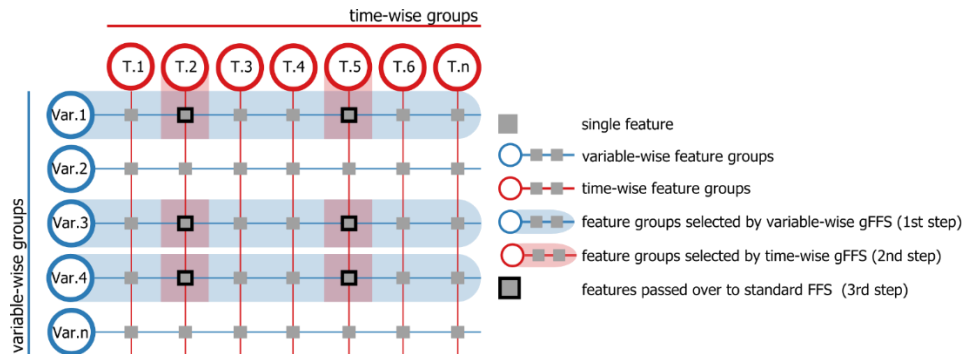
Figure IV-5. 3-step group-wise Forward Feature Selection (gFFS) approach.

In the variable-wise gFFS, the number of groups varies between the two sensors. SAR data has six (VVasc, VVdsc, VHasc, VHdsc, VV/VHasc, VV/VHdsc), optical data has 13 (B2, B3, B4, B5, B6, B7, B8,

B8A, B11, B12, NDVI, NDWI, NDYI), and consequently, the optical-SAR combination has 19 groups. In the time-wise gFFS, the number of groups is the same for both sensors and for their combination (29 time steps).

Feature selection was performed using the Random Forest (Breiman, 2001b) model with standard settings of scikit-learn implementation (Pedregosa et al., 2011) of the algorithm but with 500 trees. Each of three gFFS procedures was stopped if adding a new feature or feature group did not increase the F1-score five times in a row. Open source Python packages such as MLxtend (version 0.17.2) (Raschka, 2018), eo-box (version 0.3.10) (Mack, 2020), pandas (version 1.0.3) (McKinney, 2010) and NumPy (version 1.18.1) (Oliphant, 2006) were used in the implementation of the presented 3-step gFFS.

4. Results

4.1. Overall Classification Accuracies

4.1.1. Accuracies without Spatial Transfer (Reference Systems)

Without considering a spatial transfer, the accuracies based on all features exceed those based on features selected using spatial gFFS (Figure 6, 'Ref.System'). This pattern is common for all three feature sets. The highest median F1-score of 0.85 was reached with all features from optical and SAR sensors. The classification accuracies based on spatial gFFS showed lower median accuracies and higher accuracy ranges among reference systems, compared with the results using 'All features'. On average, F1-scores were 0.02 lower, and the F1-score range (here, maximum-minimum) was reaching 0.03 of a difference.
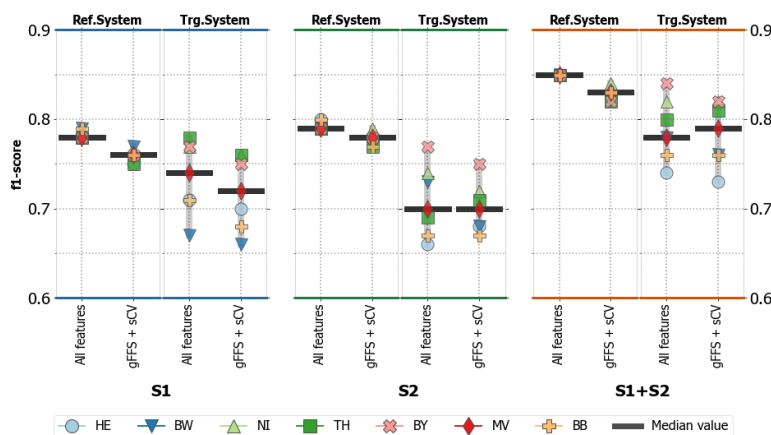
Figure IV-6. Overall F1-scores of classifications based on three feature sets, with variants of using all features and spatial feature selection. The abbreviation keys are provided in Table 2.

The classification accuracies based on S1 features were marginally lower than those based on only S2 features when not considering spatial transfer; the difference of corresponding experiments never exceeded 0.05 F1-score.

4.1.2. Accuracies for Spatially Transferred Models (Target Systems)

When spatially transferring the models (Figure 6, 'Trg.System'), the highest median F1-score of 0.79 was reached with the combination of optical and SAR features selected using 'gFFS+sCV'. But the median F1-score differences between no feature selection and spatial feature selection approaches remained below 3% for all sensor groups. The ranges of F1-scores in target systems are, on average, ten times higher than in reference systems. No distinct pattern was found indicating one approach's superiority or inferiority in target systems.

Figure 7 shows that the highest average accuracy loss across all six experiments was observed in target tiles HE, BB, and BW (F1-score loss−−0.09). The lowest average F1-score reduction of 0.01 occurred in target tile BY.

The spatial transfer experiments based on S1 features showed higher accuracy values than those based on only S2 features (Figure 6). The lower accuracy losses in models based on S1 features and S1+S2 were received compared to those based on S2 (Figure 7). The median F1-score loss with S1 features is −0.04. For S2 features, it equals −0.08, and for the combination of two sensors, it is −0.06. In the experiments with S1+S2, performing spatial feature selection helped to reduce the average accuracy loss to 0.04 across the seven target systems.
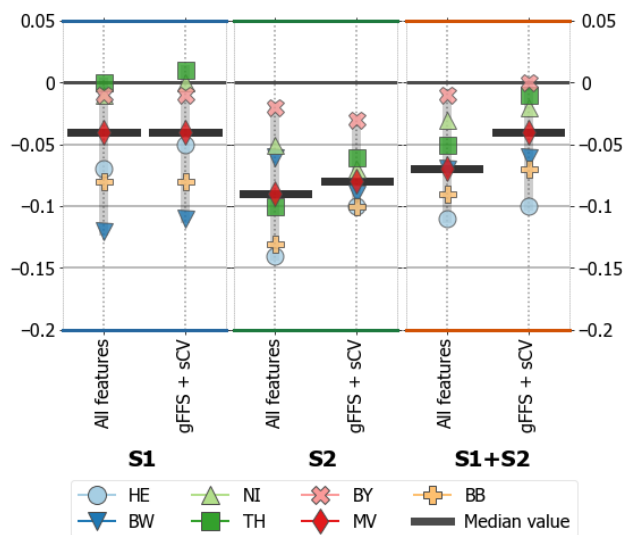
Figure IV-7. Classification accuracy (F1-score) losses in target systems compared to reference system accuracies for all six experiments based on three sensor inputs and two feature selection approaches. The abbreviation keys are provided in Table 2.

4.2. Class-Specific Classification Accuracies

4.2.1. Accuracies without Spatial Transfer (Reference Systems)

Without considering a spatial model transfer, class-specific accuracies were the highest when using optical and SAR features in combination (see supplementary material B). Except for summer cereals and winter barley, the highest median accuracies within the reference systems were reached when all features were used to build the models. In addition to Figure 8, we provide a table with median class-specific F1-scores for reference and target systems in supplementary material C.

The highest average range of F1-score values across the seven reference systems was observed for the class sunflowers (mean variation–0.06) and the lowest for the class winter rape (mean variation–0.01). Classification accuracies were higher for grasslands, maize, alfalfa, and summer cereals with only S2 features than with S1 or S1+S2. For winter cereals, the difference in median F1-scores between runs based on S2 or S1 features did not exceed 0.05 for corresponding feature selection approaches. For a detailed plot with all six experiments, see supplementary material B.
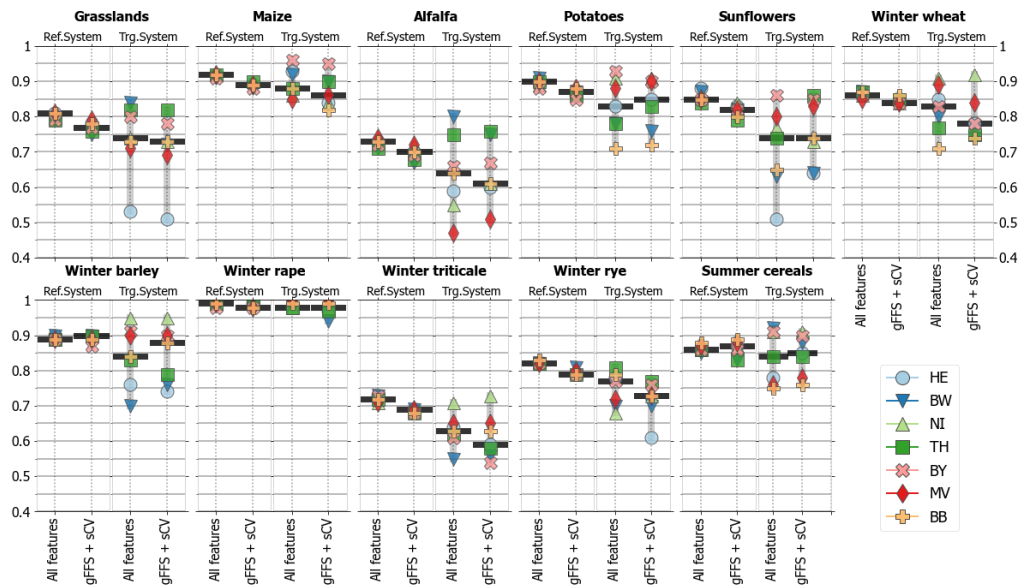
Figure IV-8. Crop-specific classification accuracies (F1-score) in reference and target systems based on the optical-SAR combination and two feature selection approaches. The abbreviation keys are provided in Table 2. For a complete plot with three sensor groups, check supplementary material B.

4.2.2. Accuracies for Spatially Transferred Models (Target Systems)

As in the reference systems, the average accuracy values in the target systems were the highest when we combined optical and SAR features (see Supplementary Material C). The potato and winter rape classes have shown equally high median accuracies with only S1 features as

with features combinations (S1+S2). In target systems, the average accuracy dropped for all classes except for winter rape when compared with the reference systems (Figure 9). For alfalfa, sunflowers, and winter triticale, the accuracy losses were most significant. The maximum F1-score losses for these three classes reached the following: −0.27 for alfalfa (target tile MV); −0.38 for sunflowers (target tile HE); −0.28 for winter triticale (target tile BW). Moreover, increased confusion among winter cereals and between alfalfa and grasslands was observed with 'gFFS+sCV'.
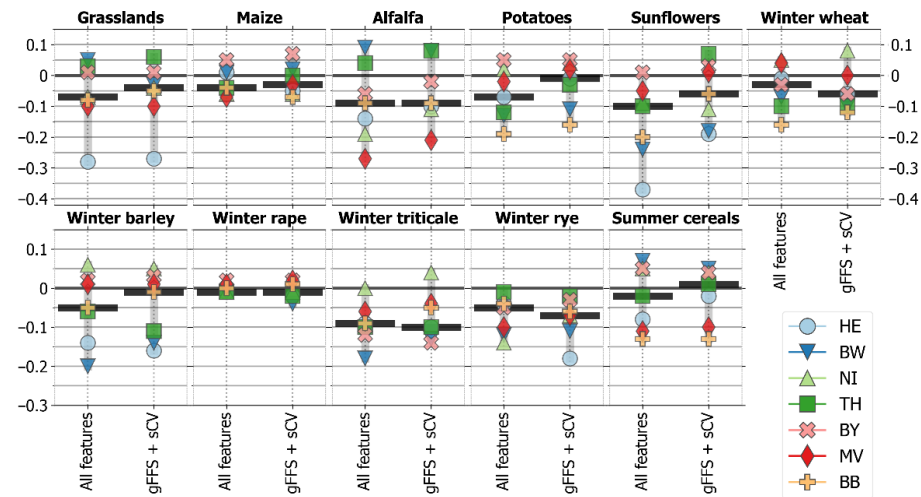
Figure IV-9. Crop-specific accuracy losses in the target systems for models using a combination of optical and SAR features (S1+S2). The abbreviation keys are provided in Table 2. For a complete plot with three sensor groups, check Supplementary Material D.

The accuracy range (maximum-minimum) between the target systems was, on average, six times higher than that between the reference systems for the corresponding experiment sets (Figure 8). Among crop classes, the highest F1-score range across the seven target systems was observed for alfalfa (0.30), sunflowers (0.28), grasslands (0.27), and potatoes (0.21). For grasslands, the high variation resulted from tile HE, which showed a substantial accuracy loss when it is set as a target tile. Almost half of the grassland samples (84 samples) in target tile HE were misclassified as alfalfa, resulting in a very low F1-score.

The classes with high accuracy variations in the target systems also showed significant alteration of NDVI temporal profiles across the seven study sites. For example, alfalfa fields are harvested several times during the growing period, with varying harvest event patterns across the country. This results in various reflectance and backscatter patterns in the time-series and increases the within-class variance, which complicates the identification of the alfalfa fields (Figure 10). The NDVI temporal profiles for all considered crop classes and tiles are provided in Supplementary Material E.
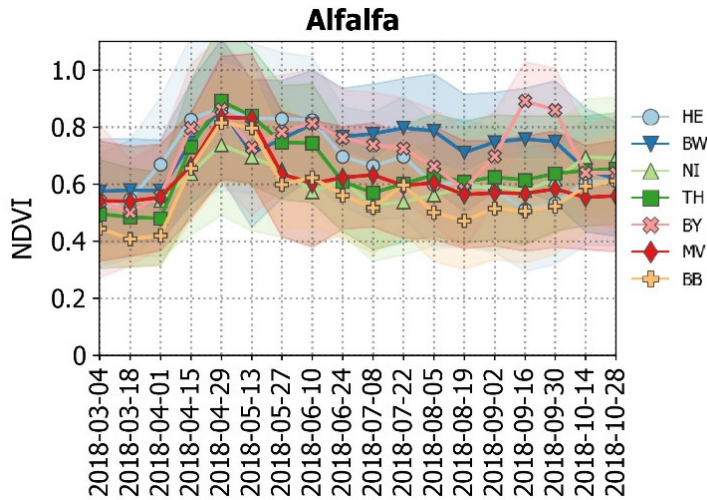
Figure IV-10. NDVI temporal profile of class alfalfa across seven study sites.

Except for sunflowers and potatoes, no clear pattern was observed indicating the superiority or inferiority of a particular model building approach. For sunflowers and potatoes, the use of spatial feature selection on optical-SAR features reduced the median accuracy losses in target systems to 0.01 for potatoes (from an F1-score of 0.07 with 'All features') and 0.06 for sunflowers (from an F1-score of 0.10 with 'All features').

The models built using only SAR features showed the lowest accuracy losses in target systems for the following seven classes: grasslands, alfalfa, sunflowers, winter wheat, winter barley, winter rape, and winter triticale (Figure 9). The remaining four classes (maize, potatoes, winter rye, and summer cereals) showed the lowest accuracy losses with the combination of optical and SAR features.

4.3. Features Selected with Spatial gFFS

For runs with only S2 features and S1+S2, the average number of selected single features was lower than with S1 features (Table 3). The dissimilarities were present not only in the number of selected single features or groups but also in the repeatedly selected variables (Figure 11). Among optical variables, NDVI, NDYI, B6, and B11 were selected the most in the S2 and S1+S2 runs. All six SAR variables were selected more than four times in the S1 and S1+S2 runs.

The temporal groups covering the period from mid-April (15 April 2018) to the beginning of August (5 August 2018) were selected the most in all three sensor combinations. This period covers the most critical agro-phenological phases (e.g., plant emergence, plant height development, flowering) and land management activities (e.g., hay cut, harvest) across all study sites. The temporal groups in the two autumn seasons (autumn 2017, autumn 2018) were rarely selected with spatial CV.

102

Table IV-3. The average number of selected single features or feature groups using 3-step group-wise Forward Feature Selection (gFFS) and the corresponding total number of model evaluation runs.

Sensors	Feature Selection Approach	Avg. Number of Selected Variable Groups	Avg. Number of Selected Time Groups	Avg. Number of Selected Single Variables	Total Number of Model Evaluation Runs	Number of Needed Runs If Standard FFS Applied
S1	All features	-	-	174	-	-
	gFFS+sCV	6	17	53	4474	7965
S2	All features	-	-	377	-	-
	gFFS+sCV	7	13	31	2807	11,568
S1+S2	All features	-	-	551	-	-
	gFFS+sCV	11	14	46	6649	24,816
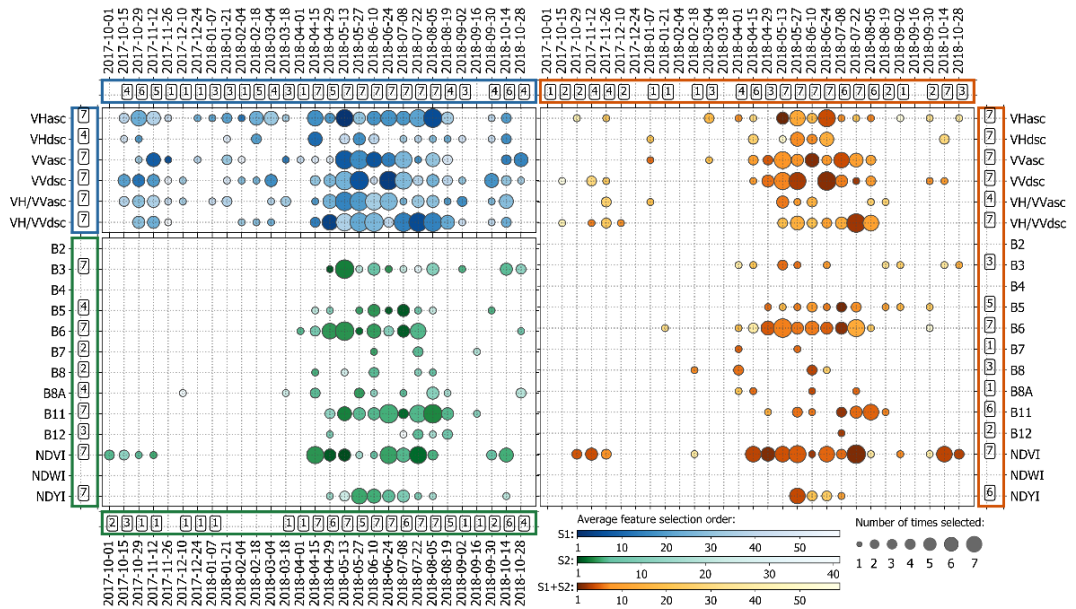
Figure IV-11. Analysis of feature selection results with spatial feature selection. The abbreviation keys are provided in Table 2. **Outer boxes:** The number of times a variable group (*y*-axis) or a time group (*x*-axis) was selected by spatial gFFS on runs with only SAR (blue borders), only optical (green border), and optical-SAR feature combinations (orange borders). **Inner boxes:** The single features, selected in the last step of spatial gFFS. The circle sizes represent the number of times a feature was selected, and the colour intensity represents the median order (sequence) at which it was selected.

4.4. (Potential) Influences of Environmental Settings

As illustrated in Figure 12a, the sizes of the parcels in the study sites (MV, BB, and TH) located in Eastern Germany are bigger than those (BW, HE, NI, and BY) located in Western Germany. The smallest average parcel sizes were recorded in the southwestern two tiles—HE (1.2 ha) and BW (1.1 ha). These tiles showed high accuracy losses when a classification model built on all other regions was transferred to them. The parcel sizes also substantially vary depending on crop types

(see Supplementary Material F). The crops with smaller average parcel sizes, such as grasslands and alfalfa, showed higher accuracy ranges when a model was spatially transferred to the unseen area than those crops grown on larger parcels.
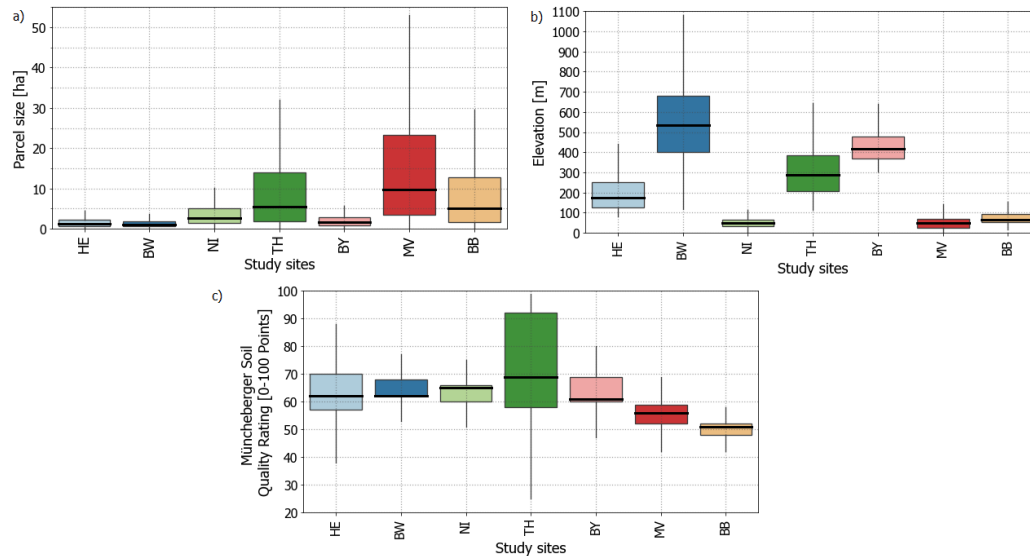
Figure IV-12. Distribution of (**a**) parcel sizes, (**b**) surface elevation, and (**c**) Müncheberger soil quality values among the seven study sites. Each boxplot contains all sampled 5500 points.

Moreover, a considerable difference in surface elevation values is present across the study sites (Figure 12b). The highest average elevation was recorded for the tile BW, and the lowest elevation values were observed for three tiles located in the Northern German Lowlands (NI, MV, and BB). We observed considerable accuracy losses when Random Forest models were transferred to the tile BW.

The average values of the Müncheberger soil quality rating range between 60 and 70 points (Figure 12c), except for two northern tiles (MV, BB). The data from those two tiles showed the lowest average values, which indicate lower soil suitability for cropping purposes and potentially reduced plant vitality or biomass development. High accuracy losses were recorded when trained models were spatially transferred to those two northern tiles (Figure 7).

The temporal shifts in the timing of phenophases or field management activities (e.g., harvest) across the seven study sites can be observed from the phenological observation data acquired from DWD (Figure 13). For example, the average harvesting time for maize in tile BB happened approximately 3.5 days (minimum difference with tile MV) and 23 days (maximum difference with tile BW) earlier (Figure 13). When the Random Forest model was spatially transferred to tile BB, we obtained high accuracy losses for the maize class (see Figure 9). The same is true for summer barley, which is part of the summer cereals class (Figure 13). For winter wheat, notable temporal

shifts (variation of median values: 27 days) were present in crop sowing events; For winter barley, higher dissimilarities were present in harvest time (17 days) than in the average timing of sowing events (8 days). The accuracy losses were higher for winter barley, with more temporal dissimilarities in harvest occurrence than for winter wheat (Figure 9). However, accuracy losses in the winter rape class were minimal despite similar differences in average harvest (14 days) and sowing (10 days) days across seven regions. More information is available in Supplementary Material G.
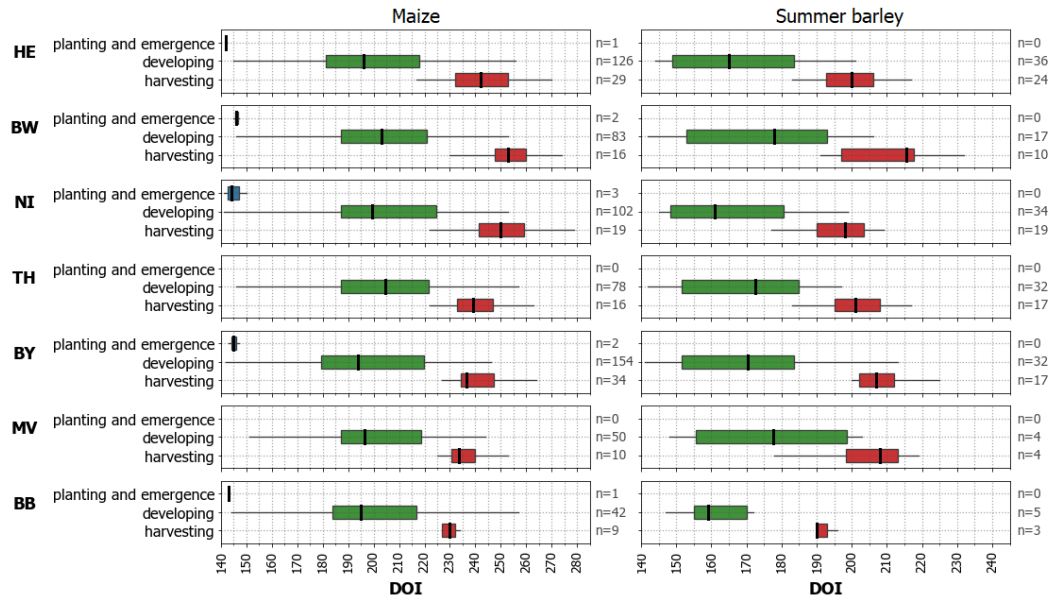
Figure IV-13. Phenological phase observations for maize and summer barley (part of summer cereals class) located within the seven study sites (data source: DWD, 2018d). The following grouping of recorded phenological phases was applied: 'Planting and emergence' include 'beginning of tilling sowing drilling' and 'beginning of emergence' for maize; 'Developing' includes 'beginning of flowering', 'beginning of mil ripeness', 'beginning of wax-ripe stage', 'yellow ripeness', 'tip of tassel visible', 'beginning of growth in height' for maize and 'beginning of heading', 'yellow ripeness', 'beginning of shooting' for summer barley; 'Harvesting' includes 'harvest' for both classes. No observation was found for the 'Planting and emergence' stage for summer barley.

5. Discussion

Due to the environmental, climatic, and phenological differences across the study sites, classification accuracy losses in target systems are inevitable. While such regional differences and the lack of representative training samples are the main drivers of the reduced performance of Random Forest models in target sites, the quality and relevance of input remote-sensing features are other important aspects affecting the spatial transferability of the models. Our study demonstrated that the optical-SAR combination outperforms the classification results based on single sensors in both reference and target systems (Figure 6). The superior performance of the optical-SAR

combination for crop type classification in model training sites is well known (Forkuor et al., 2014; Orynbaikyzy et al., 2019; Van Tricht et al., 2018). Our finding adds that the optical-SAR combination outperforms the single sensor datasets also in geographic spaces unseen by the model. A combination of optical and SAR features should be preferred when performing large-scale crop type mapping with spatially limited training data.

The classification accuracies of the models based on only optical features were marginally better than those based on only SAR features in the training sites (Figure 6). This goes in line with the available comparative literature on the application of optical and SAR data for crop type mapping (Orynbaikyzy et al., 2019). However, in the target systems, it flipped to the opposite: models based on SAR features showed better accuracies than those based on optical. This resulted in lower accuracy losses with only SAR data compared to only optical, or a combination of both. This is a new finding that is relevant for real-world crop mapping scenarios where training data often has limited spatial coverage. It might be more important to select an approach or dataset that is more robust regarding spatial transferability than the best in the reference system.

The presented results show that the models built with SAR features are more robust (i.e., have lower accuracy losses in transfer systems) than optical or optical-SAR combinations. The majority of the investigations to date have reported SAR data's suitability for crop type mapping in the training sites (Bargiel, 2017). The results of our comparative study support the findings of (Gilcher and Udelhoven, 2021) that SAR data is well suited for building spatially transferable crop type classification models and add that, in similar climatic conditions as in Germany, SAR-based models are more spatially transferable than those based on optical data. Recent studies from Woźniak et al. (2022) (Woźniak et al., 2022) and d'Andrimont et al., 2021 (d'Andrimont et al., 2021) illustrated that detailed mapping of crops at the country and continental scale with good accuracy is possible using only SAR data.

A reason for the lower accuracy losses of models based on SAR data could be the availability of consistent valid observations due to its all-weather sensing capabilities, which is crucial for successfully classifying various crops. Whereas the availability of valid optical data highly depends on the weather conditions of the sensed locations. For example, Ghassemi et al. (2022) (Ghassemi et al., 2022) reported that generating monthly composites for the entire Europe was not possible due to the persistent cloud presence in some regions. Knowing that data from both sensors can successfully replicate the agro-phenological development phases of crops (Veloso et al., 2017), it is reasonable that SAR-based models with more usable observations across large areas show better performance for spatial transfer than those based on

optical data only. However, for areas with no persistent cloud cover issues, such spatial transferability study outcomes could be the opposite. Also, the capacity for spatial transferability of models based on only optical features could be different when other data compositing approaches are applied, as proposed by Preidl et al. (2020), or a combination of two or more optical sensors is used, as shown by Griffiths et al. (2019).

Contrary to the findings of (Meyer et al., 2019) on small scale land use and land cover classification, performing spatial feature selection had no substantial effect on the spatial transferability of Random Forest models for crop type mapping. Nonetheless, spatial feature selection helped to eliminate irrelevant for the classifier features and to build much simpler models that are, based on the classification accuracies, comparable or even better (in the case of optical-SAR combination) than those based on all features. The models built using only eight percent from all single optical-SAR features (gFFS+sCV, Table 3) showed marginally improved absolute accuracies (Figure 6) and reduced accuracy losses (Figure 7) in target systems. Reduction of accuracy losses in target systems was also recorded for only optical features when spatial feature selection is applied (Figure 7). The models built with fewer predictor variables showed better spatial transferability as have already been reported by (Ferraciolli et al., 2019; Wenger and Olden, 2012). This underpins the relevance of spatial feature selection, especially for large-scale crop type mapping studies where an increased number of predictor variables decreases computational feasibility and requires substantial storage capacity.

As anticipated, accuracy losses vary among crop classes. Crops that are harvested several times during the growing period (e.g., grasslands and alfalfa), classes with a small number of parcels (e.g., sunflowers and potatoes), and classes with high genetic variability (e.g., wheat and rye) showed high accuracy losses in target systems. The survey of German farmers (Macholdt and Honermeier, 2017) indicated that the choice of cultivars is mainly driven by environmental variables such as soil quality. Consequently, this results in the spatially clustered representation of cultivars across the region, which could negatively affect a model's ability to correctly predict the unseen cultivar. However, spatially transferring models for mapping of, e.g., maize and winter rape across large areas was possible with low accuracy losses. This supports the recent findings of Gilcher and Udelhoven (2021) where acceptable spatial generalisation was possible for binary maize vs. non-maize classification with CNN. In future research, specific features designed to express generalised patterns, such as cutting event indicators that are independent of a specific moment in time, or various texture features based on optical and SAR data, should be considered for testing their usefulness for improving spatial transferability.

The spatial feature selection emphasised the importance of all SAR features along with NDVI, NDYI, B6 and B11. The relevance of NDVI, red-edge (B6) and short-wave infrared (B11) information for crop type mapping has already been reported in earlier studies (Griffiths et al., 2019; Orynbaikyzy et al., 2020). In this study, NDYI from the May-June period, which corresponds to the rapeseed flowering phase in Germany, showed high importance. For mapping rapeseed crops with high accuracy, it is highly suggested to consider NDYI which has been also successfully applied for mapping rapeseed flowering events in Germany (d'Andrimont et al., 2020). Contrary to the finding (Sun et al., 2003) where NDWI was among the top six important features, no NDWI features were selected in our study by spatial feature selection, indicating their irrelevance. As for SAR features, based on the outcomes of spatial feature selection (see Figure 11), we advise using a VH/VV ratio. The advantages of using the polarization ratio were reported by Veloso et al. (2017) for separating maize and sunflowers during the flowering phase and by Inglada et al. (2016) for early crop type mapping.

The results underpin that mapping dynamic land-use classes such as croplands at a larger scale without well-distributed training data is challenging and complex. Many abiotic and anthropogenic factors influence the development of the crops throughout the growing period (Bajocco et al., 2021). Expectedly, those influencing factors enormously vary across geographic regions within Germany.

The variations in phenological crop development stages across the seven study sites (Figure 12) seem to be among the main drivers of the reduced spatial transferability of the tested Random Forest models. For example, due to a mild climate in the region of the upper Rhine valley (tiles HE and BW), winter crops reach maturity earlier (Wizemann et al., 2014) than in other regions of the country. In the German low mountain ranges (tile TH) and northern sites (tile MV), phenological development stages could occur later for a few days or even weeks. Phenological observation records presented in this study have shown considerable temporal differences in their occurrence across seven study sites. Consequently, when models are spatially transferred to the study sites with prominent differences in phenological development phases, the model will fail to predict crop labels accurately. Identifying an 'area of applicability' could be a potential approach to accounting for the fitness of the machine learning or deep learning models to new geographic areas for crop type mapping tasks (Meyer and Pebesma, 2020).

The surface elevation varies substantially across the study sites (see Figure 12b). The tile BW with the highest average altitude (in sampled areas) above 500 m has also shown higher accuracy losses than other tiles. Similar results were reported by Stoian et al. (2019), where higher misclassifications were recorded when models were

spatially transferred to the high altitude zones with more complex topography. Besides shifts in phenology, in comparison to warmer lowlands, geometric distortions in the SAR data such as layover, foreshortening, shadow and high precipitation frequency (see Figure 2) in tile BW could have had a negative effect on the quality of SAR features in our study. This most likely explains the substantial accuracy losses in this tile when only SAR features are used.

Parcels sizes much smaller than those from the remaining tiles (Figure 12a) could have increased accuracy losses in the two southern tiles—HE and BW. As reported in earlier studies (Arias et al., 2020; Löw and Duveiller, 2014), small parcels are harder to classify due to the increased amount of mixed pixels and potential differences in field management. In our case, variations in farm management (e.g., seeding dates, management decisions) or type of farming (Bichler et al., 2005) could be the reason for higher accuracy losses than the parcel sizes themselves.

The uneven spatial distribution of extreme drought in 2018 (Reinermann et al., 2019) combined with low soil quality (Figure 12c) could drive higher misclassification in tiles BB and MV. The north-eastern part of Germany is characterised by sandy soils and lower water holding capacities (Mueller et al., 2014). Thus, in severe drought events such as in 2018, these areas were often hit the hardest (Lüttger and Feike, 2018). Especially, temperature-sensitive crops such as potatoes and sunflowers, and crops with high water demands, such as alfalfa, are among those most affected by drought. While the year 2018 was a particularly interesting case for evaluating the spatial transferability of the Random Forest models under varying climatic and environmental conditions, it would be beneficial to perform such a transferability analysis for other years with more typical climatic conditions across all study sites. Moreover, testing not only spatial but also temporal transferability of machine learning models using multi-sensor features and multi-year crop type information could advance our understanding of model transferability across space and time.

The study design was structured in the typical nested cross-validation manner where an outer loop is used for accuracy estimation using a spatially independent test-set while an inner loop is used for tuning the models via spatial feature selection (Figure 4). The training data distribution across Germany partially reflects the existing studies (Griffiths et al., 2019; Preidl et al., 2020) where dense reference data is available for large areas but completely missing over other large areas, for example, entire federal states. In future studies, the permutation of not all study sites but all possible combinations (site-to-site transfer, 2 × 2 split) could be considered to better understand and maybe compensate for the reasons for accuracy losses in transfer regions.

The proposed 3-step gFFS method showed a good ability to select the relevant features with substantially reduced computational costs

than the original Forward Feature Selection method. However, the main limitation of the proposed method is the risk of losing informative single features in omitted groups.

The accuracy losses were measured to assess the transferability and accuracy declines of the models in unseen geographic spaces by subtracting the F1-score acquired in the target system from that in the reference system. Here, F1-score values in each of the systems were based on the same number of validation points per crop type in each fold (tile). However, in studies with a varying number of samples for different crop types in the folds, the accuracy losses for the target system could be influenced by the dominant presence of 'easy to predict' (e.g., winter rape, maize) or small and complex crops (e.g., alfalfa, grasslands).

6. Conclusions

The presented research examines the spatial transferability of Random Forest models by analysing only optical, only SAR, and optical-SAR feature combinations, and testing if transferability could be improved by spatial feature selection. Based on the study outcomes, the following conclusions were drawn for our crop type mapping case in Germany:

- Random Forest models based on optical-SAR combinations outperform models based on single sensor data in training sites and geographic spaces unseen by the model;
- SAR-based models show the lowest accuracy losses when transferred to an area outside the training regions;
- Performing spatial feature selection on feature sets with only optical data and optical-SAR combination reduces classification accuracy losses in areas where the models were not trained;
- Small classes, grasslands, and alfalfa show high accuracy losses in areas outside the training regions;
- Environmental and geographic variables could aid in explaining or anticipating poor spatial transferability for specific regions.

Remote sensing data is undoubtedly one of the primary information sources for successful crop type mapping. Thus, understanding the strengths and weaknesses of different elements in classification approaches (e.g., datasets, derived features, and classifiers) with respect to spatial model transferability is an important issue, for those faced in practise with a situation where transferability is required.

V

CHAPTER V: SYNTHESIS

1. Findings

This thesis explored the potential of the synergetic use of optical data from Sentinel-2 and SAR data from Sentinel-1 for crop type mapping. The comprehensive review on the topic (Chapter II) led to two detailed studies focusing on Sentinel-1 and Sentinel-2 combinations for mapping a wide range of crops (Chapter III) and testing the spatial transferability of the Random Forest models (Chapter IV). The current Chapter V presents the main findings and concludes this manuscript.

1.1. Objective I: Comprehensive Review

Chapter II presents a detailed analysis of 75 peer-reviewed research publications focusing on the optical-SAR combination for crop type mapping. Most studies reported that the combined use of optical and SAR data produces higher classification accuracies than those based on single sensors. As for single sensors, optical data mostly outperform results based on only SAR data. However, it is necessary to consider that those comparative studies were based on several cloud-free optical and few SAR scenes. SAR data were mainly considered as complementary to the more traditional crop type mapping approach based on only optical data. The usage of dense time-series features was rare, mainly due to the absence of freely available optical and SAR datasets with regular observations over desired locations.

As anticipated in the review study (published in early 2019), the launch of Sentinel-1 and Sentinel-2 satellites proving global coverage with the high spatial-temporal resolution, boosted the research on crop type mapping using the optical-SAR combination. The recent studies on the country (e.g., Blickensdörfer et al., 2022; Van Tricht et al., 2018), continent (e.g., Venter and Sydenham, 2021) and global scale (e.g., World Cereal www.esa-worldcereal.org) based on Sentinel-1 and Sentinel-2 data is an excellent example of it.

It is vital to have representative crop reference data for the study site, along with remote sensing data. The study highlighted the unique role of reference data from the Integrated Administration and Control System (IACS) of the Common Agricultural Policy (CAP). Freely

available, high quality and high spatial precision reference data from IACS would become essential for detailed research on the application of machine learning techniques for crop type mapping at various scales in Europe. For example, all recently published German-wide crop type mapping studies were based on IACS (commonly known as LPIS data) (Asam et al., 2022; Blickensdörfer et al., 2022; Griffiths et al., 2019; Preidl et al., 2020).

Among three data fusion levels, categorized by Pohl and van Genderen (1998), feature-level fusion, i.e., simple band stacking, was used the most compared to pixel-level or decision-level fusion techniques. The review also revealed that most studies on optical-SAR combinations were performed on a small spatial extent (less than 15,000 km²). The studies mainly focused on the major crop types, where the top five include maize, wheat, rice, soya beans and barley. Minor crops (crops grown in a small amount and typically have reduced spatial coverage), such as legumes and nuts, were less researched.

Until the publication of the review study, no research was published on machine learning models' spatial or temporal transferability for crop type mapping with the optical-SAR combination. However, the issue of model transferability is of paramount relevance for accurate crop type mapping, especially when performing large-scale crop type mapping with spatially limited reference (ground truth) data.

The research gaps identified in the review study (Chapter II) were used to develop the research objectives of the following studies presented in Chapter III and Chapter IV.

1.2. Objective II: Large-Scale Detailed Crop Type Map using Sentinel-1 and Sentinel-2 Combination

As emphasised in the review study, less research was done on the optical-SAR combination for mapping a wide variety of crops (including minor crops) at a large scale and using dense time-series features from both sensors. These research gaps were set as research objectives for the study presented in Chapter III, where the analysis of large scale (~30,000 km²), detailed (16 crop classes) crop type classification results based on the dense time-series (bi-weekly features) of the optical-SAR combination were presented. Additional to classification accuracies from three feature sets, the influence of feature selection, parcel sizes and optical data availability on the accuracies were investigated. The study site covered the whole Brandenburg federal state (Germany) territory, and the analysis was done in 2017.

The study outcomes supported the existing knowledge that the optical-SAR combination outperforms the single sensor data. However, contrary to available literature, the performance of SAR data

showed marginally higher accuracies than those based on optical data. The following reasons were used as a possible explanation: (a) SAR data were able to better separate spectrally similar crops (e.g., summer or winter cereals, legumes), which were more challenging for optical data; (b) The study was focused only on crop classes, and no other LULC class (such as water, urban, forest etc.) were considered, where optical data could have been more informative than SAR data; (c) The number of cloud-free optical scenes were very limited in 2017 and affected the quality of the optical features.

The feature selection showed that mono-temporal optical scenes are more informative than mono-temporal SAR scenes. However, with the addition of features, the increase in accuracies of SAR-based models was much higher than those based on optical data only.

In line with the previous studies (Van Tricht et al., 2018), the Random Forest models showed that they could handle the high feature dimensionality. An increasing number of features did not harm the classification accuracies (Figure III-11).

Permanent grasslands	Potatoes	Winter rape	Summer oat
Temporal grasslands	Sugar beets	Winter barley	Legume mixture
Maize	Winter wheat	Winter triticale	Peas-Beans
Sunflowers	Winter rye	Summer barley	Lupins
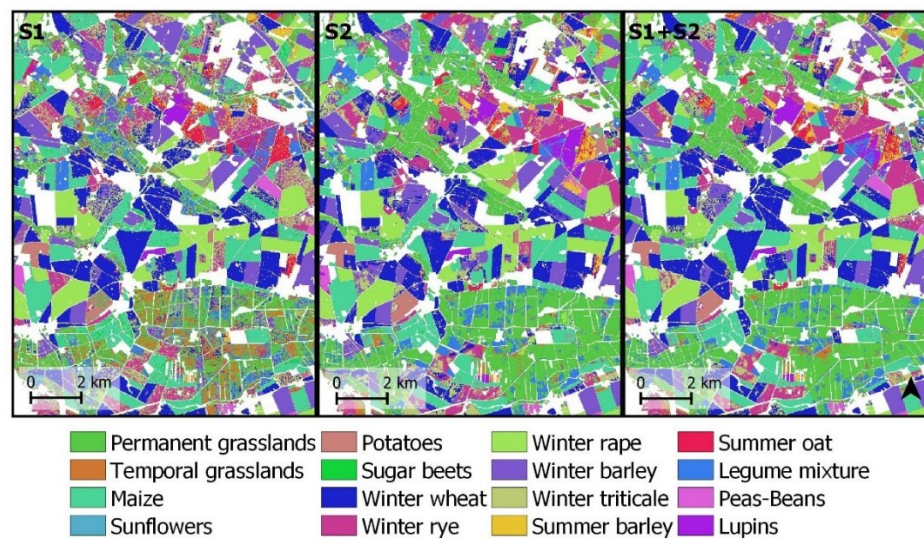
Figure V-1. Crop type classification results from Chapter III based on only SAR (S1), only optical (S2) and optical-SAR feature combination (S1+S2).

Crop type maps based on the optical-SAR combination show a more homogenous representation of the crop fields than optical or SAR only crop type maps (Figure V-1). The reduction of the 'salt-and-pepper effect', which is clearly visible on maps based on only SAR and only optical data, could be explained due to the fewer misclassifications, particularly among minor crops and winter cereals.

As anticipated, on average, small parcels had higher misclassifications than big parcels among all investigated crop classes except for grasslands (see Figure III-13). Also, crop classes with a large share of small fields, such as potato and lupins (with an average field size of 10 ha), showed higher misclassifications than major crops with

more than 20 ha as winter rapeseed or sugar beets. This makes minor crops that are usually planted on small-sized fields harder to classify.

This study's training and test samples were randomly distributed across study sites, which in our case was the entire state of Brandenburg. Suppose a trained machine learning model is used to classify croplands in another state where no reference data is available. In that case, the resulting accuracies will likely differ from those calculated on the training sites in Brandenburg. The accuracy variation results from the absence of representative samples in the new sites. However, accuracy variations are also possible due to the input data quality, crop types and complexity of the machine learning models. The issue of spatial transferability of Random Forest models was investigated in Chapter IV.

1.3. Objective III: Spatial Transferability of Crop Type Classification Models based on Sentinel-1 and Sentinel-2 Combination

In **Chapter IV**, the spatial transferability of Random Forest models was tested using three input datasets and two feature selection methods. As reported in Chapter III, the combination of optical-SAR features showed higher classification accuracies in training sites than with single sensors. The new finding was that the combination outperforms single sensors in unseen-by-model study sites. This new finding highlights another essential advantage of using the optical-SAR combination over single sensor data for crop type mapping tasks.

The study experiments also showed that SAR-based models have lower accuracy losses than models based on optical data or optical-SAR data combinations. When spatially transferring the model to unseen by model areas, i.e., new environmental and management conditions with no training samples, the SAR-based models show better accuracies than those based on optical data. SAR provide weather-independent and regular land surface observations. Optical data is weather dependent, and the quality is directly related to the presence of clouds or cloud shadows. No such comparative study quantitively proved that the models based on SAR data outperform those based on optical data in sites unseen by the model. This finding was one of the central in the manuscript. However, one should note that this finding is valid for regions where persistent cloud cover is an issue, like Germany.

The study results showed that spatial feature selection helps to reduce classification accuracy losses in sites with no training data when only optical data or the combination of optical and SAR data is used. As previously reported (Meyer et al., 2019), performing spatial feature selection helps to remove redundant input features that carry site-specific information. This substantially lessens the number of features while preserving the most informative ones and allows to build simpler machine-learning models.

The performance, i.e. losses in classification accuracy, of models based on features selected using spatial and random cross-validation, the resampling method used to evaluate model performance, did not differ when using only SAR data. However, only optical and optical-SAR combination using spatial feature selection helped to reduce the accuracy losses in unseen sites.

The reduction of classification accuracies when the machine learning model is spatially transferred to a previously unseen site is not surprising. The causes of such accuracy reduction are mainly due to the absence of representative samples from the target sites. When the target site's environmental and climatic conditions differ substantially from the reference sites, the model fails to predict the crop labels correctly. The analysis of auxiliary variables, such as land surface elevation, soil quality rating, phenological observations and parcel sizes, helped to understand the possible driving forces of higher misclassifications in regions across Germany. For example, higher elevation and more complex topography in the southern part of the country could affect the quality of the SAR features, resulting in low classification accuracies of the experiments based on solo SAR data. Crop classes with a higher number of small parcels (e.g., sunflower, potatoes) have shown a more significant misclassification rate than those with bigger parcels (e.g. winter cereals). Regions with a lower soil quality rating (north-eastern Germany) were more affected by the drought in 2018, resulting in substantially different phenological growth than the rest of the country. This led to high misclassification in the region.

Such analysis of environmental and phenological variables is vital to understand the driving forces behind the high misclassifications in particular sites. When large-scale crop type classification work is carried out, such analysis could also facilitate understanding which region could be error-prone due to differing climatic or environmental conditions. On the other hand, using environmental and climatic variables could be helpful before reference data collection, promoting data collection from more diverse regions.

2. Discussion and Outlook

Detailed crop-type mapping with dense time-series data and spatial model transferability were the central topics of two experimental studies (Chapters III and IV). In addition to each detailed discussion provided in Chapter III and Chapter IV, the current sub-chapter offers a general discussion of the research outcomes in the context of recent literature.

2.1.   *Crop Type Mapping with Sentinel-1 and Sentinel-2*

2.1.1.   *Sentinel-1 and Sentinel-2: Different but Complementary*

Optical and SAR sensors collect information at different regions of the electromagnetic spectrum and capture substantially different characteristics of plants. Combining optical and SAR data gives more diverse information about sensed targets rather than single sensor data or a combination of sensors with similar data acquisition properties (e.g., optical-optical, SAR-SAR). This is supported by studies reporting classification accuracy increases when SAR data is added to the combination of optical sensors such as Sentinel-2 and Landsat-8 (e.g., Blickensdörfer et al., 2022; Song et al., 2021).

Another fundamental difference between the two datasets is their weather dependence. In scenarios where persistent cloud cover does not allow building good quality optical features over specific time windows, SAR data can still deliver relevant land surface observations. Notably, combining optical-SAR data should be the preferred option for countries like Germany, where persistent cloud cover could be an issue.

The data from Sentinel-1 and Sentinel-2, with their global and frequent acquisitions, give great opportunities for consistent mapping of global agricultural lands. These remote sensing datasets became the most frequently used ones for crop type mapping tasks in the last few years. The new developments towards cloud-based solutions could be one of the triggers.

For the last few years, accessing the analysis-ready Sentinel-1 and Sentinel-2 data has required little effort. With the shift towards cloud-native computing and data storage, the combination of satellite data became more straightforward. As one of the core freely available Earth Observation datasets, Sentinel-1 and Sentinel-2 data is part of many free (for research purposes) and commercial cloud-computing services such as Data and Information Access Service (DIAS) from ESA, Google Earth Engine from Google, Planetary Computer from Microsoft, UP42 geospatial platform from UP42, Sentinel Hub from Synergise. With the growing popularity of such geospatial services, combining both sensors for various application domains would soon become the new normality.

2.1.2.   *Freely Available Reference Data*

Additional to the remote sensing data, the availability of good quality reference data for the supervised machine learning tasks is critical. In recent years, more crop type reference data have become available to the research community and the general public. For example, Radiant MLHub (https://mlhub.earth/datasets) from Radian Earth gathers and freely distributes the crop reference data collected worldwide, but

with a focus on African countries. At the EU level, LPIS data became more accessible for research purposes. Moreover, some EU member states such as Austria, Denmark and Slovenia annually publish country-wide LPIS data. The new initiative on harmonizing the crop reference data at the European level has been started by the EuroCrops project (https://www.eurocrops.tum.de/). The availability of such openly accessible, high quality and quantity reference data enables more detailed research on crop type mapping as presented in Chapters III and IV.

The unique role of LPIS data for studies on crop type mapping in EU countries could be recognized by the number of published papers where LPIS data is used as a reference data source (Asam et al., 2022; Blickensdörfer et al., 2022; Griffiths et al., 2019; Preidl et al., 2020; Woźniak et al., 2022).

2.1.3. Mapping Minor Crops

The inclusion of minor crops to the classification scheme should be more often considered by the crop type classification studies in order to gain more knowledge on the advantages and limitations of various datasets and model-building approaches. For example, the results of the second paper (Chapter III) illustrated that some minor crops in Germany are better mapped with only SAR data (e.g., pean-beans, lupins) than with only optical and wise-versa (e.g., legume mixture). The results of the third paper (Chapter IV) indicated that crops with small field sizes (mainly minor crops) show higher accuracy losses in unseen sites compared to major crops with, on average, larger field sizes.

While minor crops were included in the analysis in Chapters III and IV, no particular strategy for improving the accuracies of minor crops was conducted as in Waldner et al. (2019). This was mainly due to the availability of reference data that allowed balanced stratified sampling. However, the availability of a large amount of reference data is often not the case. Thus, other sampling designs should be tested to understand their effect on mapping minor crops with the optical-SAR combination.

2.1.4. Analysis of Auxiliary Variables

The analysis of auxiliary information presented in two papers (Chapters III and IV) helped to explain the classification accuracies to a certain extent. By-products of input reference and remote sensing data such as field size, reference pixel location within the field and cloud cover probability are often omitted and not considered for post-classification analysis. Also, auxiliary environmental variables such as soil quality or surface elevation could be used to understand the output

classification results better. The post-classification analysis should be more often considered along with the classification accuracies.

2.1.5. *Moving from Random Forest to Deep Learning Models'*

Both studies presented in the thesis (Chapters III and IV) use the Random Forest classifier. The choice was motivated by its reported superiority over other machine learning classifiers (Sun et al., 2019) and its robustness against sample noise or large feature space (Belgiu and Drăgu, 2016). However, a comparison to other classifiers such as XGboost or Convolutional Neural Networks (CNN) would be desired to understand their advantages or disadvantages over Random Forest for crop type mapping with optical-SAR data combination.

2.2. *Spatial Transferability of Random Forest Models*

2.2.1. *The complexity of Random Forest Models*

As illustrated in the third paper (Chapter IV), simpler models with fewer features are more spatially transferable than complex models. This finding aligns with Ferraciolli et al. (2019), where simpler models performed better on an independent test set for sugarcane yield predictions. Apart from more robust performance concerning spatial transferability, simpler models can be better interpreted and explained (Roscher et al., 2020). While Random Forest can handle large feature space without negatively affecting its performance in the training sites (see Chapter III), it is advised to reduce the feature space for cases when the spatial transfer of the models is expected (Orynbaikyzy et al., 2022).

2.2.2. *Groups-wise Feature Selection with Spatial Cross-Validation*

The grouping of the features by bands and dates helped to substantially reduce the computational complexity of the feature selection procedure while analyzing all features from Sentinel-2 and Sentinel-1 simultaneously. The results from both papers (Chapter III and IV) showed a similar pattern by emphasizing the importance of short-wave infrared, red-edge bands, vegetation indices (e.g., NDVI, NDYI, PSRI) from Sentinel-2 together with VV, VH bands with their ratio from Sentinel-1. To avoid the unnecessary burden of handling all features from two sensors, using only a subset of proven and effective features is suggested. In such cases, it will not be necessary to group features.

In line with the findings of Hengl et al. (2021) and Meyer et al. (2019), the results of Chapter IV have shown that feature selection with spatial cross-validation can remove redundant features and reduce classification accuracy losses in previously unseen sites. However, reducing features also reduces absolute accuracies in reference and

target systems. The reduction of accuracies in the reference system could indicate that some relevant features were also removed during feature selection.

When partitioning the reference data into cross-validation folds, the spatial distance was used as a primary indicator of dissimilarity. However, as it was emphasized by Yates et al. (2018), the model transferability is more dependent on the environmental dissimilarity between reference and target regions rather than spatial. Future research should investigate the spatial or spatial-temporal transferability of machine learning models across diverse environments and climatic regions.

2.2.3. Accuracy Estimation with Spatially Independent Test-Set: When Relevant?

A spatially independent test set is expected to deliver unbiased estimates of the model's performance outside the training sites. In recent years, more and more research papers within the geospatial domain are calling for the usage of accuracy estimation techniques, such as e.g. spatial cross-validation, that account for spatial dependence present in the data (Ferraciolli et al., 2019; Karasiak et al., 2022, 2019; Ploton et al., 2020). However, spatial cross-validation could be the wrong strategy for estimating the crop type map quality. When training sites are a large part of the general map extent, using spatially independent test sets outside the training sites may result in negatively biased accuracy estimates. This was recently argued by Wadoux et al. (2021), where random cross-validation led to lower bias than spatial cross-validation for large-scale above-ground biomass mapping. Designing the sampling strategy as presented in Pohjankukka et al. (2017), where training and test samples lie within a certain spatial-temporal distance from each other, should be considered for getting unbiased accuracy estimates for large-scale map products.

3. Conclusion

This thesis addressed the advantages and limitations of the synergetic usage of optical data from Sentinel-2 and SAR data from Sentinel-1 satellites for crop type mapping.

The comprehensive review of the state-of-the-art research helped to structure existing scientific knowledge on the subject and uncover the research gaps. The anticipated substantial increase of studies on optical-SAR combination after the launch of Sentinel-1 and Sentinel-2 satellites became a reality during the last few years.

The investigation of the joint use of dense time-series features from both satellites revealed the superior performance of optical-SAR

feature combination over single sensor datasets. The proposed feature selection workflow helped to emphasize optical and SAR features with high relevance for the classifier. Also, the study illustrated the added values of SAR features in large-scale mapping scenarios where building consistent and good quality optical features could be challenging. Particularly in cloud-prone regions such as northern Germany.

In the context of large-scale crop type mapping, the spatial transferability of the Random Forest models was addressed. It has been shown that accuracy losses outside of the training sites are not only driven by geographic and environmental dissimilarities but also by the type of input remote sensing data and the way the machine learning classifier was built.

Current research intends to contribute to the scientific effort to automate the generation of high-quality spatial-temporal information on agricultural lands. Such information derived from remote sensing datasets is essential in successful food production and supply chain management.

REFERENCES

Abdikan, S., Bilgin, G., Sanli, F.B., Uslu, E., Ustuner, M., 2015. Enhancing land use classification with fusing dual-polarized TerraSAR-X and multispectral RapidEye data. J. Appl. Remote Sens. 9, 096054. https://doi.org/10.1117/1.jrs.9.096054

Abdikan, S., Sanli, F.B., 2012. Comparison of different fusion algorithms in urban and agricultural areas using sar (palsar and radarsat) and optical (spot) images. Bol. Ciências Geodésicas 18, 509–531. https://doi.org/10.1590/s1982-21702012000400001

Ahern, F.J., Goodenough, D.G., Grey, A.L., Ryerson, R.A., Vilbikaitis, R.J., Goldberg, M., 1978. Simultaneous microwave and optical wavelength observations of agricultural targets. Can. J. Remote Sens. 4, 127–142. https://doi.org/10.1080/07038992.1978.10854975

Ajadi, O.A., Barr, J., Liang, S.-Z., Ferreira, R., Kumpatla, S.P., Patel, R., Swatantran, A., 2021. Large-scale crop type and crop area mapping across Brazil using synthetic aperture radar and optical imagery. Int. J. Appl. Earth Obs. Geoinf. 97, 102294. https://doi.org/10.1016/j.jag.2020.102294

Aneece, I., Thenkabail, P.S., 2021. Classifying crop types using two generations of hyperspectral sensors (Hyperion and DESIS) with machine learning on the cloud. Remote Sens. 13. https://doi.org/10.3390/rs13224704

Arias, M., Campo-Bescós, M.Á., Álvarez-Mozos, J., 2020. Crop classification based on temporal signatures of Sentinel-1 observations over Navarre province, Spain. Remote Sens. 12. https://doi.org/10.3390/rs12020278

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

Atzberger, C., 2013. Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. Remote Sens. 5, 949–981. https://doi.org/10.3390/rs5020949

Azzari, G., Jain, M., Lobell, D.B., 2017. Remote Sensing of Environment Towards fi ne resolution global maps of crop yields : Testing multiple methods and satellites in three countries. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2017.04.014

Bajocco, S., Vanino, S., Bascietto, M., Napoli, R., 2021. Exploring the drivers of sentinel-2-derived crop phenology: The joint role of climate, soil, and land use. Land 10. https://doi.org/10.3390/land10060656

Ban, Y., Hu, H., Rangel, I.M., 2010. Fusion of Quickbird MS and RADARSAT SAR data for urban land-cover mapping: Object-based and knowledge-based approach. Int. J. Remote Sens. 31, 1391–1410. https://doi.org/10.1080/01431160903475415

Bargiel, D., 2017. A new method for crop classification combining time series of radar images and crop phenology information. Remote Sens. Environ. 198, 369–383. https://doi.org/10.1016/j.rse.2017.06.022

Bazzi, H., Ienco, D., Baghdadi, N., Zribi, M., Demarez, V., 2020. Distilling before Refine: Spatio-Temporal Transfer Learning for Mapping Irrigated Areas Using Sentinel-1 Time Series. IEEE Geosci. Remote Sens. Lett. 17, 1909–1913. https://doi.org/10.1109/LGRS.2019.2960625

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future köppen-geiger climate classification maps at 1-km resolution. Sci. Data 5, 1–12. https://doi.org/10.1038/sdata.2018.214

Belgiu, M., Bijker, W., Csillik, O., Stein, A., 2021. Phenology-based sample generation for supervised crop type classification. Int. J. Appl. Earth Obs. Geoinf. 95, 102264. https://doi.org/10.1016/j.jag.2020.102264

Belgiu, M., Csillik, O., 2018. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote Sens. Environ. 204, 509–523. https://doi.org/10.1016/j.rse.2017.10.005

Belgiu, M., Drăgu, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Benos, L., Tagarakis, A.C., Dolias, G., Berruto, R., Kateris, D., Bochtis, D., 2021. Machine learning in agriculture: A comprehensive updated review. Sensors 21, 1–55. https://doi.org/10.3390/s21113758

Betbeder, J., Laslier, M., Corpetti, T., Pottier, E., Corgne, S., Hubert-Moy, L., 2014. Multi-temporal optical and radar data fusion for crop monitoring: Application to an intensive agricultural area in BRITTANY(France). pp. 1493–1496. https://doi.org/10.1109/IGARSS.2014.6946720

BGR, 2013. The Product Center of the Federal Institute for Geosciences and Natural Resources (BGR) [WWW Document]. URL https://produktcenter.bgr.de/terraCatalog/Start.do (accessed 6.9.21).

Bichler, B., Lippert, C., Häring, A.M., Dabbert, S., 2005. The determinants of the spatial distribution of organic farming in Germany. Berichte Uber Landwirtschaft 83, 50–75.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg.

Blaes, X., Vanhalle, L., Defourny, P., 2005. Efficiency of crop identification based on optical and SAR image time series. Remote Sens. Environ. 96, 352–365. https://doi.org/10.1016/j.rse.2005.03.010

Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., Hostert, P., 2022. 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. 269, 112831. https://doi.org/10.1016/j.rse.2021.112831

Brandenburg Surveying and Geospatial Information Office, 2017. Daten aus dem Agrarförderantrag [WWW Document].

Breiman, L., 2001a. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., 2001b. Random forests. Random For. 1–122. https://doi.org/10.1201/9780367816377-11

Brisco, B., Brown, R.J., Manore, M.J., 1989. Early season crop discrimination

with combined SAR and TM data. Canadian journal of remote sensing.

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sens. Environ. 210, 35–47. https://doi.org/10.1016/j.rse.2018.02.045

Chavez, P.S., Kwarteng, A.Y., 1989. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. Photogramm. Eng. Remote Sens. 55, 339–348.

Cheng, Y., Yu, L., Cracknell, A.P., Gong, P., 2016. Oil palm mapping using Landsat and PALSAR: a case study in Malaysia. Int. J. Remote Sens. 37, 5431–5442. https://doi.org/10.1080/01431161.2016.1241448

Clauss, K., Ottinger, M., Kuenzer, C., 2018. Mapping rice areas with Sentinel-1 time series and superpixel segmentation. Int. J. Remote Sens. 39, 1399–1420. https://doi.org/10.1080/01431161.2017.1404162

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.C., Skakun, S. V., Justice, C., 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote Sens. Environ. 219, 145–161. https://doi.org/10.1016/j.rse.2018.09.002

Colditz, R.R., Wehrmann, T., Bachmann, M., Steinnocher, K., Schmidt, M., Strunz, G., Dech, S., 2006. Influence of image fusion approaches on classification accuracy: A case study. Int. J. Remote Sens. 27, 3311–3335. https://doi.org/10.1080/01431160600649254

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Conrad, C., Dech, S., Dubovyk, O., Fritsch, S., Klein, D., Löw, F., Schorcht, G., Zeidler, J., 2014. Derivation of temporal windows for accurate crop discrimination in heterogeneous croplands of Uzbekistan using multitemporal RapidEye images. Comput. Electron. Agric. 103, 63–74. https://doi.org/10.1016/j.compag.2014.02.003

Conrad, C., Fritsch, S., Zeidler, J., Rücker, G., Dech, S., 2010. Per-field irrigated crop classification in arid Central Asia using SPOT and ASTER data. Remote Sens. 2, 1035–1056. https://doi.org/10.3390/rs2041035

Conrad, C., Löw, F., Lamers, J.P.A., 2017. Mapping and assessing crop diversity in the irrigated Fergana Valley, Uzbekistan. Appl. Geogr. 86, 102–117. https://doi.org/10.1016/j.apgeog.2017.06.016

d'Andrimont, R., Taymans, M., Lemoine, G., Ceglar, A., Yordanov, M., van der Velde, M., 2020. Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and -2 time series. Remote Sens. Environ. 239, 111660. https://doi.org/10.1016/j.rse.2020.111660

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

Dalla Mura, M., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J., Benediktsson, J.A., 2015. Challenges and Opportunities of Multimodality and Data Fusion in Remote Sensing. Proc. IEEE 103,

1585–1601. https://doi.org/10.1109/JPROC.2015.2462751

De Alban, D.J., Connette, M.G., Oswald, P., Webb, L.E., 2018. Combined Landsat and L-Band SAR Data Improves Land Cover Classification and Change Detection in Dynamic Tropical Landscapes. Remote Sens. . https://doi.org/10.3390/rs10020306

Defourny, P., Bontemps, S., Bellemans, N., Cara, C., Dedieu, G., Guzzonato, E., Hagolle, O., Inglada, J., Nicola, L., Rabaute, T., Savinaud, M., Udroiu, C., Valero, S., Bégué, A., Dejoux, J.F., El Harti, A., Ezzahar, J., Kussul, N., Labbassi, K., Lebourgeois, V., Miao, Z., Newby, T., Nyamugama, A., Salh, N., Shelestov, A., Simonneaux, V., Traore, P.S., Traore, S.S., Koetz, B., 2019a. 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. 221, 551–568. https://doi.org/10.1016/j.rse.2018.11.007

Defourny, P., Moreau, I., Wolter, J., 2019b. ECoLaSS Evolution of Copernicus Land Services based on Sentinel data D8.2 'D33.1b-Time Series Analysis for Thematic Classification (Issue 2)'.

Defourny, P., Moreau, I., Wolter, J., Khalil, E., Gallaun, H., Miletich, P., Puhm, M., Villerot, S., Pennec, A., Lhernould, A., Courmont, Y., De-Wyse, T., Herrman, D., Schweitz, K., Leutner, B., 2019c. D33.1b-Time Series Analysis for Thematic Classification (Issue 2) [WWW Document]. URL https://www.ecolass.eu/project-deliverables (accessed 2.19.21).

Demarez, V., Helen, F., Marais-Sicre, C., Baup, F., 2019. In-season mapping of irrigated crops using Landsat 8 and Sentinel-1 time series. Remote Sens. 11. https://doi.org/10.3390/rs11020118

Denize, J., Hubert-Moy, L., Betbeder, J., Corgne, S., Baudry, J., Pottier, E., 2018. Evaluation of using sentinel-1 and -2 time-series to identifywinter land use in agricultural landscapes. Remote Sens. 11. https://doi.org/10.3390/rs11010037

Dong, J., Zhuang, D., Huang, Y., Fu, J., 2009. Advances in multi-sensor data fusion: Algorithms and applications. Sensors 9, 7771–7784. https://doi.org/10.3390/s91007771

Dusseux, P., Corpetti, T., Hubert-Moy, L., Corgne, S., 2014. Combined use of multi-temporal optical and Radar satellite images for grassland monitoring. Remote Sens. 6, 6163–6182. https://doi.org/10.3390/rs6076163

DWD, 2018a. DWD Climate Data Center (CDC): Grids of monthly averaged daily air temperature (2m) over Germany, version v1.0. [WWW Document]. URL ftp://opendata.dwd.de/climate_environment/CDC/grids_germany/monthly/air_temperature_mean/ (accessed 12.21.20).

DWD, 2018b. DWD Climate Data Center (CDC): Grids of monthly total precipitation over Germany, version v1.0. [WWW Document]. URL ftp://opendata.dwd.de/climate_environment/CDC/grids_germany/monthly/precipitation/ (accessed 12.21.20).

DWD, 2018c. The Weather in Germany in 2018 12–13.

DWD, 2018d. Climate Data Center (CDC): Phenological observations of crops from sowing to harvest (annual reporters, recent), Version v006. Last accessed: 26.05.2021 [WWW Document].

Ehlers, M., 1991. Multisensor image fusion techniques in remote sensing. ISPRS J. Photogramm. Remote Sens. 46, 19–30. https://doi.org/10.1016/0924-2716(91)90003-E

Erasmi, S., Twele, A., 2009. Regional land cover mapping in the humid tropics using combined optical and SAR satellite data - A case study from Central Sulawesi, Indonesia. Int. J. Remote Sens. 30, 2465–2478. https://doi.org/10.1080/01431160802552728

FAO, 2009. How to Feed the World in 2050.

Feingersh, T., Gorte, B.G.H., Leeuwen, H.J.C. Van, 2001. Fusion of SAR and SPOT image data for crop mapping, in: IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217). pp. 873–875 vol.2. https://doi.org/10.1109/IGARSS.2001.976665

Ferraciolli, M.A., Bocca, F.F., Rodrigues, L.H.A., 2019. Neglecting spatial autocorrelation causes underestimation of the error of sugarcane yield models. Comput. Electron. Agric. 161, 233–240. https://doi.org/10.1016/j.compag.2018.09.003

Filella, I., Peñuelas, J., 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. Int. J. Remote Sens. 15, 1459–1470. https://doi.org/10.1080/01431169408954177

Firouzabadi, P.Z., Sadidy, J., 2006. Paddy rice mapping of the Caspian Sea coast using microwave and optical remotely sensed data. Remote Sens. Agric. Ecosyst. Hydrol. VIII 6359, 63591A. https://doi.org/10.1117/12.690750

Fiumara, A., Pierdicca, N., 1989. Evaluation of classification results obtained with combined multitemporal optical and microwave data 2, 787–790.

Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., 2011. Solutions for a cultivated planet. Nature 478, 337–342. https://doi.org/10.1038/nature10452

Fontanelli, G., Crema, A., Azar, R., Stroppiana, D., Villa, P., Boschetti, M., 2014. Agricultural crop mapping using optical and SAR multi-temporal seasonal data: A case study in Lombardy region, Italy. Int. Geosci. Remote Sens. Symp. 1489–1492. https://doi.org/10.1109/IGARSS.2014.6946719

Foody, G., See, L., Fritz, S., Moorthy, I., Perger, C., Schill, C., Boyd, D., 2018. Increasing the accuracy of crowdsourced information on land cover via a voting procedure weighted by information inferred from the contributed data. ISPRS Int. J. Geo-Information 7. https://doi.org/10.3390/ijgi7030080

Foody, G.M., Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. Remote Sens. Environ. 103, 179–189. https://doi.org/10.1016/j.rse.2006.04.001

Foody, G.M., McCulloch, M.B., Yates, W.B., 1995. The effect of training set size and composition on artificial neural network classification. Int. J. Remote Sens. 16, 1707–1723. https://doi.org/10.1080/01431169508954507

Foody, G.M., Pal, M., Rocchini, D., Garzon-Lopez, C.X., Bastin, L., 2016. The sensitivity of mapping methods to reference data quality: Training supervised image classifications with imperfect reference data. ISPRS Int. J. Geo-Information 5. https://doi.org/10.3390/ijgi5110199

Forkuor, G., Conrad, C., Thiel, M., Landmann, T., Barry, B., 2015. Evaluating the sequential masking classification approach for improving crop discrimination in the Sudanian Savanna of West Africa. Comput. Electron. Agric. 118, 380–389. https://doi.org/10.1016/j.compag.2015.09.020

Forkuor, G., Conrad, C., Thiel, M., Ullmann, T., Zoungrana, E., 2014. Integration of optical and synthetic aperture radar imagery for improving crop mapping in northwestern Benin, West Africa. Remote Sens. 6, 6472–6499. https://doi.org/10.3390/rs6076472

Franklin, S.E., Blodgett, C.F., 1993. An example of satellite multisensor data fusion. Comput. Geosci. 19, 577–583. https://doi.org/10.1016/0098-3004(93)90083-H

Frantz, D., Haß, E., Uhl, A., Stoffels, J., Hill, J., 2018. Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. Remote Sens. Environ. 215, 471–481. https://doi.org/10.1016/j.rse.2018.04.046

Frantz, D., Röder, A., Stellmes, M., Hill, J., 2017. Phenology-adaptive pixel-based compositing using optical earth observation imagery. Remote Sens. Environ. 190, 331–347. https://doi.org/10.1016/j.rse.2017.01.002

Gadiraju, K.K., Vatsavai, R.R., 2020. Comparative analysis of deep transfer learning performance on crop classification, Proceedings of the 9th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, BIGSPATIAL 2020. Association for Computing Machinery. https://doi.org/10.1145/3423336.3431369

Gangkofner, U.G., Pradhan, P.S., Holcomb, D.W., 2008. Optimizing the high-pass filter addition technique for image fusion. Photogramm. Eng. Remote Sensing 74, 1107–1118. https://doi.org/10.14358/PERS.74.9.1107

Geary, R.C., 1954. The Contiguity Ratio and Statistical Mapping. Society 5, 115–127.

Gebhardt, S., Huth, J., Nguyen, L.D., Kuenzer, C., 2013. A comparison of TerraSAR-X Quadpol backscattering with RapidEye multispectral vegetation indices over rice fields in the Mekong Delta , Vietnam. Int. J. Remote Sens. 37–41.

German National Weather Service, 2019. German National Weather Service (Deutscher Wetterdienst - DWD) [WWW Document].

Gerstmann, H., Doktor, D., Gläßer, C., Möller, M., 2016. PHASE: A geostatistical model for the Kriging-based spatial prediction of crop phenology using public phenological and climatological observations. Comput. Electron. Agric. 127, 726–738. https://doi.org/10.1016/j.compag.2016.07.032

Ghassemi, B., Dujakovic, A., Żółtak, M., Immitzer, M., Atzberger, C., Vuolo, F., 2022. Designing a European-Wide Crop Type Mapping Approach Based on Machine Learning Algorithms Using LUCAS Field Survey and Sentinel-2 Data. Remote Sens. 14, 541. https://doi.org/10.3390/rs14030541

Gibril, M.B.A., Bakar, S.A., Yao, K., Idrees, M.O., Pradhan, B., 2017. Fusion of RADARSAT-2 and multispectral optical remote sensing data for LULC extraction in a tropical agricultural area. Geocarto Int. 32, 735–748. https://doi.org/10.1080/10106049.2016.1170893

Gilcher, M., Udelhoven, T., 2021. Field geometry and the spatial and temporal generalization of crop classification algorithms—a randomized approach to compare pixel based and convolution based methods. Remote Sens. 13, 1–20. https://doi.org/10.3390/rs13040775

Godfray, H. Charles J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C., 2010. Food security: The challenge of feeding 9 billion people. Science (80-.). 327, 812–818. https://doi.org/10.1126/science.1185383

Godfray, H Charles J, Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C., 2010. Food security: the challenge of feeding 9 billion people. Science (80-.). 327, 812–818.

Goodenough, D.G., Narendra, P.M., O'Neill, K., 1978. Feature subset selection in remote sensing. Can. J. Remote Sens. 4, 143–148. https://doi.org/10.1080/07038992.1978.10854976

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Griffiths, P., Nendel, C., Hostert, P., 2019. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. Remote Sens. Environ. 220, 135–151. https://doi.org/10.1016/j.rse.2018.10.031

Gutzler, C., Helming, K., Balla, D., Dannowski, R., Deumlich, D., Glemnitz, M., Knierim, A., Mirschel, W., Nendel, C., Paul, C., Sieber, S., Stachow, U., Starick, A., Wieland, R., Wurbs, A., Zander, P., 2015. Agricultural land use changes - A scenario-based sustainability impact assessment for Brandenburg, Germany. Ecol. Indic. 48, 505–517. https://doi.org/10.1016/j.ecolind.2014.09.004

Guyon, I., Elisseeff, A., 2011. An Introduction to Variable and Feature Selection 3, 1–26.

H, M., B, B., 2004. The Application of C-Band Polarimetric SAR for Agriculture: A Review. Can. J. Remote Sens. 30, 525.

Hagolle, O., Huc, M., Desjardins, C., Auer, S., Richter, R., 2017. MAJA ATBD Algorithm Theoretical Basis Document. Development 0–39.

Haldar, D., Patnaik, C., 2010. Synergistic use of multi-temporal Radarsat SAR and AWiFS data for Rabi rice identification. J. Indian Soc. Remote Sens. 38, 153–160. https://doi.org/10.1007/s12524-010-0006-x

Haldar, D., Patnaik, C., Mohan, S., Chakraborty, M., 2012. Jute and tea discrimination through fusion of sar and optical data. Prog. Electromagn. Res. B 337–354. https://doi.org/10.2528/PIERB11123011

Hall, D.L., McMullen, S.A.H., 2004. Mathematical techniques in multisensor data fusion. Artech House.

Hao, P., Di, L., Zhang, C., Guo, L., 2020. Transfer Learning for Crop classification with Cropland Data Layer data (CDL) as training samples.

Sci. Total Environ. 733, 138869. https://doi.org/10.1016/j.scitotenv.2020.138869

Haralik, R.M., 1979. Statistical and structured approaches to the description of textures. TIIRE pp.98-118.

Harris, J.R., Murray, R., Hirose, T., 1990. IHS transform for the integration of radar imagery with other remotely sensed data. Photogramm. Eng. Remote Sens. 56, 1631–1641.

Hatfield, J.L., Prueger, J.H., 2010. Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. Remote Sens. 2, 562–578. https://doi.org/10.3390/rs2020562

Hill, M.J., Ticehurst, C.J., Lee, J. Sen, Grunes, M.R., Donald, G.E., Henry, D., 2005. Integration of optical and radar classifications for mapping pasture type in Western Australia. IEEE Trans. Geosci. Remote Sens. 43, 1665–1680. https://doi.org/10.1109/TGRS.2005.846868

Hong, G., Zhang, A., Zhou, F., Brisco, B., 2014. Integration of optical and synthetic aperture radar (SAR) images to differentiate grassland and alfalfa in Prairie area. Int. J. Appl. Earth Obs. Geoinf. 28, 12–19. https://doi.org/10.1016/j.jag.2013.10.003

Hong, G., Zhang, A., Zhou, F., Townley-Smith, L., Brisco, B., Olthof, I., 2011. Crop-type identification potential of Radarsat-2 and MODIS images for the Canadian prairies. Can. J. Remote Sens. 37, 45–54. https://doi.org/10.5589/m11-026

Hong, G., Zhang, Y., Mercer, B., 2009. A wavelet and IHS integration method T fuse high resolution SAR with moderai resolution multispectral images. Photogramm. Eng. Remote Sensing 75, 1213–1223. https://doi.org/10.14358/PERS.75.10.1213

Huang, X., Zhang, L., Li, P., 2007. Classification and Extraction of Spatial Features in Urban Areas Using High-Resolution Multispectral Imagery 4, 260–264.

Ianninia, L., Molijn, R.A., Hanssen, R.F., 2013. Integration of multispectral and C-band SAR data for crop classification. Remote Sens. Agric. Ecosyst. Hydrol. XV 8887, 88871D. https://doi.org/10.1117/12.2029330

Immitzer, M., Neuwirth, M., Böck, S., Brenner, H., Vuolo, F., Atzberger, C., 2019. Optimal input features for tree species classification in Central Europe based on multi-temporal Sentinel-2 data. Remote Sens. 11. https://doi.org/10.3390/rs11222599

Immitzer, M., Vuolo, F., Atzberger, C., 2016. Firast experience with Sentinel-2 data for crop and tree species classifications in central Europe. Remote Sens. 8. https://doi.org/10.3390/rs8030166

Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre, G., Bontemps, S., Defourny, P., Koetz, B., 2015. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. Remote Sens. 7, 12356–12379. https://doi.org/10.3390/rs70912356

Inglada, J., Vincent, A., Arias, M., Marais-Sicre, C., 2016. Improved early crop type identification by joint use of high temporal resolution sar and optical image time series. Remote Sens. 8. https://doi.org/10.3390/rs8050362

128

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. Remote Sens. 9, 95. https://doi.org/10.3390/rs9010095

Jain, A., 1997. Feature selection: evaluation, application, and small sample performance. IEEE Trans. Pattern Anal. Mach. Intell. 19, 153–158. https://doi.org/10.1109/34.574797

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M.R., Kuemmerle, T., Meyfroidt, P., Mitchard, E.T.A., Reiche, J., Ryan, C.M., Waske, B., 2016. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. Remote Sens. 8, 1–23. https://doi.org/10.3390/rs8010070

Kandaswamy, U., Adjeroh, D.A., Lee, M.C., 2005. Efficient texture analysis of SAR imagery. IEEE Trans. Geosci. Remote Sens. 43, 2075–2083. https://doi.org/10.1109/TGRS.2005.852768

Karasiak, N., Dejoux, J.F., Fauvel, M., Willm, J., Monteil, C., Sheeren, D., 2019. Statistical stability and spatial instability in mapping forest tree species by comparing 9 years of satellite image time series. Remote Sens. 11. https://doi.org/10.3390/rs11212512

Karasiak, N., Dejoux, J.F., Monteil, C., Sheeren, D., 2022. Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing. Mach. Learn. 111, 2715–2740. https://doi.org/10.1007/s10994-021-05972-1

Kenduiywo, B.K., Bargiel, D., Soergel, U., 2018. Crop-type mapping from a sequence of Sentinel 1 images. Int. J. Remote Sens. 39, 6383–6404. https://doi.org/10.1080/01431161.2018.1460503

Khosravi, I., Safari, A., Homayouni, S., 2018. MSMD: maximum separability and minimum dependency feature selection for cropland classification from optical and radar data. Int. J. Remote Sens. 39, 2159–2176. https://doi.org/10.1080/01431161.2018.1425564

Klages, S., Heidecke, C., Osterburg, B., 2020. The impact of agricultural production and policy on water quality during the dry year 2018, a case study from Germany. Water (Switzerland) 12, 1–20. https://doi.org/10.3390/W12061519

Klonus, S., Ehlers, M., 2007. Image fusion using the Ehlers spectral characteristics preservation algorithm. GIScience Remote Sens. 44, 93–116. https://doi.org/10.2747/1548-1603.44.2.93

Kurosu, T., Uratsuka, S., Maeno, H., Kozu, T., 1999. Texture statistics for classification of land use with multitemporal JERS-1 SAR single-look imagery. IEEE Trans. Geosci. Remote Sens. 37, 227–235. https://doi.org/10.1109/36.739157

Kussul, N., Lavreniuk, M., Shelestov, A., Yailymov, B., 2016a. Along the season crop classification in Ukraine based on time series of optical and SAR images using ensemble of neural network classifiers, in: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). pp. 7145–7148. https://doi.org/10.1109/IGARSS.2016.7730864

Kussul, N., Lemoine, G., Gallego, F.J., Skakun, S. V, Lavreniuk, M., Shelestov, A.Y., 2016b. Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 2500–2508. https://doi.org/10.1109/JSTARS.2016.2560141

Kussul, N., Mykola, L., Shelestov, A., Skakun, S., 2018. 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. 51, 627–636. https://doi.org/10.1080/22797254.2018.1454265

Larrañaga, A., Álvarez-Mozos, J., Albizua, L., 2011. Crop classification in rain-fed and irrigated agricultural areas using Landsat TM and ALOS/PALSAR data. Can. J. Remote Sens. 37, 157–170. https://doi.org/10.5589/m11-022

Laso Bayas, J.C., Lesiv, M., Waldner, F., Schucknecht, A., Duerauer, M., See, L., Fritz, S., Fraisl, D., Moorthy, I., McCallum, I., Perger, C., Danylo, O., Defourny, P., Gallego, J., Gilliams, S., Akhtar, I.U.H., Baishya, S.J., Baruah, M., Bungnamei, K., Campos, A., Changkakati, T., Cipriani, A., Das, Krishna, Das, Keemee, Das, I., Davis, K.F., Hazarika, P., Johnson, B.A., Malek, Z., Molinari, M.E., Panging, K., Pawe, C.K., Pérez-Hoyos, A., Sahariah, P.K., Sahariah, D., Saikia, A., Saikia, M., Schlesinger, P., Seidacaru, E., Singha, K., Wilson, J.W., 2017. A global reference database of crowdsourced cropland data collected using the Geo-Wiki platform. Sci. Data 4, 1–10. https://doi.org/10.1038/sdata.2017.136

Le Hégarat-Mascle, S., Quesney, A., Vidal-Madjar, D., Taconet, O., Normand, M., Loumagne, C., 2000. Land cover discrimination from multitemporal ERS images and multispectral Landsat images: A study case in an agricultural area in France. Int. J. Remote Sens. 21, 435–456. https://doi.org/10.1080/014311600210678

Li, R.Y., Ulaby, F.T., Eyton, J.R., 1980. Crop Classification With a Landsat/Radar Sensor Combination. Proc. Soc. Photo-Optical Instrum. Eng. 78–87.

Liang, W., Abidi, M., Carrasco, L., McNelis, J., Tran, L., Li, Y., Grant, J., Liang, W., 2020. Mapping vegetation at species level with high-resolution multispectral and lidar data over a large spatial area: A case study with Kudzu. Remote Sens. 12. https://doi.org/10.3390/rs12040609

Lobo, A., Chic, O., Casterad, A., 1996. Classification of mediterranean crops with multisensor data: Per-pixel versus per-object statistics and image segmentation. Int. J. Remote Sens. 17, 2385–2400. https://doi.org/10.1080/01431169608948779

López-Andreu, F.J., López-Morales, J.A., Erena, M., Skarmeta, A.F., Martínez, J.A., 2022. Monitoring System for the Management of the Common Agricultural Policy Using Machine Learning and Remote Sensing. Electron. 11, 1–18. https://doi.org/10.3390/electronics11030325

Löw, F., Duveiller, G., 2014. Defining the spatial resolution requirements for crop identification using optical remote sensing. Remote Sens. 6, 9034–9063. https://doi.org/10.3390/rs6099034

Löw, F., Michel, U., Dech, S., Conrad, C., 2013. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using Support Vector Machines. ISPRS J. Photogramm. Remote Sens. 85, 102–119. https://doi.org/10.1016/j.isprsjprs.2013.08.007

Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. 28, 823–870. https://doi.org/10.1080/01431160600746456

Lucas, B., Pelletier, C., Schmidt, D., Webb, G.I., 2020. UNSUPERVISED

DOMAIN ADAPTATION TECHNIQUES FOR CLASSIFICATION OF SATELLITE IMAGE TIME SERIES Faculty of Information Technology , Monash University , Melbourne , Australia Universit ´ e Bretagne Sud , IRISA , UMR CNRS 6074 , Vannes , France 1074–1077.

Lucas, B., Pelletier, C., Schmidt, D., Webb, G.I., Petitjean, F., 2021. A Bayesian-inspired, deep learning-based, semi-supervised domain adaptation technique for land cover mapping, Machine Learning. Springer US. https://doi.org/10.1007/s10994-020-05942-z

Lussem, U., Hütt, C., Waldhoff, G., 2016. Combined analysis of Sentinel-1 and RapidEye data for improved crop type classification: An early season approach for rapeseed and cereals. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 41, 959–963. https://doi.org/10.5194/isprsarchives-XLI-B8-959-2016

Lüttger, A.B., Feike, T., 2018. Development of heat and drought related extreme weather events and their effect on winter wheat yields in Germany. Theor. Appl. Climatol. 132, 15–29. https://doi.org/10.1007/s00704-017-2076-y

Lyons, M.B., Keith, D.A., Phinn, S.R., Mason, T.J., Elith, J., 2018. A comparison of resampling methods for remote sensing classification and accuracy assessment. Remote Sens. Environ. 208, 145–153. https://doi.org/10.1016/j.rse.2018.02.026

Macholdt, J., Honermeier, B., 2017. Yield stability in winter wheat production: A survey on German farmers' and advisors' views. Agronomy 7. https://doi.org/10.3390/agronomy7030045

Mack, B., 2020. EO-BOX: Open Source Python Package [WWW Document]. URL https://github.com/benmack/eo-box (accessed 6.8.21).

Mack, B., 2019. eo-box [WWW Document]. GitHub.

Mansaray, L.R., Huang, W., Zhang, D., Huang, J., Li, J., 2017. Mapping rice fields in urban Shanghai, southeast China, using Sentinel-1A and Landsat 8 datasets. Remote Sens. 9. https://doi.org/10.3390/rs9030257

Maus, V., Câmara, G., Appel, M., Pebesma, E., 2019. dtwSat: Time-weighted dynamic time warping for satellite image time series analysis in R. J. Stat. Softw. 88. https://doi.org/10.18637/jss.v088.i05

Maxwell, A.E., Warner, T.A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: An applied review. Int. J. Remote Sens. 39, 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

McKinney, W., 2010. Data structures for statistical computing in python, in: Proceedings of the 9th Python in Science Conference. pp. 51–56.

McNairn, H., Brisco, B., 2004. The application of C-band polarimetric SAR for agriculture: A review. Can. J. Remote Sens. 30, 525–542. https://doi.org/10.5589/m03-069

McNairn, Heather, Champagne, C., Shang, J., 2007. The value of SAR multi-polarization data in delivering annual crop inventories. Int. Geosci. Remote Sens. Symp. 1397–1400. https://doi.org/10.1109/IGARSS.2007.4423067

McNairn, H, Champagne, C., Shang, J., 2007. The value of SAR Multi-polarization data in delivering annual crop inventories, in: 2007 IEEE

International Geoscience and Remote Sensing Symposium. pp. 1397–1400. https://doi.org/10.1109/IGARSS.2007.4423067

McNairn, H., Champagne, C., Shang, J., Holmstrom, D., Reichert, G., 2009a. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. ISPRS J. Photogramm. Remote Sens. 64, 434–449. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2008.07.006

McNairn, H., Champagne, C., Shang, J., Holmstrom, D., Reichert, G., 2009b. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. ISPRS J. Photogramm. Remote Sens. 64, 434–449. https://doi.org/10.1016/j.isprsjprs.2008.07.006

McNairn, H., Ellis, J., Van der Sanden, J.J., Hirose, T., Brown, R.J., 2002. Providing crop information using RADARSAT-1 and satellite optical imagery. Int. J. Remote Sens. 23, 851–870. https://doi.org/10.1080/01431160110070753

McNairn, H., Kross, A., Lapen, D., Caves, R., Shang, J., 2014. Early season monitoring of corn and soybeans with TerraSAR-X and RADARSAT-2. Int. J. Appl. Earth Obs. Geoinf. 28, 252–259. https://doi.org/10.1016/j.jag.2013.12.015

Meyer, H., Pebesma, E., 2020. Predicting into unknown space? Estimating the area of applicability of spatial prediction models 0–16.

Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., Nauss, T., 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environ. Model. Softw. 101, 1–9. https://doi.org/10.1016/j.envsoft.2017.12.001

Meyer, H., Reudenbach, C., Wöllauer, S., Nauss, T., 2019. Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction. Ecol. Modell. 411. https://doi.org/10.1016/j.ecolmodel.2019.108815

Michelson, D.B., Liljeberg, B.M., Pilesjö, P., 2000. Comparison of algorithms for classifying Swedish landcover using Landsat TM and ERS-1 SAR data. Remote Sens. Environ. 71, 1–15. https://doi.org/10.1016/S0034-4257(99)00024-3

Millard, K., Richardson, M., 2015. On the importance of training data sample selection in Random Forest image classification: A case study in peatland ecosystem mapping. Remote Sens. 7, 8489–8515. https://doi.org/10.3390/rs70708489

Mueller, L., Schindler, U., Graham Shepherd, T., Ball, B.C., Smolentseva, E., Pachikin, K., Hu, C., Hennings, V., Sheudshen, A.K., Behrendt, A., Eulenstein, F., Dannowski, R., 2014. The muencheberg soil quality rating for assessing the quality of global farmland. Environ. Sci. Eng. (Subseries Environ. Sci. 235–248. https://doi.org/10.1007/978-3-319-01017-5_13

Müller-Wilm, U., 2016. Sen2Cor Configuration and User Manual, Ref. S2-PDGS-MPC-L2A- SUM-V2.3. Eur. Sp. Agency.

Müller, H., Rufin, P., Griffiths, P., Barros Siqueira, A.J., Hostert, P., 2015. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. Remote Sens. Environ. 156, 490–499. https://doi.org/10.1016/j.rse.2014.10.014

NASA-JPL, 2013. NASA Shuttle Radar Topography Mission Global 1 arc second. 2013, distributed by NASA EOSDIS Land Processes DAAC [WWW Document]. URL https://doi.org/10.5067/MEaSUREs/SRTM/SRTMGL1.003. (accessed 6.9.21).

Nelson, A., Setiyono, T., Rala, A.B., Quicho, E.D., Raviz, J. V., Abonete, P.J., Maunahan, A.A., Garcia, C.A., Bhatti, H.Z.M., Villano, L.S., Thongbai, P., Holecz, F., Barbieri, M., Collivignarelli, F., Gatti, L., Quilang, E.J.P., Mabalay, M.R.O., Mabalot, P.E., Barroga, M.I., Bacong, A.P., Detoito, N.T., Berja, G.B., Varquez, F., Wahyunto, Kuntjoro, D., Murdiyati, S.R., Pazhanivelan, S., Kannan, P., Nirmala Mary, P.C., Subramanian, E., Rakwatin, P., Intrman, A., Setapayak, T., Lertna, S., Minh, V.Q., Tuan, V.Q., Duong, T.H., Quyen, N.H., Van Kham, D., Hin, S., Veasna, T., Yadav, M., Chin, C., Ninh, N.H., 2014. Towards an operational SAR-based rice monitoring system in Asia: Examples from 13 demonstration sites across Asia in the RIICE project. Remote Sens. 6, 10773–10812. https://doi.org/10.3390/rs61110773

Nikparvar, B., Thill, J.-C., 2021. Machine Learning of Spatial Data. ISPRS Int. J. Geo-Information. https://doi.org/10.3390/ijgi10090600

Nowakowski, A., Mrziglod, J., Spiller, D., Bonifacio, R., Ferrari, I., Mathieu, P.P., Garcia-Herranz, M., Kim, D.-H., 2021. Crop type mapping by using transfer learning. Int. J. Appl. Earth Obs. Geoinf. 98, 102313. https://doi.org/10.1016/j.jag.2021.102313

Nünez, J., Otazu, X., Fors, O., Prades, A., Palà, V., Arbiol, R., 1999. Multiresolution-based image fusion with additive wavelet decomposition. IEEE Trans. Geosci. Remote Sens. 37, 1204–1211. https://doi.org/10.1109/36.763274

Ok, A.O., Akyurek, Z., 2012. A segment-based approach to classify agricultural lands by using multi-temporal optical and microwave data. Int. J. Remote Sens. 33, 7184–7204. https://doi.org/10.1080/01431161.2012.700423

Okamoto, K., 1999. Estimation of rice-planted area in the tropical zone using a combination of optical and microwave satellite sensor data. Int. J. Remote Sens. 20, 1045–1048. https://doi.org/10.1080/014311699213091

Okamoto, K., Kawashima, H., 1999. Estimation of rice-planted area in the tropical zone using a combination of optical and microwave satellite sensor data. Int. J. Remote Sens. 20, 1045–1048. https://doi.org/10.1080/014311699213091

Oliphant, T.E., 2006. A guide to NumPy. Trelgol Publishing USA.

Onojeghuo, A.O., Blackburn, G.A., Wang, Q., Atkinson, P.M., Kindred, D., Miao, Y., 2018. Mapping paddy rice fields by applying machine learning algorithms to multi-temporal sentinel-1A and landsat data. Int. J. Remote Sens. 39, 1042–1067. https://doi.org/10.1080/01431161.2017.1395969

Orynbaikyzy, A., Gessner, U., Conrad, C., 2022. Spatial Transferability of Random Forest Models for Crop Type Classification Using Sentinel-1 and Sentinel-2. Remote Sens. 14, 1493. https://doi.org/10.3390/rs14061493

Orynbaikyzy, A., Gessner, U., Conrad, C., 2019. Crop type classification using a combination of optical and radar remote sensing data: a review. Int. J.

Remote Sens. 40, 6553–6595. https://doi.org/10.1080/01431161.2019.1569791

Orynbaikyzy, A., Gessner, U., Mack, B., Conrad, C., 2020. 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. 12. https://doi.org/10.3390/RS12172779

Overmars, K.P., De Koning, G.H.J., Veldkamp, A., 2003. Spatial autocorrelation in multi-scale land use models. Ecol. Modell. 164, 257–270. https://doi.org/10.1016/S0304-3800(03)00070-X

Park, S., Im, J., 2016. Classification of croplands through fusion of optical and sar time series data. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 41, 703–704. https://doi.org/10.5194/isprsarchives-XLI-B7-703-2016

Park, Seonyoung, Im, J., Park, Seohui, Yoo, C., Han, H., Rhee, J., 2018. Classification and mapping of paddy rice by combining Landsat and SAR time series data. Remote Sens. 10, 1–22. https://doi.org/10.3390/rs10030447

Parks, S.M., 2012. Synthetic Aperture Radar (Sar) and Optical Imagery Data Fusion: Crop Yield Analysis in Southeast Asia. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXIX-B7, 403–406. https://doi.org/10.5194/isprsarchives-xxxix-b7-403-2012

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Pinheiro, A., Carrão, H., Caetano, M., 2007. Evaluation of ASAR and optical data synergy for high resolution land cover mapping in Portugal. Int. Geosci. Remote Sens. Symp. 1517–1520. https://doi.org/10.1109/IGARSS.2007.4423097

Ploton, P., Mortier, F., Réjou-méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-fleury, S., Pélissier, R., 2020. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nat. Commun. 1–11. https://doi.org/10.1038/s41467-020-18321-y

Pohjankukka, J., Pahikkala, T., Nevalainen, P., Heikkonen, J., 2017. Estimating the prediction performance of spatial models via spatial k-fold cross validation. Int. J. Geogr. Inf. Sci. 31, 2001–2019. https://doi.org/10.1080/13658816.2017.1346255

Pohl, C., 2016. Multisensor image fusion guidelines in remote sensing. IOP Conf. Ser. Earth Environ. Sci. 34. https://doi.org/10.1088/1755-1315/34/1/012026

Pohl, C., van Genderen, J., 2015. Structuring contemporary remote sensing image fusion. Int. J. Image Data Fusion 6, 3–21. https://doi.org/10.1080/19479832.2014.998727

Pohl, C., Van Genderen, J., 2016. Remote sensing image fusion: A practical guide. Crc Press.

Pohl, C., Van Genderen, J.L., 1998. Review article Multisensor image fusion

in remote sensing: Concepts, methods and applications, International Journal of Remote Sensing. https://doi.org/10.1080/014311698215748

Preidl, S., Lange, M., Doktor, D., 2020. Introducing APiC for regionalised land cover mapping on the national scale using Sentinel-2A imagery. Remote Sens. Environ. 240, 111673. https://doi.org/10.1016/j.rse.2020.111673

Presutti, M.E., Franklin, S.E., Moskal, L.M., Dickson, E.E., 2001. Supervised Classification of Multisource Satellite Image Spectral and Texture Data for Agricultural Crop Mapping in Buenos Aires Province, Argentina. Can. J. Remote Sens. 27, 679–684. https://doi.org/10.1080/07038992.2001.10854910

Prins, A.J., Van Niekerk, A., 2020. Crop type mapping using LiDAR, Sentinel-2 and aerial imagery with machine learning algorithms. Geo-Spatial Inf. Sci. 24, 1–13. https://doi.org/10.1080/10095020.2020.1782776

Qi, J., Wang, C., Inoue, Y., Zhang, R., Gao, W., 2004. Synergy of optical and radar remote sensing in agricultural applications. Proc. SPIE - Int. Soc. Opt. Eng. 5153. https://doi.org/10.1117/12.514562

Qi, J.G., Wang, C.Z., Inoue, Y., Zhang, R.F., Gao, W., Cao, G.Z., 2004. Synergy of optical and radar remote sensing in agricultural applications. Dianbo Kexue Xuebao/Chinese J. Radio Sci. 19, 399–404. https://doi.org/10.1117/12.514562

Qiao, C., Daneshfar, B., Davidson, A., Jarvis, I., Liu, T., Fisette, T., 2014. Integration of optical and polarimetric SAR imagery for locally accurate crop classification. pp. 1485–1488. https://doi.org/10.1109/IGARSS.2014.6946718

Qiao, C., Daneshfar, B., Davidson, A.M., 2017. The application of discriminant analysis for mapping cereals and pasture using object-based features. Int. J. Remote Sens. 38, 5546–5568. https://doi.org/10.1080/01431161.2017.1325530

Raghavswamy, V., Gautam, N.C., Padmavathi, M., Badarinath, K.V.S., 1996. Studies on microwave remote sensing data in conjunction with optical data for land use/land cover mapping and assessment. Geocarto Int. 11, 25–31. https://doi.org/10.1080/10106049609354558

Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., 2006. Are niche-based species distribution models transferable in space? J. Biogeogr. 33, 1689–1703. https://doi.org/10.1111/j.1365-2699.2006.01466.x

Raschka, S., 2018. MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack. J. Open Source Softw. 3. https://doi.org/10.21105/joss.00638

Reinermann, S., Gessner, U., Asam, S., Kuenzer, C., Dech, S., 2019. The effect of droughts on vegetation condition in Germany: An analysis based on two decades of satellite earth observation time series and crop yield statistics. Remote Sens. 11. https://doi.org/10.3390/rs11151783

Richards, J.A., 2005. Analysis of remotely sensed data: The formative decades and the future. IEEE Trans. Geosci. Remote Sens. 43, 422–432. https://doi.org/10.1109/TGRS.2004.837326

Richards, J.A., 1999. Remote sensing digital image analysis. Springer.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton,

D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography (Cop.). 40, 913–929. https://doi.org/10.1111/ecog.02881

Rocha, A.D., Groen, T.A., Skidmore, A.K., Darvishzadeh, R., Willemen, L., 2018. Machine learning using hyperspectral data inaccurately predicts plant traits under spatial dependency. Remote Sens. 10. https://doi.org/10.3390/rs10081263

Roscher, R., Bohn, B., Duarte, M.F., Garcke, J., 2020. Explainable Machine Learning for Scientific Insights and Discoveries. IEEE Access 8, 42200–42216. https://doi.org/10.1109/ACCESS.2020.2976199

Saeys, Y., Inza, I., Larrañaga, P., 2007. A review of feature selection techniques in bioinformatics. Bioinformatics 23, 2507–2517. https://doi.org/10.1093/bioinformatics/btm344

Salehi, B., Daneshfar, B., Davidson, A.M., 2017. Accurate crop-type classification using multi-temporal optical and multi-polarization SAR data in an object-based image analysis framework. Int. J. Remote Sens. 38, 4130–4155. https://doi.org/10.1080/01431161.2017.1317933

Sandholt, I., 2001. The combination of polarimetric SAR with satellite SAR and optical data for classification of agricultural land. Geogr. Tidsskr. J. Geogr. 101, 21–32. https://doi.org/10.1080/00167223.2001.10649448

Sanli, F.B., Abdikan, S., Esetlili, M.T., Sunar, F., 2017. Evaluation of image fusion methods using PALSAR, RADARSAT-1 and SPOT images for land use/ land cover classification. J. Indian Soc. Remote Sens. 45, 591–601. https://doi.org/10.1007/s12524-016-0625-y

Schistad Solberg, A.H., Taxt, T., Jain, A.K., 1996. A markov random field model for classification of multisource satellite imagery. IEEE Trans. Geosci. Remote Sens. 34, 100–113. https://doi.org/10.1109/36.481897

Schmedtmann, J., Campagnolo, M.L., 2015. Reliable crop identification with satellite imagery in the context of Common Agriculture Policy subsidy control. Remote Sens. 7, 9325–9346. https://doi.org/10.3390/rs70709325

Schmitt, M., Zhu, X.X., 2016. Data Fusion and Remote Sensing: An ever-growing relationship. IEEE Geosci. Remote Sens. Mag. 4, 6–23. https://doi.org/10.1109/MGRS.2016.2561021

Schowengerdt, R.A., 1980. Reconstruction of multispatial, multispectral image data using spatial frequency content. Photogramm. Eng. Remote Sens. 46, 1325–1334.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., Brenning, A., 2018. Performance evaluation and hyperparameter tuning of statistical and machine-learning models using spatial data. https://doi.org/10.1016/j.ecolmodel.2019.06.002

Shackelford, A.K., Davis, C.H., 2003. A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas. IEEE Trans. Geosci. Remote Sens. 41, 1920–1932. https://doi.org/10.1109/TGRS.2003.814627

Shang, J., McNairn, H., Champagne, C., Jiao, X., 2008. Contribution of Multi-Frequency, Multi-Sensor, and Multi-Temporal Radar Data to Operational Annual Crop Mapping, in: IGARSS 2008 - 2008 IEEE

136

International Geoscience and Remote Sensing Symposium. pp. III-378-III–381. https://doi.org/10.1109/IGARSS.2008.4779362

Sheoran, A., Haack, B., 2013. Classification of California agriculture using quad polarization radar data and Landsat Thematic Mapper data. GIScience Remote Sens. 50, 50–63. https://doi.org/10.1080/15481603.2013.778555

Sitokonstantinou, V., Papoutsis, I., Kontoes, C., Arnal, A.L., Andrés, A.P.A., Zurbano, J.A.G., 2018. Scalable parcel-based crop identification scheme using Sentinel-2 data time-series for the monitoring of the common agricultural policy. Remote Sens. 10. https://doi.org/10.3390/rs10060911

Skakun, S., Kussul, N., Yu Shelestov, A., Lavreniuk, M., Kussul, O., 2016. Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine " Stimulating Innovation for Global Monitoring of Agriculture and its Impact on the Environment . Ieee J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 3712–3719.

Solberg, A.H., Jain, A.K., Taxt, T., 1994. Multisource Classification of Remotely Sensed Data: Fusion of Landsat TM and SAR Images. IEEE Trans. Geosci. Remote Sens. 32, 768–778. https://doi.org/10.1109/36.298006

Solberg, A.H.S., 2006. Data fusion for remote-sensing applications. Signal image Process. Remote Sens. 249–271.

Song, X.-P., Huang, W., Hansen, M.C., Potapov, P., 2021. An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. Sci. Remote Sens. 3, 100018. https://doi.org/10.1016/j.srs.2021.100018

Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., Mochizuki, K. ichiro, 2017. Assessing the suitability of data from Sentinel-1A and 2A for crop classification. GIScience Remote Sens. 54, 918–938. https://doi.org/10.1080/15481603.2017.1351149

Soria-Ruiz, J., Fernandez-Ordonez, Y., McNairm, H., Bugden-Storie, J., 2007. Corn monitoring and crop yield using optical and RADARSAT-2 images. Int. Geosci. Remote Sens. Symp. 3655–3658. https://doi.org/10.1109/IGARSS.2007.4423638

Soria-Ruiz, J., Fernandez-Ordońez, Y., Woodhouse, I.H., 2010. Land-cover classification using radar and optical images: A case study in Central Mexico. Int. J. Remote Sens. 31, 3291–3305. https://doi.org/10.1080/01431160903160777

Steele-Dunne et al., S., 2017. Radar Remote Sensing of Agriciltural Canopies. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sen 10, 1–25.

Stehman, S. V., Foody, G.M., 2019. Key issues in rigorous accuracy assessment of land cover products. Remote Sens. Environ. 231, 111199. https://doi.org/10.1016/j.rse.2019.05.018

Stoian, A., Poulain, V., Inglada, J., Poughon, V., Derksen, D., 2019. Land cover maps production with high resolution satellite image time series and convolutional neural networks: Adaptations and limits for operational systems. Remote Sens. 11, 1–26. https://doi.org/10.3390/rs11171986

Sukawattanavijit, C., Chen, J., 2015. Fusion of multi-frequency SAR data with THAICHOTE optical imagery for maize classification in Thailand. pp.

617–620. https://doi.org/10.1109/IGARSS.2015.7325839

Sulik, J.J., Long, D.S., 2015. Spectral indices for yellow canola flowers. Int. J. Remote Sens. 36, 2751–2765. https://doi.org/10.1080/01431161.2015.1047994

Sun, C., Bian, Y., Zhou, T., Pan, J., 2019. Using of multi-source and multi-temporal remote sensing data improves crop-type mapping in the subtropical agriculture region. Sensors (Switzerland) 19, 1–23. https://doi.org/10.3390/s19102401

Sun, L., Chen, J., Guo, S., Deng, X., Han, Y., 2020. Integration of time series sentinel-1 and sentinel-2 imagery for crop type mapping over oasis agricultural areas. Remote Sens. 12. https://doi.org/10.3390/RS12010158

Sun, W., Heidt, V., Gong, P., Xu, G., 2003. Information fusion for rural land-use classification with high-resolution satellite imagery. IEEE Trans. Geosci. Remote Sens. 41, 883–890. https://doi.org/10.1109/TGRS.2003.810707

Tang, F., Ishwaran, H., 2017. Random Forest Missing Data Algorithms. Physiol. Behav. 176, 139–148. https://doi.org/10.1016/j.physbeh.2017.03.040

Tardy, B., Inglada, J., Michel, J., 2017. Fusion approaches for land cover map production using high resolution image time series without reference data of the corresponding period. Remote Sens. 9. https://doi.org/10.3390/rs9111151

Taxt, T., Schistad Solberg, A.H., 1997. Information fusion in remote sensing. Vistas Astron. 41, 337–342. https://doi.org/10.1016/S0083-6656(97)00036-6

Tetteh, G.O., Gocht, A., Erasmi, S., Schwieder, M., Conrad, C., 2021. Evaluation of Sentinel-1 and Sentinel-2 feature sets for delineating agricultural fields in heterogeneous landscapes. IEEE Access 9, 116702–116719. https://doi.org/10.1109/ACCESS.2021.3105903

Thenkabail, P.S., Schull, M., Turral, H., 2005. Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data. Remote Sens. Environ. 95, 317–341. https://doi.org/10.1016/j.rse.2004.12.018

Thomas, I.L., Ching, N.P., Benning, V.M., D'Aguanno, J.A., 1987. A review of multi-channel indices of class separability. Int. J. Remote Sens. 8, 331–350. https://doi.org/10.1080/01431168708948645

Tobler, W.R., 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. Econ. Geogr. 46, 234. https://doi.org/10.2307/143141

Torbick, N., Chowdhury, D., Salas, W., Qi, J., 2017. Monitoring rice agriculture across myanmar using time series Sentinel-1 assisted by Landsat-8 and PALSAR-2. Remote Sens. 9. https://doi.org/10.3390/rs90201019

Torbick, N., Salas, W., Xiao, X., Ingraham, P., Fearon, M.G., Biradar, C., Zhao, D., Liu, Y., Li, P., Zhao, Y., 2011. Integrating SAR and optical imagery for regional mapping of paddy rice attributes in the Poyang Lake Watershed, China. Can. J. Remote Sens. 37, 17–26. https://doi.org/10.5589/m11-020

Touzi, R., Lopes, A., Bruniquel, J., Vachon, P.W., 1999. Coherence estimation for SAR imagery. IEEE Trans. Geosci. Remote Sens. 37, 135–149.

138

https://doi.org/10.1109/36.739146

Tuia, D., Persello, C., Bruzzone, L., 2016. Domain adaptation for the classification of remote sensing data: An overview of recent advances. IEEE Geosci. Remote Sens. Mag. 4, 41–57. https://doi.org/10.1109/MGRS.2016.2548504

Ulaby, F.T., Kouyate, F., Brisco, B., Lee Williams, T.H., 1986. Textural Information in SAR Images. IEEE Trans. Geosci. Remote Sens. GE-24, 235–245. https://doi.org/10.1109/TGRS.1986.289643

Ulaby, F.T., Li, R.Y., Shanmugan, K.S., 1982. Crop Classification Using Airborne Radar and Landsat Data. IEEE Trans. Geosci. Remote Sens. 20, 42–51. https://doi.org/10.1109/TGRS.1982.4307519

UN, 2017. World Population Prospects: 2017 Revision.

Van Der Meer, F., 1997. What does multisensor image fusion add in terms of information content for visual interpretation? Int. J. Remote Sens. 18, 445–452. https://doi.org/10.1080/014311697219187

Van Genderen, J.L., Pohl, C., 1994. Image fusion: Issues, techniques and applications.

Van Tricht, K., Gobin, A., Gilliams, S., Piccard, I., 2018. Synergistic use of radar sentinel-1 and optical sentinel-2 imagery for crop mapping: A case study for Belgium. Remote Sens. 10. https://doi.org/10.3390/rs10101642

Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.F., Ceschia, E., 2017. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. Remote Sens. Environ. 199, 415–426. https://doi.org/10.1016/j.rse.2017.07.015

Venter, Z.S., Sydenham, M.A.K., 2021. Continental-scale land cover mapping at 10 m resolution over europe (Elc10). Remote Sens. 13. https://doi.org/10.3390/rs13122301

Vescovi, F.D., Gomarasca, M.A., 1999. Integration of Optical and Microwave Remote Sensing Data for Agricultural Land Use Classification. Environ. Monit. Assess. 58, 133–149. https://doi.org/10.1023/A:1006047906601

Villa, P., Stroppiana, D., Fontanelli, G., Azar, R., Brivio, P.A., 2015. In-season mapping of crop type with optical and X-band SAR data: A classification tree approach using synoptic seasonal features. Remote Sens. 7, 12859–12886. https://doi.org/10.3390/rs71012859

Vrabel, J., 1996. Multispectral imagery band sharpening study. Photogramm. Eng. Remote Sensing 62, 1075–1083.

Wadoux, A.M.J.C., Heuvelink, G.B.M., de Bruin, S., Brus, D.J., 2021. Spatial cross-validation is not the right way to evaluate map accuracy. Ecol. Modell. 457. https://doi.org/10.1016/j.ecolmodel.2021.109692

Wald, L., 1999. Some terms of reference in data fusion. IEEE Trans. Geosci. Remote Sens. 37, 1190–1193. https://doi.org/10.1109/36.763269

Waldner, F., Chen, Y., Lawes, R., Hochman, Z., 2019. Needle in a haystack: Mapping rare and infrequent crops using satellite imagery and data balancing methods. Remote Sens. Environ. 233, 111375. https://doi.org/10.1016/j.rse.2019.111375

Wang, J., Xiao, X., Qin, Y., Dong, J., Zhang, G., Kou, W., Jin, C., Zhou, Y.,

Zhang, Y., 2015. Mapping paddy rice planting area in wheat-rice double-cropped areas through integration of Landsat-8 OLI, MODIS, and PALSAR images. Sci. Rep. 5, 1–11. https://doi.org/10.1038/srep10088

Waske, B., Menz, G., Benediktsson, J.A., 2007. Fusion of support vector machines for classifying SAR and multispectral imagery from agricultural areas. Int. Geosci. Remote Sens. Symp. 4842–4845. https://doi.org/10.1109/IGARSS.2007.4423945

Waske, B., Van Der Linden, S., 2008. Classifying multilevel imagery from SAR and optical sensors by decision fusion. IEEE Trans. Geosci. Remote Sens. 46, 1457–1466. https://doi.org/10.1109/TGRS.2008.916089

Weiss, M., Jacob, F., Duveiller, G., 2020. Remote sensing for agricultural applications: A meta-review. Remote Sens. Environ. 236, 111402. https://doi.org/10.1016/j.rse.2019.111402

Wenger, S.J., Olden, J.D., 2012. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. Methods Ecol. Evol. 3, 260–267. https://doi.org/10.1111/j.2041-210X.2011.00170.x

Whitcraft, A.K., Becker-Reshef, I., Justice, C.O., Gifford, L., Kavvada, A., Jarvis, I., 2019. No pixel left behind: Toward integrating Earth Observations for agriculture into the United Nations Sustainable Development Goals framework. Remote Sens. Environ. 235, 111470. https://doi.org/10.1016/j.rse.2019.111470

White, J.C., Wulder, M.A., Hobart, G.W., Luther, J.E., Hermosilla, T., Griffiths, P., Coops, N.C., Hall, R.J., Hostert, P., Dyk, A., Guindon, L., 2014. Pixel-based image compositing for large-area dense time series applications and science. Can. J. Remote Sens. 40, 192–212. https://doi.org/10.1080/07038992.2014.945827

Wizemann, H.D., Ingwersen, J., Högy, P., Warrach-Sagi, K., Streck, T., Wulfmeyer, V., 2014. Three year observations of water vapor and energy fluxes over agricultural crops in two regional climates of Southwest Germany. Meteorol. Zeitschrift 24, 39–59. https://doi.org/10.1127/metz/2014/0618

Woźniak, E., Rybicki, M., Kofman, W., Aleksandrowicz, S., Wojtkowski, C., Lewiński, S., Bojanowski, J., Musiał, J., Milewski, T., Slesiński, P., Łączyński, A., 2022. Multi-temporal phenological indices derived from time series Sentinel-1 images to country-wide crop classification. Int. J. Appl. Earth Obs. Geoinf. 107. https://doi.org/10.1016/j.jag.2022.102683

Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 122, 2–10. https://doi.org/10.1016/j.rse.2012.01.010

Xu, W., Wu, B., Tian, Y., Huang, J., Zhang, Y., 2004. Synergy of multitemporal radarsat SAR and landsat ETM data for extracting agricultural crops structure. Int. Geosci. Remote Sens. Symp. 6, 4073–4076. https://doi.org/10.1109/igarss.2004.1370026

Yates, K.L., Bouchet, P.J., Caley, M.J., Mengersen, K., Randin, C.F., Parnell, S., Fielding, A.H., Bamford, A.J., Ban, S., Barbosa, A.M., Dormann, C.F., Elith, J., Embling, C.B., Ervin, G.N., Fisher, R., Gould, S., Graf, R.F., Gregr, E.J., Halpin, P.N., Heikkinen, R.K., Heinänen, S., Jones, A.R., Krishnakumar, P.K., Lauria, V., Lozano-Montes, H., Mannocci, L., Mellin, C., Mesgaran, M.B., Moreno-Amat, E., Mormede, S., Novaczek,

E., Oppel, S., Ortuño Crespo, G., Peterson, A.T., Rapacciuolo, G., Roberts, J.J., Ross, R.E., Scales, K.L., Schoeman, D., Snelgrove, P., Sundblad, G., Thuiller, W., Torres, L.G., Verbruggen, H., Wang, L., Wenger, S., Whittingham, M.J., Zharikov, Y., Zurell, D., Sequeira, A.M.M., 2018. Outstanding Challenges in the Transferability of Ecological Models. Trends Ecol. Evol. 33, 790–802. https://doi.org/10.1016/j.tree.2018.08.001

Yu, L., Fu, H., Wu, B., Clinton, N., Gong, P., 2016. Exploring the potential role of feature selection in global land-cover mapping. Int. J. Remote Sens. 37, 5491–5504. https://doi.org/10.1080/01431161.2016.1244365

Zeng, Y., Zhang, J., Van Genderen, J.L., 2006. Comparison and analysis of remote sensing data fusion techniques at feature and decision levels, in: ISPRS Commission VII Mid-Term Symp., Remote Sens.: From Pixels to Process.

Zhang, J., 2010. Multi-source remote sensing data fusion: Status and trends. Int. J. Image Data Fusion 1, 5–24. https://doi.org/10.1080/19479830903561035

Zhong, L., Gong, P., Biging, G.S., 2014. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. Remote Sens. Environ. 140, 1–13. https://doi.org/10.1016/j.rse.2013.08.023

Zhong, L., Hu, L., Yu, L., Gong, P., Biging, G.S., 2016. Automated mapping of soybean and corn using phenology. ISPRS J. Photogramm. Remote Sens. 119, 151–164. https://doi.org/10.1016/j.isprsjprs.2016.05.014

Zhou, T., Pan, J., Zhang, P., Wei, S., Han, T., 2017. Mapping winter wheat with multi-temporal SAR and optical images in an urban agricultural region. Sensors (Switzerland) 17, 1–16. https://doi.org/10.3390/s17061210

Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. Remote Sens. Environ. 159, 269–277. https://doi.org/10.1016/j.rse.2014.12.014

APPENDIX

A.1 LIST OF PUBLICATIONS

Peer-Reviewed Journals

- Orynbaikyzy, A., Gessner, U., Conrad, C., 2022. Spatial Transferability of Random Forest Models for Crop Type Classification Using Sentinel-1 and Sentinel-2. Remote Sens. 14, 1493. https://doi.org/10.3390/rs14061493

- Orynbaikyzy, A., Gessner, U., Conrad, C., 2019. Crop type classification using a combination of optical and radar remote sensing data: a review. Int. J. Remote Sens. 40, 6553–6595. https://doi.org/10.1080/01431161.2019.1569791

- Orynbaikyzy, A., Gessner, U., Mack, B., Conrad, C., 2020. 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. 12. https://doi.org/10.3390/RS12172779

- Orynbaikyzy, A., Plank, S., Vetrita, Y., Martinis, S., Santoso, I., 2023. Joint use of Sentinel-2 and Sentinel-1 data for rapid mapping of volcanic eruption deposits in Southeast Asia. Int. J. Appl. Earth Obs. Geoinf. 116, 103166. https://doi.org/10.1016/j.jag.2022.103166

- Zorn, E.U., Orynbaikyzy, A., Plank, S., Babeyko, A., Darmawan, H., Robbany, I.F., 2022. Identification and ranking of subaerial volcanic tsunami hazard sources in Southeast Asia. Nat. Hazards Earth Syst. Sci. 3083–3104. https://doi.org/10.5194/nhess-22-3083-2022

Conference Proceedings as a First Author

- Orynbaikyzy, A., Gessner, U., Conrad, C., 2019. Investigating Dense Time Series of Optical and SAR Earth Observation data for crop type classification. EARSeL Symposium 2019, 1-4 July, 2019, Salzburg, Austria

- Orynbaikyzy, A., Gessner, U., Mack, B., Conrad, C., 2019. Optimal combination of optical and SAR features for a crop type classification. Living Planet Symposium. 13-17 May, 2019, Milan, Italy

- Orynbaikyzy, A., Plank, S. M., Martinis, S., Tjahjaningsih, A., Ismanto, R. D., Chusnayah, F., Santoso, I., Suwarsono, Khomarudin, M. R., Vetrita, Y., 2022. SAR-based Automated

Unsupervised Rapid Mapping of Land Surface Changes after Volcanic Eruptions in Indonesia. EARSeL Symposium 2022, 13-16 Sep 2022, Paphos, Cyprus

A.2 CURRICULUM VITAE

Work Experience

- 2021 July – Present: **Research Associate.** Department of Geo-Risks and Civil Security, German Remote Sensing Data Center (DFD), German Aerospace Center (DLR). Wessling, Germany.
- 2017 November – 2020 October: **Doctoral Student.** Department of Land Surface Department, German Remote Sensing Data Center (DFD), German Aerospace Center (DLR). Wessling, Germany.
- 2017 January – 2017 October: **GIS Analyst.** The Republican Centre for State Urban Planning and Cadastre. Astana, Kazakhstan.
- 2016 July – 2016 September: **Research Intern.** Interfaculty Department of Geoinformatics (Z_GIS), Paris Lodron University of Salzburg. Salzburg, Austria.

Education

- 2019 September – Present: **Doctoral Candidate.** Institute of Geosciences and Geography, Martin Luther University Halle-Wittenberg. Halle, Germany.
- 2014 October – 2014 September: **Master of Science in "Applied Geoinformatics".** Interfaculty department of Geoinformatics (Z_GIS), Paris Lodron University of Salzburg. Salzburg, Austria.
- 2010 September – 2014 June: **Bachelor in "Cartography and Geodesy".** Institute of Cartography and Geoinformatics, Al-Farabi Kazakh National University. Almaty, Kazakhstan.

A.3 REPRINTING PERMISSIONS OF THE PUBLISHERS

MDPI Open Access Information and Policy

All articles published by MDPI are made immediately available worldwide under an open access license. This means:

- everyone has free and unlimited access to the full-text of *all* articles published in MDPI journals;

- everyone is free to re-use the published material if proper accreditation/citation of the original publication is given;

- open access publication is supported by the authors' institutes or research funding agencies by payment of a comparatively low **Article Processing Charge (APC)** for accepted articles.

Permissions

Eidesstattliche Erklärung / Declaration under Oath

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

_____   _____

Datum / Date   Unterschrift des Antragstellers /Signature of the applicant