# Phenological and spectral optimisation of multi-temporal land use classification

## 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

Herrn Gerstmann, Henning
geb. am 27. Mai 1989  in Berlin

Gutachter:
1.  Prof. Dr. Cornelia Gläßer
2.  Prof. Dr. Birgit Kleinschmit

Tag der Verteidigung:  Halle (Saale), 20. Dezember 2018

# PHENOLOGICAL AND SPECTRAL OPTIMISATION OF MULTI-TEMPORAL LAND USE CLASSIFICATION

Using plant phenological phases as indicators for temporal windows of high spectral separability of agricultural crops

HENNING GERSTMANN

June 2018

# ACKNOWLEDGMENTS

# ABSTRACT

Global change and the continuously growing global population challenge scientists and political decision-makers. These challenges require efficient management tools for yield forecasting, estimation of usable water resources or to administratively regulate crop cultivation via agricultural subsidies.

These tools require highly accurate and timely information about current land use and land use change. Especially in agricultural areas, these data can often be produced only with a large effort, because ground surveys are very cost- and labour-intensive or limited by political crises and inefficient administrative structures. Monitoring initiatives based on remote sensing data, which classify crop types according to their different spectral reflectance characteristics, provide valuable contributions to such mapping tasks. The increasing temporal density of available satellite data of rising geometrical and spectral resolution can exceed computational limits for large-scale land use classifications. Contrarily, persistent periods of cloud coverage are an issue in many regions of the world, which can strongly reduce the amount of available and usable data sets. Inaccurate land use data hamper the reliability and accuracy of the respective models and hence decrease public acceptance of these tools.

As shown in recent studies, specifically adjusted vegetation indices and plant phenological phases are effective indicators to address the negative effects of temporal data gaps or to discriminate spectrally highly similar crop types. Reflectance properties of crops are species-specific, which results in temporal windows of different spectral separability. These temporal windows can be employed for the selection of decisive acquisition dates and spectral regions to prevent computational problems.

In this study, a methodology to optimise spectral features automated and to detect stable patterns of their spectral separability was developed. These patterns are systematically linked to phenological phases which are understood as indicators for time frames of maximal spectral separability of crops. The thesis on hand shows that the application of the phenological indicator time frames to select the most suitable acquisition dates and the optimisation of vegetation indices can increase land use classification accuracy. As a consequence, ground surveying effort can be reduced. Phenological indicator time frames could be derived for the dominating crop types in Germany. Since the methodology is flexibly applicable to other crop types and study regions, it can enhance national or other large-scale remote sensing based monitoring initiatives.

v

## ZUSAMMENFASSUNG

Der globale Wandel und die stetig wachsende Weltbevölkerung stellen Akteure aus Wissenschaft und Politik vor neue Herausforderungen. Dies erfordert effiziente Verfahren zur Prognose landwirtschaftlicher Erträge und Wasserressourcen oder zur behördlichen Steuerung des Feldfruchtanbaus durch Agrarsubventionen.

Solche Werkzeuge benötigen genaue, hochaktuelle Landnutzungsdaten. Diese sind oftmals nur schwer zu erhalten, da der Erhebungsaufwand kosten- und arbeitsintensiv ist oder durch politische Krisen und ineffiziente Verwaltungen erhöht wird. Fernerkundungsbasiertes Monitoring, welches unterschiedliche spektrale Reflexionseigenschaften zur Unterscheidung nutzt, leistet hier einen wichtigen Beitrag zur Kartierung. Durch die zunehmende temporale Dichte von Satellitendaten immer höherer geometrischer und spektraler Auflösung können Klassifikationen großer Flächen an Rechenkapazitätsgrenzen stoßen. Zusätzlich wird die Verfügbarkeit nutzbarer Datensätze durch längere Zeiträume starker Wolkenbedeckung limitiert. Unzureichend genaue Landnutzungsdaten erschweren jedoch Aussagekraft, Modellgüte und Akzeptanz der genannten Werkzeuge.

Aktuelle Forschungsergebnisse konnten zeigen, dass angepasste spektrale Maße und pflanzenphänologische Phasen dabei Werkzeuge sind, um die negativen Effekte von Datenlücken zu mindern oder spektral ähnliche Fruchtarten zu unterscheiden. Fruchtarten weisen phasenabhängige Reflexionseigenschaften auf, wodurch Zeitfenster unterschiedlich hoher Trennbarkeit entstehen. Diese können zur Bestimmung entscheidender Aufnahmezeitpunkte oder Spektralbereiche verwendet werden, um rechentechnischen Problemen vorzubeugen. In dieser Arbeit wurde eine Methode entwickelt, um spektrale Maße automatisiert zu optimieren und dabei zeitlich stabile Muster spektraler Trennbarkeit zu detektieren. Die spektralen Trennbarkeitsprofile werden systematisch phänologischen Phasen zugeordnet, die anschließend als Indikatoren für Zeitfenster maximale Trennbarkeit verschiedener Feldfrüchte betrachtet werden können.

Die vorliegende Arbeit zeigt, dass die Anwendung von phänologischen Indikatorphasen zur Auswahl der möglichst aussagekräftigen Satellitendaten zusammen mit der Optimierung von Vegetationsindizes die Genauigkeiten von Landnutzungsklassifikationen erhöhen kann. Dabei konnten Trennbarkeitszeitfenster für die dominierenden Feldfruchtarten in Deutschland ausgewiesen werden. Da die Methode flexibel auf andere Fruchtarten und Untersuchungsgebiete anwendbar ist, kann sie einen wertvollen Beitrag zu überregionalen Monitoringinitiativen leisten.

# CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# LIST OF ABBREVIATIONS

BBCH      Biologische Bundesanstalt, Bundessortenamt und
          Chemische Industrie

EU        European Union

EC        European Commission

ESA       European Space Agency

DEM       Digital Elevation Model

DEMMIN    Durable Environmental Multidisciplinary Monitoring
          Information Network

DLR       Deutsches Zentrum für Luft- und Raumfahrt e.V.

DOY       Day of Year

DWD       German Weather Service / German: *Deutscher Wetterdienst*

EVI       Enhanced Vegetation Index

LPIS      Land Parcel Identification System

LULC      Land Use / Land Cover

MODIS     Moderate Resolution Imaging Spectroradiometer

NDVI      Normalized Difference Vegetation Index

NIR       near infrared

PBC       phenology-based classification

PHASE     Phenological model for Application in Spatial and
          Environmental sciences

RESA      RapidEye Science Archive

SAR       Synthetic Aperture Radar

SRTM      Shuttle Radar Topography Mission

SWIR      short wave infrared

TERENO    Terrestrial Environmental Observatories

WDRVI     Wide-dynamic range vegetation index

VI        Vegetation index

# INTRODUCTION

The following sections briefly describe the outline of this thesis. The motivation for this research (Section 1.1), which leads to the determines the overall aims and scope of this thesis (Section 1.2), is presented along with its general structure (Section 1.3).

## 1.1 MOTIVATION

Accurate information on Land Use / Land Cover (LULC) gains increasing importance in several contexts such as food and water security for a growing global population or monitoring of phenomena related to global and climate change (e.g. Feddema et al. 2005). The LULC is traditionally assessed by ground surveys, that can naturally only focus on an area of limited spatial size or by reporting obligations that are costly and error-prone. The analysis of remote sensing imagery became a powerful tool to provide these information at moderate costs and larger scales (Congalton et al. 2014) up to continental and global products. The spectrum of applications that require these information ranges from catchment hydrological modelling, estimation of carbon storage for climate change modelling to crop yield estimation for food and biofuel security assessment and control mechanisms for area-based subsidies (Blaes et al. 2005).

The suitability of remote sensing methods for these topics is primarily due to its independence of administrative borders, continuous data supply and (almost) global coverage. This is facilitated by the continuously increasing number of satellites that have been launched to provide image data of the Earth's surface for analysing processes related to these topics (Belward & Skøien 2015). Beside the increasing temporal resolution, these new super-spectral sensors record spectral information in more spectral regions than the established, multi-spectral satellite sensors.

Effective strategies to deal with the unprecedented data amounts require the integration of additional data sources. Furthermore, methods are required that are robust against situations of persistent cloud coverage, which can significantly limit the performance of land use classifications. Phenological information is one of the most employed types of data to improve classification results. However, due to vari-

ous environmental factors that influence the spectral signature of crops, spatially explicit phenological information of high thematic depth are required. Until now, no operational method is currently available that efficiently provides this type of phenological information and integrates phenology into classification frameworks.

## 1.2 AIMS AND SCOPE

This thesis aims on the development of methods to improve the applicability of remote sensing methods for detailed, accurate and fast monitoring of cropland areas. These methods are intended to be universally applicable in various regions and for different satellite sensors and application scenarios.

To achieve this, the established strategies of vegetation indices and vegetation phenology as separability-determining variables are consolidated.

The goals of this thesis are therefore:

- to investigate how spectral indices and other spectral features can be optimised to enable optimal feature selections;

- to employ volunteered geographic information for modelling and to investigate how they can be coupled with spectral reflectance patterns;

- to develop a geostatistical model to generate spatially explicit phenological information of high thematic depth;

- to detect of phenological time frames for optimal crop species separability;

- to optimise data set selection for crop classification.

## 1.3 STRUCTURE OF THIS THESIS

This thesis is based on three research papers, which have been published in 2016 and 2018. The three papers follow the general aim that was described in Section 1.2. Prior to the presentation of the papers, an introduction is given to the thematic context of the research. This includes a description of the primary investigation areas (Chapter 3.1), and the three components of the methodology, which are land use classification (Section 2.1.1), vegetation phenology (Section 2.1.2) and vegetation indices (Section: 2.1.3). Further, Section 2.1.1.2 outlines the limitations of current land use classification schemes and presents strategies to overcome them.

The three papers form the chapters 4 to 6. Each of the papers is summarised briefly and the essence of its results and main findings are combined to contribute to the overall research aims.

In Chapter 7, a synthesis of the results presented in the individual chapters is given. This synthesis presents results of further application and methodological enhancements regarding thematic depth and usability of the demonstrated methods (Chapter 7). All presented methods and results are comprehensively discussed in Chapter 8 together with the outline of possible future research and final conclusions derived from this thesis.

# 2

SCIENTIFIC BACKGROUND

This chapter outlines the scientific background of this study. It addresses the state of the art (Section 2.1) in remote sensing-based land use classification and crop phenology, the definition of core terms and the developed research hypothesis of this thesis. The investigation areas and study sites are introduced in Section 3.1 and the data basis for the conducted research is described in Section 3.2. The chapter concludes in Section 2.2 in which the base workflow of developed methodology is presented.

## 2.1 STATE OF THE ART

The following sections describe the scientific background of the conducted research. It includes a literature review that addresses remote sensing methods for land use classification, crop phenology and current trends to integrate these research areas. This literature review leads to the formulation of the main research hypothesis, which are formulated in Section 2.1.5.

The core terms *plant phenology*, *phenological phases* and *phenological time frames* are central throughout this thesis and therefore differentiated initially. First, *plant phenology* is defined as "annually and periodically reappearing events (phases) in growth and development of plants" (Schmidt et al. 2014) and their interrelation to phases of the same or different species.

Consequently, a *phenological phase* is defined as the period between two subsequent phenological events. Synonyms for *phenological phases*, which appear in this thesis, are *phenophases*, *phenological stages* and *growing stages*.

A *phenological time frame* or *phenological (temporal) window* is understood as a temporal period, during which a single phenological phase dominates all fields covered by the respective crop type within an area, delimited by criteria such as landscape units, administrative division or data availability.

### 2.1.1 *Land use classification*

Among the applications of remote sensing-based methods that monitor the Earth's land surface, land use classification is one of the most frequently applied and consequently one of the applications on which the largest research activities were focussed. In this context, *land cover* describes the "biophysical cover of the Earth's terrestrial surface, identifying vegetation, inland water, bare soil or human infrastructure" (Gómez et al. 2016). Beyond this, *land use* describes the subdivision of anthropogenic land cover types in more detailed thematic classes to capture how people use the Earth's surface.

In principle, classification is the distinction of specific classes within a remote sensing image. Classification uses a statistical or machine-learning-based classifier that employs the class-specific reflectance differences (Schowengerdt 2006) to distinguish between classes. Land use classification uses material-specific reflectance properties of natural and artificial surfaces to find clusters of similar composition. The level of detail of these classes depends on the scale of observation and can range from aggregated classes such as built-up area, forests and cropland to detailed classes such as individual tree or crop species, vegetation communities or different roof materials in built-up areas (Schowengerdt 2006).

Land use classification is mainly performed using optical remote sensing data that describe the solar radiation that is reflected by the Earth's surface in wavelengths between 400 nm and 2500 nm. Optical techniques show limitations because they only provide usable data for classifications during daytime and cloud-free weather conditions (Blaes et al. 2005). The probability, that enough cloud-free images for a classification are available, varies between ecoregions (Whitcraft et al. 2015). Additionally, the content of valuable information decreases with decreasing sun inclination angles, because of high percentages of the image are affected by shadows. Furthermore, Vegetation indices (VIs) and especially the Normalized Difference Vegetation Index (NDVI) tend to increase when sun inclination angle is low (Goodin et al. 2004). As a consequence for mid-latitudes, images acquired during winter only contain sparse amounts of valid information.

#### 2.1.1.1 *Types of land use classification*

A huge scientific effort has been put successfully in the optimisation of classification methods, including approaches that use images from multiple acquisition dates (e. g. Guerschman et al. 2003; Lunetta et al. 2006), incorporation of Synthetic Aperture Radar (SAR) data (e. g. Blaes et al. 2005; Joshi et al. 2016) and other remote sensing data sources. Further effort was made on the development, adaptation and comparison of different classifiers (e. g. Löw et al. 2015; Tehrany et al. 2014) in order to evaluate classification performance, robustness

or flexibility. The following paragraphs describe different scientific trends in classification optimisation regarding temporality of the approaches, supervision and geometric characteristics of the input data.

TEMPORALITY    The traditional way to classify remote sensing images is to consider an image acquired on one specific point in time, usually around noon of a day. Depending on the application requirements in terms of level of detail, target classes, structure, location or size of the site, the resulting classified image may be sufficient if the acquisition date of the single image is timed well. If the application includes a higher class number and is applied in regions of rapidly changing or fragmented landscapes or over a large gradient of climatic conditions, seasonality must be considered for accurate mapping (e. g. Blaes et al. 2005). To do so, images of multiple acquisition dates provide the necessary classification information with usually higher accuracy than mono-temporal classifications. Time series of satellite imagery are also essential for image compositing techniques that result in cloud-free image composites of persistently cloud-covered study regions (e. g. Frantz et al. 2017; Griffiths et al. 2013).

SUPERVISION    Land use mapping can be performed either incorporating ground information or without any ground references (e. g. Mather & Tso 2016). The first case is denoted as supervised, the latter as unsupervised classification. Unsupervised classifications perform a class decision by only grouping pixels of similar reflectance properties. Supervised methods extract the reflectance patterns of areas which land cover is known and classify all pixels to the ground truth class that is most similar. Support vector machines, artificial neural networks, classification and regression trees, random forests and maximum likelihood classifiers are the most frequently applied supervised classifiers, while k-means and fuzzy clustering are popular examples for unsupervised classifiers (Rodriguez-Galiano et al. 2012).

GEOMETRIC BASE    Remote sensing sensors record the observed reflectance as mean value for an artificial rectangle (*pixel*) of edge lengths constantly defined in meters or arc degrees. Naturally, the pixel is often used as object for which a class decision is made, although post-processing after the classification reduces the effects of irregular misclassification or intra-field variability of field crops. To avoid these effects and to reduce the processing time, natural objects are an alternative of increasing importance (e. g. Blaschke 2010). These natural objects are formed of averaged pixel information for an area of similar characteristics, such as elevation, aspect, slope or other terrain attributes or vegetation patterns. These objects can be derived by image segmentation as additional pre-processing step.

### 2.1.1.2  *Limiting factors for classification accuracy*

Despite these immense efforts, factors of remote sensing methods still persist that limit the classification accuracy. First, no satellite sensor exists that provides data that reaches all requirements for accurate, fast, and large scale application. These requirements are

- large swath widths to reduce the requirement for data mosaicking that is related with systematic reflectance differences between the parts of the mosaic due to differing conditions at the time of image acquisition,

- high spectral resolution, i. e. the monitoring in a number of spectral bands exceeding approximately 30 instead of a small number to capture all relevant parts of the electromagnetic spectrum,

- high radiometric resolution to capture small but significant reflectance differences between materials,

- high spatial resolution to minimise the influence of pixels that contain spectral information of different materials, and

- continuous monitoring of regular repetition intervals.

All currently both freely and commercially available optical sensors fail at least one of these requirements, due to their mission design that must balance data amount and processing needs. The satellite sensors of the Landsat family for instance fulfil regular repetition, large swath widths and high radiometric resolution, but its $30\,\text{m} \times 30\,\text{m}$ pixel sizes limits the applicability in fragmented landscapes. Further, the spectral bands do not record reflectance in the red edge region, which is crucial for accurate vegetation monitoring (Schuster et al. 2012). Contrarily, hyperspectral sensors like HyMap have high spatial and spectral resolution but small swaths and no regular repetition interval due to its airborne design. The upcoming hyperspectral EnMap sensor, which also has a large number of spectral bands, only will provide data of $20\,\text{m} \times 20\,\text{m}$ pixel size on demand.

### 2.1.1.3  *Crop mapping as special case of land cover classification*

Crop classification is a more challenging task than for example monitoring of forest decline or flood detection due to often very similar spectral response of the crops. Furthermore, in many regions of intense agriculture, the cultivated crops change every season or even within seasons due to crop rotation practices, which narrows the temporal window for classification. Since accurate crop monitoring is required for food security, ecological monitoring and for policy makers (Courault et al. 2016; Gómez et al. 2016), accurate methods must be

developed that work operationally and produce results of sufficient reliability during short acquisition periods. Thus, efforts focus on various strategies to improve classification accuracies. These strategies include, among others, the incorporation of phenological parameters (e.g. Haralick et al. 1980; Knight et al. 2006; Lloyd 1990; Son et al. 2013; Zhong et al. 2011), pre-classification stratification of the study area (Franklin & Wulder 2002), incorporation of multi-sensor remote sensing data such as combinations of multi-spectral, SAR and LIDAR data (Chen et al. 2017; Joshi et al. 2016; Siachalou et al. 2015; Waldner et al. 2015) or texture (Khatami et al. 2016) as additional criteria besides reflectance or improved feature selection.

Two of these strategies, which have been identified as key methods in land use mapping (Morain 1974), are investigated within this thesis. Namely these are incorporation of plant phenology and improved feature selection. Two reasons determine the selection of the strategies:

1. The drivers of phenological development are mainly climatic and the modelling of the plant physiological responses has been studied intensively during the last decades. Consequently, findings related to phenology are probably transferable and adjustable for wide range of different landscapes.

2. There is no doubt in the remote sensing community that appropriate selection of spectral features such as vegetation indices improve classification accuracies and most recent satellite sensors allow the computation of much more sophisticated spectral features.

2.1.2   *Vegetation phenology*

According to Lieth (2013), phenology is defined as "the study of the timing of recurring biological events, the causes of their timing with regard to biotic and abiotic forces, and the interrelation of phases of the same or different species". Since this definition includes periodical events of all kinds of organisms (e.g. seasonal variations of bacterioplankton community composition, beginning of the rotting season of ruminants), the more specific term *plant phenology* is defined as the examination of annually and periodically reappearing stages in growth and development of plants, represented by clearly defined events (Schmidt et al. 2014).

The main developmental phases of plants, that can be observed for almost every species (emergence, growing, blooming, ripening and senescence), can be subdivided into shorter, species-specific phases for a more detailed description of the plant phenological cycle. These sub-phases are of importance for several applications such as yield prediction of cereal crops (Bolton & Friedl 2013), timing of fertilising actions (Filella et al. 1995) or for monitoring pollen contamin-

ation to define allergy risk time frames (Beggs 2004). For continuous monitoring, uniquely defined standards for sub-phase definitions are required to allow comparable observations, especially for volunteer-based programmes. Different numeric schemes were developed over the past decades. The development of these standards is described in detail in Meier (2003). A globally accepted numeric code system is the Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (BBCH) scale (Bleiholder et al. 2001), that defines ten main phases (Tab. 1) with up to ten sub-phases for each main phase.

Table 1: Main phases of the BBCH numerical code for plant phenological phases (Bleiholder et al. 2001)

| Codes | Main phase |
| --- | --- |
| 00...09 | Germination / sprouting / bud development |
| 10...19 | Leaf development (main shoot) |
| 20...29 | Formation of side shoots / tillering |
| 30...39 | Stem elongation or rosette growth / shoot development (main shoot) |
| 40...49 | Development of harvestable vegetative plan |
| 50...59 | Inflorescence emergence (main shoot) / heading |
| 60...69 | Flowering (main shoot) |
| 70...79 | Development of fruit |
| 80...89 | Ripening or maturity of fruit and seed |
| 90...99 | Senescence, beginning of dormancy |

An alternative phenological scale is used for the German Weather Service / German: *Deutscher Wetterdienst* (DWD) phenological monitoring programme, which also assigns numerical codes of one or two digits to each phase. Since the Phenological model for Application in Spatial and Environmental sciences (PHASE), that is developed within this study and presented in Chapter 5, is based on the DWD programme, the DWD scale (DWD 2015) is used for phase coding. Most but not all of the phases have an exact equivalent in the BBCH scale. The DWD and BBCH codes for the phases of agricultural crops that are currently monitored by the DWD programme are listed in Tab. 2.

Various applications benefit from plant phenological information, both from spatially explicit models and point-based observations. Key application fields for plant phenological information were identified by Chuine et al. (2013):

- Agrometeorology: Cereal yield or primary production, timing of agricultural management actions such as fertiliser or pesticide applications;

- Public health: Prediction of atmospheric pollen content for allergy alerts, prediction of algal blooms;

- Environmental monitoring: Detection of trends in phenological timings as consequence of climate change related variations of growing conditions.

Table 2: Crop phenological phases that are observed by the DWD (DWD 2015) phenological monitoring programme and the BBCH (Bleiholder et al. 2001) equivalents.

| Phase | DWD Code | BBCH Code |
|---|---|---|
| Greenup | 1 | - |
| Beginning of flowering | 5 | 61 |
| Full flowering | 6 | - |
| End of flowering | 7 | 69 |
| Beginning of sowing | 10 | 0 |
| Emergence | 12 | 10 |
| Closed stand | 13 | 35 |
| 4$^{th}$ leaf unfolded | 14 | 14 |
| Beginning of shooting / stem elongation | 15/67 | 31 |
| Beginning of leaf formation | 16 | 12 |
| Beginning of bud formation | 17 | 50 |
| Beginning of heading / tassel emergence | 18/66 | 51 |
| Beginning of milk ripening | 19 | 75 |
| Early dough ripening | 20 | 83 |
| Beginning of yellow ripening | 21 | 87 |
| Beginning of full ripening | 22 | - |
| Harvest | 24 | - |
| 1$^{st}$ cut for hay | 25 | - |
| 1$^{st}$ cut for silage | 26 | - |
| Tip of tassel emergence | 65 | 53 |

The basic assumption for the majority of the currently existing phenological models is that variations of the phenological development of crops and grapevine are determined by temperature differences. This dependency was already presented by Réaumur (1735). With increasing computational capacities and knowledge of climate change, phenological modelling became more important since the 1960s (Chuine et al. 2013).

While temperature is still understood as the main driving factor for plant development, other variables such as soil moisture and day length also influence plant phenology in different degrees of impact. For example, it has been frequently proven that precipitation vari-

ations have almost no effect on timing of the phenological phases in temperate regions, while in semi-arid climates drought-induced plant stress delays or advances their timing (Fu et al. 2014).

In this context, it must be emphasised, that although *timing* of phases is not affected by precipitation patterns under temperate climatic conditions, *yield* can significantly decline as consequence of drought stress. Consequently, more advanced phenological models have been developed for application in semi-arid and arid regions. These models require a much more diverse set of input variables than only elevation and temperature. Detailed overviews on phenological models are given in Chapter 5 and e. g. by Chuine et al. (2013) and Zhao et al. (2013).

### 2.1.3 *Vegetation indices*

Phenological stages are linked to seasonal fluctuations of plant vitality or plant water as well as of chlorophyll and nitrogen contents. Each of these parameters becomes manifest in individual parts of the electromagnetic spectrum. Differences in chlorophyll and nitrogen content are most effective in the visible and near infrared (NIR) spectrum parts (e. g. Hunt Jr. et al. 2013), while general plant vitality is more obvious to describe between red and NIR spectral regions. Finally, plant water content can be best investigated using reflectance from the NIR to mid- or short wave infrared (SWIR).

Vegetation indices can be formulated based on these described effects, depending on the intended application. Vegetation indices are vegetation-specific spectral indices computed by application of more or less complex mathematical operations applied on the spectral bands of a remote sensing sensor. As demonstrated in numerous studies (e. g. Asrar et al. 1984; Löw et al. 2013), the usage of VIs instead of the single spectral channels as variables for crop mapping leads to higher classification accuracy, because VIs amplify the spectral response of vegetation and reduce effects of soil and topography without eliminating these effects (Hunt Jr. et al. 2013).

An almost unmanageable number of spectral indices for various applications can be found in the literature. Beside vegetation (e. g. Bannari et al. 1995), these applications include mineralogical and geological detection (Van der Meer et al. 2012), fire monitoring (Epting et al. 2005; Harris et al. 2011) and others. Despite this huge number of specific indices, it remains uncertain whether a proposed index is really the best performing possible index for the specific purpose (Rivera et al. 2014). The optimisation of spectral indices for a specific application is therefore a very promising approach to achieve results of maximum possible accuracy.

### 2.1.4 *Phenology as variable for crop classification*

Phenology becomes increasingly integrated into classification frameworks, since different vegetation communities or species show specific differences in their developmental phases timing. These phenology-based classification (PBC) approaches, which have been applied since decades (Haralick et al. 1980; Lloyd 1990; Morain 1974; White et al. 2014; Xu et al. 2017), traditionally extract phenological metrics directly from multi-spectral reflectance data. Field-related differences of reflectance patterns can be utilised to discriminate between crop species, cultivation practices as well as between irrigated and non-irrigated agricultural areas.

Various studies have demonstrated that classification of smaller number of data sets acquired during highly significant temporal periods lead to higher classification accuracy than the classification of all available satellite images for a specific study region (e.g. Conrad et al. 2014; Murakami et al. 2001; Van Niel & McVicar 2004; Wardlow et al. 2007). These temporal windows are a consequence of the inter-relation of phase-specific reflectance characteristics that are represented by the satellite images. This was demonstrated e.g. by Murakami et al. (2001), who investigated the optimal number of images to include into a land use classification of cereals, rice, soy, lotus and rush in Japan via NDVI profiles and spectral separability indicators. The four optimal dates, however, were only investigated without including ground or satellite-derived phenological observations.

Van Niel & McVicar (2004) studied temporal windows intensively for a study region in Australia by assessing the accuracy of several single-date classifications and various multi-date combinations of images. However, they resulted in the definition of calendar months that can be used as indicators for the highest separability and refrained to interpret the time frames against a phenological background.

Wardlow et al. (2007) addressed exactly these questions with a detailed investigation of the effects of agricultural management actions (especially irrigation) and mapping performance of six major crop types. They conducted their analysis on large-area crop classification in Kansas (United States of America) using Moderate Resolution Imaging Spectroradiometer (MODIS) data and concluded with stable vegetation index profiles which can be explained by typical phenological behaviour of the monitored crops. Additionally, Wardlow et al. (2007) investigated the pairwise spectral separability by calculation of the Jeffries-Matusita distance, concluding in time frames during which the separability of one major crop from the other major crop types is maximal. However, the temporal borders of the high separability time frames were held constant across entire Kansas. Regional phenological differences of up to one month thus stayed unconsidered and further discussion of inter-annual variations was omitted.

Förster et al. (2012) proposed a method, that uses annually updated information on the current growing advance of crops for a "phenological correction" of derived NDVI profiles. Phenological *in-situ* measurements were used as data on which the phenological correction process is based on. However, Förster et al. (2012) neither addressed the limited reliability of citizen science monitoring networks nor any kind of interpolation to estimate phenological stages in not-observed regions.

The problems of accurate timing of image acquisition and the minimum number of required time steps for accurate separation of spectrally similar crop types have been investigated by Conrad et al. (2014) and Schmidt et al. (2014). Conrad et al. (2014) concluded with a number of 5 images per season to achieve accuracies > 85 % for classifications of highly structured landscape in irrigated, agriculturally used landscapes in Uzbekistan. These findings correspond relatively well to those presented by Schmidt et al. (2014) for the classification of grassland communities, which could be classified with accuracy exceeding 80 % when only 3 acquisition dates were used.

Zhong et al. (2014) presented a promising approach that uses distinctive phenological features for classification of maize, soy beans and a mixed crop class under humid climate in Kansas, USA. Their method applies a random forest classifier trained by, among others, phenological metrics data derived in a single year and applied this classifier on data sets acquired during other years. In doing so, Zhong et al. (2014) could achieve classification accuracies of more than 80 %. They concluded that phenology is the best performing source for inter-annual transferability of separability patterns compared to spectral features and accumulated temperatures.

Another recently developed approach for multi-temporal classification is denoted as Time-weighted Dynamic Time Warping Analysis (TWDTWA; Belgiu & Csillik 2018; Maus et al. 2016). This method is especially designed for time series of high temporal density. It computes a typical NDVI profile set for known crop types in a reference study site and classifies an unknown site by comparison of the profile patterns. To account for phenological differences, the reference profiles can be shifted, warped, stretched or compressed. However, also in this approach, phenological stages are simply derived from the VI profiles. Consequently, Maus et al. (2016) classified only aggregated classes such as "double cropping", "single cropping", "pasture" and "forest". Belgiu & Csillik (2018) tested the approach based on Sentinel-2 data and single-species crop classes, but did not conduct the separation of different winter cereal crop types.

2.1.5 *Objectives and research hypothesis*

The research initiatives described in the previous sections all show shortcomings regarding the thematic depth of the investigated land use classes. First, spectrally similar species are often aggregated to classes of lower thematic depth (Belgiu & Csillik 2018; Wardlow et al. 2007). Spectrally highly similar species may possibly be cropped for completely different purposes, which must be respected in ecological modelling. Thus many applications require a more detailed cropping information for individual species. For example, cereal species are often aggregated to the thematic classes "winter-cropped cereals" and "summer-cropped cereals", which disregards the fact that e. g. winter wheat is mainly cropped for food and bio-ethanol production while winter barley is predominantly used for fodder and only of minor importance for human nutrition. Hence, methods are required that allow the species-wise separation of crops over larger areas.

Second, phenology was found to be a key variable to improve the accuracy of land use classifications and also allows to deal with data gaps during cloud-covered conditions. It was repeatedly shown that time frames of higher spectral separability are related to phenological growth stages (Zhong et al. 2011). However, phenological information is usually derived from remote sensing imagery, which can only identify a small set of clearly visible phenological phases, or estimated based on static crop calendars (Conrad et al. 2014; Wardlow et al. 2007). Phenological ground observations are an alternative (Förster et al. 2012; Schmidt et al. 2014), but were not employed to systematically model spatially explicit phenological phases. However, such a modelling might reduce the uncertainty-evoking effects of volunteered geographic information collections (Flanagin & Metzger 2008), even though the DWD as data set provider applies several quality checks to the raw observational data (Hense & Müller 2007).

Third, the development of novel VIs is still subject of remote sensing research to further improve the capability of an index to map a specific reflectance-represented ground phenomenon (e. g. Hede et al. 2015; Zhang et al. 2016). However, this procedure often results in an index that is only valid under strict conditions and requires a lot of data collection and modelling efforts (Rivera et al. 2014). Hence, for classification tasks on larger scales, broader, less specifically sensitive indices are usually preferred which limits the achievable classification accuracy.

These three shortcomings of current classification frameworks lead to three hypotheses for this thesis, which are addressed in the three papers presented in chapters 4, 5 and 6:

1. The automated selection of optimised spectral features acting as predictor variables can improve classification accuracies compared to a pre-defined set of standard vegetation indices;

2. Temporal patterns in spectral separability of crop types can be explained by the site- and weather-specific phenological development of the crops represented as phenological phases that must be spatially explicitly available;

3. Multi-annually stable separability maxima can be understood as phenological indicator phases for the data set selection to reduce the number of required data sets and ground truth observations for land use classifications.

The regional focus of this research is on central Europe with its heterogeneous cropping patterns since it is characterised by favouring climatic conditions and soils. Further, high prerequisites regarding agricultural and ecological education of the farmers as well as on relatively high technological infrastructure are fulfilled. Since for larger scales the geometric resolution of satellite sensors represents the main limiting factor for thematic depth (Wu & Li 2009, see Section 2.1.1.2), the investigations focus on an ensemble of 12 major crop types that are cultivated on significant percentage of the arable land in central Europe (Tab. 3).

Table 3: Target crop species and its main usage purpose

| Crop (trivial) | Crop (scientific) | Primary usage |
|---|---|---|
| Perennial grassland | various | Forage, biofuel |
| Winter wheat | *Triticum aestivum* | Forage, nutrition, biofuel |
| Winter barley | *Hordeum vulgare* | Forage, nutrition |
| Winter rye | *Secale cereale* | Nutrition, biofuel |
| Winter rapeseed | *Brassica napus* | Biofuel |
| Maize | *Zea mays* | Biofuel, forage, nutrition |
| Sugar beet | *Betula vulgaris* subsp. *vulgaris* | Nutrition, biofuel |
| Potatoes | *Solanum tuberosum* | Nutrition |
| Durum wheat | *Triticum durum* | Nutrition |
| Summer barley | *Hordeum vulgare* | Brewing, nutrition |
| Summer oats | *Avena sativa* | Forage |
| Summer wheat | *Triticum aestivum* | Brewing, nutrition |

The methods aim at contributing to the COPERNICUS Earth observation programme managed by the European Space Agency (ESA) and the European Commission (EC). This initiative approaches an operational crop monitoring programme for Europe.

The Sentinel-2 constellation (Drusch et al. 2012), that consists of two identically constructed satellites launched in 2015 and 2017, today provides optical imagery with 6 days repetition rate. Further, its combination of spectral and spatial resolution exceeds the parameters of the RapidEye constellation, that is classified as a "contributing mission" to be used for method development.

## 2.2 METHODOLOGICAL OVERVIEW

To assess the hypotheses that were formulated in Sections 1.2 and 2.1.5, the methodological sub-components, which are individually described in Chapters 4 to 6, were integrated. A workflow that summarises the three research papers is visualised in Fig. 1.
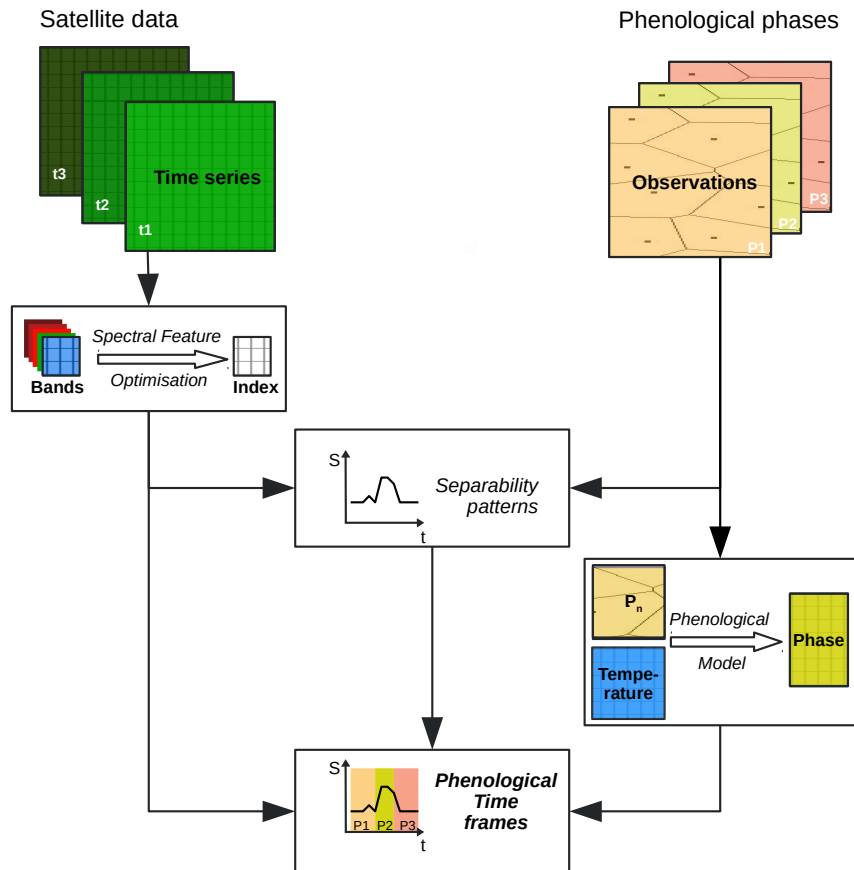


Figure 1: Workflow for the derivation of phenological time frames. $S$ represents spectral separability, $t_1$ to $t_3$ denote acquisition dates and $P_n$ represent phenological phases.

First, a method was developed that systematically finds spectral features that are most-suited for a specific land use classification scenario. Considered over time, time frames of high class-specific spectral separability are detected as an additional output. Further it is investigated, whether these time frames are related to phenological phases of the target crop classes (Chapter 4). The relationship between phenological development of crops and their spectral separability requires accurate phenological information. Volunteer-based observation programmes can provide point information of valuable input for these analysis, but show shortcomings due to small-scale climatic variations and the subjectivity of the observations. Modelling of plant-physiological responses to climate can address these limitations. Consequently, a phenological model is developed based on observations and the relationship between temperature and plant development (Chapter 5) to model the timing of important phenological phases during a growing season. These growing season-specific timings of phenological phases and spectral separability of crops are contemplated over different years. In doing so, the spatially explicit phenological information are finally linked to the separability peaks. So, phenological separability time frames and general statements on the relation of spectral separability and plant phenology are derived for the example species winter wheat, winter barley and winter rapeseed (Chapter 6).

<div style="text-align: right; font-size: 3em;">3</div>

## STUDY SITES AND DATA BASE

This chapter consist of the geographic description of the two main study sites in which the methods were developed and tested. It also summarises the characteristics of the used data sets, including remote sensing data as well as ground truth data.

### 3.1 DESCRIPTION OF THE STUDY SITES

The investigations were performed in two study sites which represent different ecological regions of Germany (Fig. 2). Both sites, one located in the Harz mountains and adjacent lowlands and the other located in the Nort-Eastern German Lowlands, are intensively investigated since they are also laboratory sites for the Terrestrial Environmental Observatories (TERENO) long-term research project (Bogena 2016)[1]. The following sections describe the two sites in detail.

#### 3.1.1 *Eastern Harz mountains and adjacent lowlands*

ADMINISTRATIVE LOCATION    The main study site is located in the German federal state Saxony-Anhalt, approximately 70 km southwest of the state capital Magdeburg and 30 km north-west of Halle (Saale). The site spans 50 km in east-west direction and 25 km from its northern to its southern border. The sites' centre is located at 51.65 °N and 11.42 °E of Greenwich (Fig. 3).

LANDSCAPE AND CLIMATIC SITUATION    The study site is part of three natural landscape units as defined by Meynen (1959). Specifically, these units are the "Northern foothills of the Harz", "Eastern foothills of the Harz and *Börde* regions" and "Harz". The climatic situation is dominated by the lee effects of the Harz mountains, resulting in comparatively low annual precipitation sums of 553 mm in the westernmost parts and 486 mm in the easternmost parts (DWD 2018a). The annual mean temperatures decrease from 8.5 °C in the west to 9.2 °C in the eastern parts of the study site (DWD 2018b).

---

1 http://teodoor.icg.kfa-juelich.de/overview-en
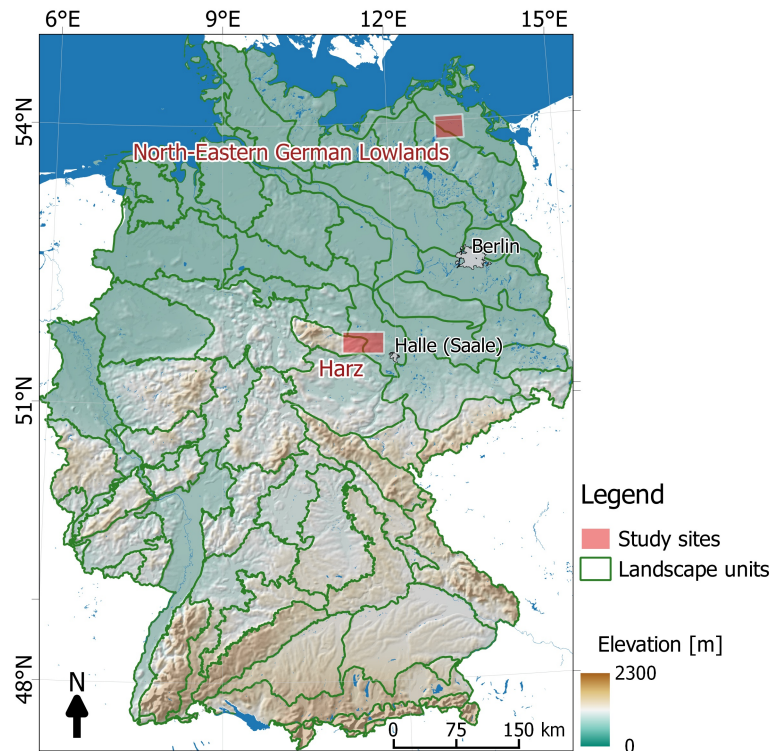
Figure 2: Location of the study sites in Germany. Data Sources: SRTM (elevation), Federal authority for environmental protection / UBA (natural units), OpenStreetMap contributers (topographic data). Projection: WGS84 UTM Zone 32N

SUBSTRATES AND SOILS    The eastern lowland parts are dominated by aeolian loess substrates that have been deposited during the Weichselian glaciation under periglacial conditions. Further, holocene fluviatile substrates can be found in the river valleys (BGR 2015a). The bedrock in the western part of the study area is dominated by various types of Devonian schist, greywacke, and quartzite conglomerates, mainly deposited during the Devonian, Ordovician and Carboniferous ages (BGR 2015b).

Due to its calcium carbonate-rich silty composition and its poriferous bedding, loess is often the primary material of highly fertile chernozemic soils, which are also the dominating soil type in the lowland parts of the study area. Fertile cambisols and regosols also cover significant percentages of the study sites. In the river valleys, alluvial gleyic soil types can be found. With increasing height to the western mountainous parts of the study site, the percentage of chernozems declines and cambisols are the most abundant soil types.

Figure 3: Location of the Harz study site and elevation. Thin grey lines represent the agricultural fields. Data sources: SRTM (elevation), ©GeoBasis-DE / BKG 2018 (topographic data). Projection: WGS84 Pseudo Mercator

LAND USE    The rural area is dominated by intensive agriculture due to the favouring conditions regarding soil and climate. The high fertility of the soils allows the cultivation of demanding crops in higher percentages than the Germany-wide average (DESTATIS 2015, see Tab. 4).

Table 4: Target crops, annual and average ($\varnothing$) cropping area in percent (%) of the total cropped area for 2010 to 2015 in the Harz test site and for entire Germany in 2015 ($DE_{2015}$)

| CROP | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | $\varnothing$ | $DE_{2015}$ |
|---|---|---|---|---|---|---|---|---|
| Winter wheat | 42.2 | 42.6 | 40.6 | 42.2 | 41.7 | 40.4 | 41.6 | 19.2 |
| Winter rapeseed | 18.8 | 16.1 | 19.5 | 19.0 | 17.8 | 17.6 | 18.1 | 7.7 |
| Perennial grassland | 11.0 | 11.2 | 10.8 | 10.6 | 10.6 | 9.6 | 10.6 | 27.9 |
| Winter barley | 8.4 | 8.4 | 8.7 | 8.8 | 8.7 | 8.5 | 8.6 | 7.5 |
| Maize | 5.3 | 6.1 | 6.3 | 7.3 | 7.3 | 7.4 | 6.6 | 12.5 |
| Sugar beet | 5.3 | 5.3 | 6.0 | 5.8 | 6.1 | 4.3 | 5.5 | 1.9 |
| Durum wheat | 2.7 | 2.6 | 2.2 | 1.0 | 1.8 | 1.5 | 2.0 | 0.1 |
| Winter rye | 0.8 | 0.8 | 0.5 | 0.4 | 0.3 | 0.3 | 0.5 | 3.7 |
| Summer barley | 0.5 | 0.9 | 0.6 | 0.3 | 0.3 | 0.3 | 0.5 | 2.2 |
| Oats | 0.4 | 0.5 | 0.4 | 0.4 | 0.6 | 0.3 | 0.4 | 0.8 |
| Summer wheat | 0.3 | 0.6 | 0.1 | 0.1 | 0.4 | 0.2 | 0.3 | 0.3 |
| Potatoes | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 1.4 |

Winter wheat is cropped on more than 40 % of the total agriculturally used area and is along with winter rapeseed, durum wheat and sugar beet over-represented by a factor of 2 to 3, while the other crop types are cropped on much less parts of the agriculturally used area than in the country-wide averages.

### 3.1.2  *North-Eastern German Lowlands*

ADMINISTRATIVE LOCATION    The second study site surrounds the town of Demmin (see Fig. 4) in the federal state Mecklenburg – West Pomerania, approximately 40 km south of the Baltic Sea and 200 km north of Berlin. The site spans 35 km each from north to south and from east to west, with its centre located at 53.93 °N and 13.15 °E of Greenwich.
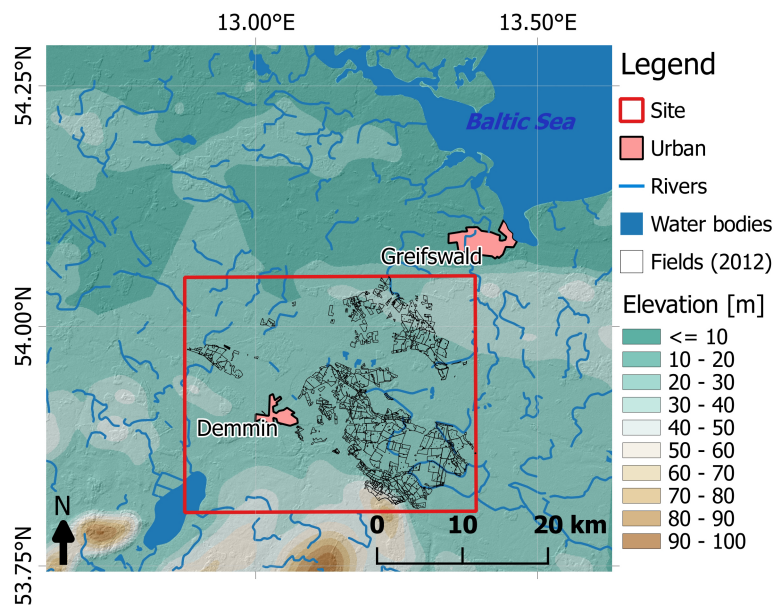


Figure 4: Location of the Northern German Lowlands test site and elevation. Thin grey lines represent the agricultural fields. Data Sources: SRTM (elevation), © GeoBasis-DE / BKG 2018 (topographic data). Projection: WGS84 Pseudo Mercator

LANDSCAPE AND CLIMATIC SITUATION    The site is completely situated within the Northern German lowlands, intersecting the natural landscape units "North-eastern Lowlands of Mecklenburg and Oder backwater area" and "Backland of the Mecklenburg-Brandenburg lake plateau" (Meynen 1959). The topography is comparatively homogeneous and planar. The climate is oceanic and cool, with annual mean temperatures of 8.7 °C (DWD 2018b) and 590 mm average annual precipitation (DWD 2018a).

SUBSTRATES AND SOILS    The bedrock is sand-, silt- and gravel-rich and has been deposited during and after the Weichselian glaciation (BGR 2015b). Consequently, the main bedrock substrates are boulder clay and boulder till. In the depressions, also glacio-fluviatile deposits and peat can be found. The soils are moderately fertile and mainly loamy gleysols, cambisols or luvisols (BGR 2015a).

LAND USE    Agriculture is the dominating land use in the study area. Due to the less fertile soils compared to the Harz study sites, the percentages of demanding crops such as winter wheat are significantly reduced and less demanding crops like winter rye are cultivated more frequently (see Tables 5 and 4).

Table 5: Target crops, annual and average (∅) cropping area in per cent(%) of the total cropped area for 2010 to 2012 in the DEMMIN test site. See Tab. 4 for Germany-wide averages.

| CROP | 2010 | 2011 | 2012 | ∅ |
|---|---|---|---|---|
| Winter wheat | 30.70 | 29.80 | 32.50 | 31.00 |
| Winter rapeseed | 23.70 | 21.10 | 22.30 | 22.40 |
| Perennial grassland | 14.20 | 18.20 | 15.40 | 15.90 |
| Winter barley | 8.30 | 6.20 | 7.10 | 7.20 |
| Maize | 8.30 | 11.40 | 7.10 | 8.90 |
| Sugar beet | 5.80 | 6.40 | 6.50 | 6.24 |
| Winter rye | 4.40 | 2.90 | 6.70 | 4.60 |
| Summer barley | 0.1 | 0.01 | 0.1 | 0.09 |
| Potatoes | 2.00 | 1.90 | 1.80 | 1.90 |

THE DEMMIN TEST FIELD    The study site is part of the TERENO long-term research project (Bogena 2016). As a part of the TERENO North-Eastern German Lowlands Observatory, the remote sensing test field is denoted as Durable Environmental Multidisciplinary Monitoring Information Network (Durable Environmental Multidisciplinary Monitoring Information Network (DEMMIN)), established and managed by the German Aerospace Centre (DLR; German: Deutsches Zentrum für Luft- und Raumfahrt e.V.).

## 3.2    DESCRIPTION OF THE DATA SETS

The data base for this research consists of remote sensing data, volunteered geographic information regarding crop phenology and cropping information per parcel for the two study sites. All these used data sets are described in the following sections.

### 3.2.1 *Remote sensing data*

RAPIDEYE DATA    Optical, multi-spectral remote sensing data recorded by the RapidEye satellite constellation are used for the spectral similarity analyses of the crops. The constellation was initially driven by the German RapidEye AG© and later acquired by Blackbridge Inc.© and Planet Inc.©. Although it is hence a commercial constellation, the data recorded by the satellites were usable free of charge for Germany-based research projects via the RapidEye Science Archive (RESA)[2] managed by the Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR). The RapidEye constellation consists of 5 individual satellites in a sun-synchronous orbit that are identically constructed (Tyc et al. 2005). The satellites were launched in 2008 and started to provide data operationally in 2009. The sensor records the reflected radiation in the 5 spectral bands blue, green, red, red edge and NIR. The detailed band characteristics are summarised in Tab. 6. The presence of a red edge band, covering the spectral region that is most sensitive to differences in chlorophyll content and vitality of plants (Schuster et al. 2012), was an unique feature when the RapidEye satellites have been launched. The Sentinel-2 sensors also include red edge bands, making RapidEye essential for the Sentinel mission preparation (Drusch et al. 2012). The RapidEye constellation is able to provide data from nadir view in revisiting times of 5 to days. Since a user-specific tasking is also possible, the effective temporal resolution of the operationally acquired images that are provided by RESA is far less than 5 days, varying between 5 an 17 acquisitions per year. The geometric resolution for the Level-1B product is 6.5 m and 5 m for the data of pre-processing Level-3A. The product of level 3A is tiled in squares of 25 km edge length, of which the tiles with the identifiers 3262921 and 3262922 correspond exactly to the Harz study site. Due to the high geometric and temporal resolutions as well as due to the presence of the red edge band, the sensor has been integrated as a contributing mission into the COPERNICUS programme driven by ESA and the EC to prepare the Sentinel missions.

Table 6: Spectral bands of the RapidEye satellites

| BAND | MINIMUM WAVELENGTH [nm] | MAXIMUM WAVELENGTH [nm] |
|---|---|---|
| Blue | 440 | 510 |
| Green | 520 | 590 |
| Red | 630 | 685 |
| Red edge | 690 | 730 |
| Near infrared | 760 | 850 |

---

2 www.resaweb.dlr.de, contract no: 653

The performed preprocessing steps and the temporal availability of data sets for the study sites are described and presented in detail in the chapters 4 and 6.

ELEVATION DATA    The Shuttle Radar Topography Mission (SRTM), performed in 2003, resulted in a Digital Elevation Model (DEM) for all latitudes between 70 °N and 70 °S (USGS 2004). The original 3-arc-seconds resolution product (Jarvis et al. 2008) was filtered and aggregated to 1 km resolution to respect the spatial uncertainty of the phenological data on the one hand and to reduce processing requirements on the other hand. The used DEM covers Germany in 358,320 raster cells. The detailed preprocessing steps are described in Chapter 5.

### 3.2.2  *Ground observation data*

PHENOLOGICAL DATA    Plant phenology in Germany is monitored nationwide by a volunteer-based observation programme managed by the DWD. Since its launch in 1951, the set monitored crop types and observed phases was only slightly modified. Approximately 1200 observers record 3 - 9 phenological stages per crop type (see Tables 1 and 2) that are meaningful for the different applications mentioned in Section 2.1.2. The observations are recorded around a pre-defined centre point enabling an estimation of their location. A detailed description of the programme and its characteristics is given in Chapter 5 and DWD (2015). The temporal consistency of the observation programme and the relatively accurate estimation of spatial uncertainty naturally defines Germany as ideal test country for phenology-related research.

LAND USE DATA    To train the models and to validate the results, highly accurate ground truth data is required. For this study, such information on field parcel level-of-detail of different data providers could be used. Those parcels are defined by three criteria (Inan et al. 2010):

1. homogeneous ownership by a single person, a company or a cooperative,

2. homogeneity regarding cultivation of one single crop species (e. g. winter wheat, sugar beet, etc.) or mixed species type (e. g. perennial grassland, winter-grain-meslin, grain-legume-meslin, etc.),

3. no interference by roads, farm tracks or other linear objects.

*Eastern Harz mountain range and adjacent lowlands:* For the entire study period from 2010 to 2012, the cultivated crops were provided for the Land Parcel Identification System (LPIS) parcels in the Harz

study site. These data were provided by the Ministry of Environment, Agriculture and Energy of Saxony-Anhalt. This data set is annually created to manage the payment of agricultural subsidies in the European Union (EU) member states. It is based on submitted forms by the farmers and interpreted orthophoto data, that represent between 80 and 100 different crop classes for every year.

*North-eastern German Lowlands:* The crop parcel data for the test site surrounding Demmin includes parcels of about 30 different crop types and covers the study site only discontinuously. The cropping information was collected in context of the TERENO research projects for the years 2010 to 2012. As consequence of the lower thematic depth, the smaller site extent, differently favouring conditions and differing project-specific requirements of the two study sites, summer barley, summer wheat and durum wheat are not recorded.

# 4

## OPTIMISATION OF SPECTRAL INDICES AND LONG-TERM SEPARABILITY ANALYSIS FOR CLASSIFICATION OF CEREAL CROPS USING MULTI-SPECTRAL RAPIDEYE IMAGERY

Paper 1 focusses on the selection of suitable spectral features or vegetation indices for land use classifications. A methodology is presented, which provides information regarding optimal spectral indices to separate two land use classes of high spectral similarity on the example of winter wheat and winter barley. The methodology tests a huge number of vegetation indices that are calculated systematically following an established vegetation index rationale with different parametrisation. To evaluate the performance of the tested indices, the effect size measure $\eta^2$ is introduced as novel indicator for spectral similarity. The optimisation procedure is applied on each data set of dense RapidEye multi-spectral time series. In doing so, spectral separability profiles could be derived that reveal temporal windows of high spectral separability. These detected time frames are interpreted against a phenological background by description of phenological differences between the target species. The utilised phenological observations reveal strong differences between closely located observation points.

# Optimization of spectral indices and long-term separability analysis for classification of cereal crops using multi-spectral RapidEye imagery

Henning Gerstmann [a,*], Markus Möller [a,b], Cornelia Gläßer [a]

[a] *Martin Luther University Halle-Wittenberg, Institute for Geosciences and Geography, Department for Remote Sensing and Cartography, Von-Seckendorff-Platz 4, 06120 Halle (Saale), Germany*
[b] *Martin Luther University Halle-Wittenberg, Institute of Agriculture and Nutrition Science, Department of Farm Management, Karl-Freiherr-von-Fritsch-Str. 4, 06120 Halle (Saale), Germany*

## ARTICLE INFO

## ABSTRACT

Crop monitoring using remotely sensed image data provides valuable input for a large variety of applications in environmental and agricultural research. However, method development for discrimination between spectrally highly similar crop species remains a challenge in remote sensing. Calculation of vegetation indices is a frequently applied option to amplify the most distinctive parts of a spectrum. Since no vegetation index exist, that is universally best-performing, a method is presented that finds an index that is optimized for the classification of a specific satellite data set to separate two cereal crop types. The $\eta^2$ (*eta-squared*) measure of association – presented as novel spectral separability indicator – was used for the evaluation of the numerous tested indices. The approach is first applied on a RapidEye satellite image for the separation of winter wheat and winter barley in a Central German test site. The determined optimized index allows a more accurate classification (97%) than several well-established vegetation indices like *NDVI* and *EVI* (<87%). Furthermore, the approach was applied on a RapidEye multi-spectral image time series covering the years 2010–2014. The optimized index for the spectral separation of winter barley and winter wheat for each acquisition date was calculated and its ability to distinct the two classes was assessed. The results indicate that the calculated optimized indices perform better than the standard indices for most seasonal parts of the time series. The red edge spectral region proved to be of high significance for crop classification. Additionally, a time frame of best spectral separability of wheat and barley could be detected in early to mid-summer.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Crop classification using remote sensing data is becoming increasingly important for land use/land cover analysis, for agricultural and environmental research and management as well as for administrative purposes on regional scale levels. However, classification of crops, especially cereal species like wheat and barley, is often restricted due to high spectral similarity of the cropland classes (Guerschman et al., 2003) and sensor-specific limitations regarding temporal, spectral and geometric resolution of the currently available satellite sensors. The first limitation can be addressed using multi-temporal classification approaches which usually lead to significantly higher accuracies (e.g., Murakami et al., 2001; Van Niel and McVicar, 2004; Guerschman et al., 2003; Prishchepov et al., 2012; Löw et al., 2015) while the latter limitation will decline as consequence of the rapid increase of available imagery of both high temporal and geometric resolution acquired by modern sensors like Sentinel-2 or RapidEye (Simmons et al., 2016).

High spectral similarity of classes can also be addressed by the selection of the most significant parts of the reflectance spectrum which is usually performed by computation of vegetation indices. These indices are combinations of mostly two or three spectral channels that amplify spectral information of high significance for a distinction of different land cover classes. The most popular indices are the *Normalized Difference Vegetation Index* (*NDVI*, Rouse et al., 1974), the *Soil-adjusted Vegetation Index* (*SAVI*, Huete, 1988) and the *Enhanced Vegetation Index* (*EVI*, Liu and Huete, 1995). The majority
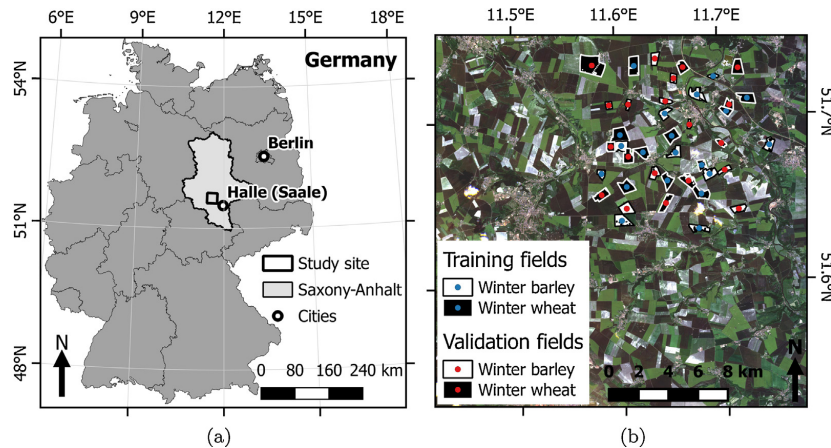
**Fig. 1.** Location of the study site (a) and training and validation parcels for the season 2014 laid over a RapidEye image acquired on 4th June 2014 (b). RGB band combination 3-2-1. Spatial reference: EPSG 32632 (Spatialreference, 2016).

of vegetation indices focuses on the spectral region between red and near infrared where reflectance of vital plants is characterized by a strong increase. A few sensors (e.g., RapidEye, Sentinel-2) have additional red edge channels between red and near infrared parts of the spectrum, and several studies demonstrated the potential of red edge reflectance information to increase land cover classification accuracies (Ramoelo et al., 2012; Schuster et al., 2012; Eitel et al., 2011).

Numerous vegetation indices have been developed over the past 40 years (see e.g., Dorigo et al., 2007; Bannari et al., 1995) for mapping of different vegetation parameters. However, Rivera et al. (2014) stated, that a specific index is not necessarily the best performing index to address the studied parameter for which the index is reported to be sensitive. Furthermore, vegetation index profiles computed from index values of different acquisition dates reveal differences in the separability of certain vegetation classes over time. To select the most sensitive index for a specific application and a specific time, Le Maire et al. (2008) proposed a brute-force approach for the selection of the optimal vegetation index to retrieve biophysical parameters of forests. Based on this idea, Rivera et al. (2014) developed a tool (*Automated Radiative Transfer Models Operator – ARTMO*) that calculates a huge number of possible vegetation indices and band combinations within an iterative loop and evaluates the performance of each tested index in regression-based estimation of a biophysical parameter. However, regression methods require the dependent variable to be at least interval-scaled, which is not given for categorical scaled land cover classes.

In this study, we first show how optimized indices can be calculated in an automatic manner for a mono-temporal data set selected out of a time series by phenological analysis. In this context, we introduce the measure of association $\eta^2$ (*eta-squared*) as novel non-parametric indicator for spectral separability. Second, we demonstrate how separability profiles can be computed using the full time series to describe separability variations within one year and different years applying spectral similarity measures. Consequently, optimal times frames for classification can finally be determined.

The approach is applied on a time series of RapidEye satellite data with high temporal resolution for the separation of winter wheat and winter barley in a Central German test site.

## 2. Study site and data

### 2.1. Site description

The study site is located in Central Germany (Fig. 1a), approximately 30 km north of the city of Halle (Saale). The site is characterized by intensive agricultural land use due to its highly fertile chernozemic soils and relatively warm climatic conditions. Cultivation data on parcel scale were obtained from the Ministry of Agriculture and Environment of Saxony-Anhalt and used as ground truth information. The most frequently grown crop types within the study site are winter wheat ($\approx$45–48% of the total cropped area between 2010 and 2014), winter oilseed rape (17–20%) and winter barley ($\approx$8–9%). Since oilseed rape can easily be separated from wheat and barley during flowering, the two cereals winter wheat and winter barley have been selected as target crops for this study due to their close genetic relationship and similar spectral behavior.

### 2.2. Satellite data

RapidEye data was obtained from the *RapidEye Science Archive*[1] which is funded by German Aerospace Center (*Deutsches Zentrum für Luft- und Raumfahrt*; DLR) and gains restricted access to RapidEye imagery. RapidEye is a commercial multi-spectral sensor constellation which consists of five identically constructed satellites (Tyc et al., 2005). The constellation provides imagery with high repetition rate, a spatial resolution of 6.5 m × 6.5 m, and five multi-spectral bands which cover the blue ($R_B$: 440–510 nm), green ($R_G$: 520–590 nm), red ($R_R$: 630–685 nm) and near infrared spectral ranges ($R_{NIR}$: 760–850 nm). In addition, a red edge band ($R_{RE}$: 690–730 nm) is available which is assumed to be sensitive to the abrupt reflectance rise caused by vegetation's chlorophyll status (Schuster et al., 2012).

The time series used in this study covers five years and consists of 43 images (Fig. 2). While the RapidEye constellation technically allows a repetition rate of less then a week, far less appropriate images are available mostly due to cloud cover and order prioritization. The temporal coverage of the time series is inconsistent

---

[1] *RESA* www.resa.blackbridge.com, project ID: 653.

**Table 1**
Well-established indices and required parameter configuration.

| $I^s$ | $B_1$ | $B_2$ | $B_3$ | $c_1$ | $c_2$ | $L$ | $G$ | Reference |
|---|---|---|---|---|---|---|---|---|
| NDVI | $R_{NIR}$ | $R_R$ | – | 1 | 1 | – | 1 | Rouse et al. (1974) |
| GNDVI | $R_{NIR}$ | $R_G$ | – | 1 | 1 | – | 1 | Gitelson et al. (1996) |
| EVI | $R_{NIR}$ | $R_R$ | $R_B$ | 6 | 7.5 | 1 | 2.5 | Huete et al. (1999) |
| EVI2 | $R_{NIR}$ | $R_R$ | – | 2.4 | 0 | 1 | 2.5 | Jiang et al. (2008) |
| SAVI | $R_{NIR}$ | $R_R$ | – | 1 | 1 | 0.5 | 1.5 | Huete (1988) |
| $EVI^{RE}$ | $R_{NIR}$ | $R_R$ | $R_B$ | 6 | 7.5 | 1 | 2.5 | – |
| $NDVI^{RE}$ | $R_{NIR}$ | $R_{RE}$ | – | 1 | 1 | – | 1 | Barnes et al. (2000) |
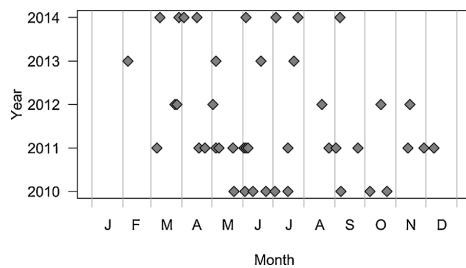


**Fig. 2.** Available RapidEye images.

throughout the years, varying from four images in 2013 to 16 scenes in 2011. No images from late spring and summer were available for 2012, while in 2010 and 2013 no early and mid-spring scenes could be used. All data sets were obtained at preprocessing level 3A, which means that radiometric, sensor, geometric corrections and geo-referencing were applied by the data provider. Each image covers 25 km × 25 km with orthorectified pixel size of 5 m × 5 m.

### 2.3. Selection of training fields

Twenty fields per year for each of the crop types winter wheat and winter barley were selected, 10 for training and validation, respectively (Fig. 1b). They were chosen under the premises that they have no *NoData* parts during the vegetation periods, no clouds and no snow coverage.

### 2.4. Phenological data

To characterize the typical growth cycle, phenological observations were analyzed. The date of entry of selected phenological phases for a variety of tree and agricultural species including winter wheat and winter barley are continuously recorded following standardized criteria by volunteer observers, managed by the German Weather Service (Deutscher Wetterdienst, 2015). The observations can be accessed via FTP[2].

## 3. Methods

### 3.1. Index permutation

The computation for the most vegetation indices is based on a normalized-difference rationale with only different spectral bands included. Equation (1) represents the structure of the formula used for the index permutation which is equivalent to the formula used for the calculation of the *EVI* (Huete et al., 1999). Within Eq. (1), $B_1$, $B_2$ and $B_3$ can be any of the spectral bands; $i$ represents all possible

---
[2] ftp.dwd.de/pub/CDC/observations_germany/phenology/.

combinations of $B_1$, $B_2$ and $B_3$ and the empirical constants $G$, $c_1$, $c_2$ and $L$.

$$I_i^{perm} = G \times \frac{B_1 - B_2}{B_1 + c_1 \times B_2 - c_2 \times B_3 + L} \quad c_1, c_2, L, G \in \mathbb{R} \qquad (1)$$

A set of well-established indices ($I^s$) was calculated for comparison. $I^s$ contains vegetation indices and red-edge-modified variants which are an element of $I_i^{perm}$ (see Table 1) as well as others which are not included. The latter are the *Wide Dynamic Range Vegetation Index* (*WDRVI*, Gitelson, 2004), which is computed similarly to *NDVI* but with a scaling of the near infrared reflectance by 0.15, the *Green Chlorophyll Index* (*CIgreen*, Wu et al., 2012), which is calculated by subtraction of 1 of the NIR/blue ratio, and the *simple NIR/Red edge ratio* (*SRR*; Ramoelo et al., 2012). Since spectrally very similar crop types are expected to show significant but narrow differences in the red edge region, the red edge variants of the established *NDVI* and *EVI* are mainly addressed within this study. Furthermore, indices calculated from the red and near infrared reflectance tend to show higher saturation effects than indices were the red band is replaced by the red edge reflectance information (e.g., Eitel et al., 2011).

### 3.2. Spectral similarity assessment

The indices calculated by permutation of Eq. (1) can be understood as interval-scaled variable, while class information is of nominal scale. Thus correlation coefficients like Pearson's $r$ and Spearman's $\rho$ cannot be used for the assessment of the index performance. However, the effect-size measure $\eta^2$ can be understood as correlation coefficient assessing the influence of class membership and index value. The metric is calculated by $\eta^2 = SS_B/SS_T$, where $SS_B$ is the sum of the squared differences of each index value to the class-specific mean, and $SS_T$ is the sum of the squared differences of each index value to the overall mean. The values range from 0 to 1, where values higher than 0.26 indicate a strong dependency of the predicted variable on the predictor (*Cohen's rule of thumb*, Cohen, 1988). In the classification case, class membership is the predictor variable on which the vegetation index value depends on.

The capability of $\eta^2$ to quantify spectral similarity is compared to the often applied *Kolmogorov–Smirnov distance* ($d^{KS}$). These two metrics were selected to address the fact that vegetation indices and remote sensing data in general are not necessarily normally distributed (Swain and Davis, 1978). Both metrics can be used to compare distributions without requiring statistical preconditions like normal distribution. Other often applied distances measures like *Jeffries–Matusita distance* (Vaiphasa et al., 2005) and *divergence* (Swain and Davis, 1978) are inappropriate since they are based on the maximum-likelihood decision rule and thus require normally distributed class samples (Richards, 2012).

The *Kolmogorov–Smirnov distance* $d^{KS}$ was proposed by Kolmogorov (1933) and is defined as the maximal difference between the empirical cumulative distribution functions (*ECDF*) of two samples (Massey, 1951). Values of $d^{KS} \simeq 1$ indicate high separability while identical spectra are characterized by $d^{KS}$ close to 0. This measure has been frequently applied in remote sensing for feature selection, class separability assessment and multi-temporal change

**Table 2**
Domains of parameters which are permuted for index calculation.

| $B_1, B_2, B_3$ | $c_1, c_2$ | $L$ | $G$ |
| --- | --- | --- | --- |
| $[R_B, R_G, R_R, R_{RE}, R_{NIR}]$ | $[-1, 0, 1, 2.4, 6, 7.5]$ | $[-1, -0.5, 0, 0.5, 1]$ | $[1, 1.5, 2.5]$ |

**Table 3**
Mean observed day of year (*DOY*) for selected phenological phases within the study site and range between earliest and latest observation between 2010 and 2014.

| | | Phases | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Year | Sowing | Shooting | Heading | Yellow ripening | Harvest |
| Winter wheat | 2010 | 262 ± 2 | 133 ± 13 | 159 ± 1 | 204 ± 5 | 218 ± 16 |
| | 2011 | 263 ± 0 | 117 ± 2 | 148 ± 5 | 194 ± 6 | 203 ± 4 |
| | 2012 | 263 ± 1 | 117 ± 4 | 141 ± 0 | 195 ± 2 | 207 ± 2 |
| | 2013 | 258 ± 2 | 128 ± 4 | 150 ± 7 | 204 ± 4 | 216 ± 3 |
| | 2014 | 259 ± 0 | 113 ± 1 | 137 ± 5 | 191 ± 5 | 205 ± 1 |
| Winter barley | 2010 | 261 ± 0 | 119 ± 5 | 144 ± 2 | 180 ± 2 | 198 ± 2 |
| | 2011 | 262 ± 4 | 120 ± 4 | 131 ± 1 | 167 ± 3 | 187 ± 7 |
| | 2012 | 264 ± 1 | 115 ± 2 | 131 ± 1 | 171 ± 2 | 191 ± 1 |
| | 2013 | 248 ± 16 | 126 ± 5 | 149 ± 9 | 177 ± 2 | 199 ± 1 |
| | 2014 | 265 ± 0 | 108 ± 0 | 112 ± 0 | 167 ± 0 | 199 ± 0 |

detection applications or geometric accuracy assessment (Rounds, 1980; Möller et al., 2012; Tang et al., 2011; Möller et al., 2013).

A practical difference of the two calculated spectral separability indicators is that $\eta^2$ can potentially be used for a multi-class classification scenario, while the *Kolmogorov–Smirnov distance* can only be used to evaluate spectral differences of two different classes.

### 3.3. Classification and validation

The optimized indices were calculated for training and validation fields. An unsupervised *k-means* clustering with five iterations was applied to the indices and compared to the actual class memberships of the pixels. A confusion matrix was calculated that contains the number of pixels which were classified correctly and which pixels were assigned to a wrong class. The percentage of correctly classified pixels to the total number of pixels represents the overall classification accuracy *A*.

The classification accuracy corresponding to the highest separability metric value ($A([\eta^2 \vee d^{KS}]_{max})$) is compared to the maximum classification accuracy ($A_{max}$) to evaluate the capability of the metric to select the most significant parameter combination for a classification problem.

### 3.4. Implementation

The satellite data were atmospherically corrected using the *Fast line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH)* algorithm (Anderson et al., 2002). The remote sensing software package ENVI® 5.2 version (Exelis Visual Information Solutions Inc., Boulder, CO, USA) and its FLAASH® module were used for atmospheric correction.

The actual methodology was implemented within the statistical computing environment *R* (R Core Team, 2015). First, the satellite data sets were cropped and masked to the extent of the training sites for the respective harvesting season. In doing so, only the test fields, on which winter wheat and winter barley were harvested in 2011, were extracted from all satellite data sets acquired later than August 2010 and for the seasons 2012, 2013 and 2014 accordingly.

All combinations of the parameters described in Table 2 within Eq. (1) were calculated. The domain of the parameters $c_1$ and $c_2$ was chosen in a way that it includes the empirical values determined by Huete et al. (1999) and Jiang et al. (2008). The similarity measure $\eta^2$ was computed using the function *etasq()* which is included in the *R* package *heplots* (Fox et al., 2015).

The function *kmeans()* that was used for validation is included in *R*'s base distribution, while the functions *confusion.matrix()* and *similarity.index()* are included in the package *clv* (Nieweglowski, 2013).

## 4. Results

The first objective of this study was to develop a framework for the detection of the optimal band and parameter combination for the index permutation approach (Section 4.1). The actual optimization procedure is performed on a single data set, but this data set has to be selected out of time series of satellite images using comprehensible criteria, e.g., phenological aspects. The algorithm is then applied to generate multi-annual separability profiles (Section 4.2).

### 4.1. Mono-temporal index optimization

#### 4.1.1. Phenological development
Winter wheat and winter barley show differences in their phenological behavior. The phenotypes of the target species are highly similar until heading, when grains and, especially for barley, awns are built. Hence, an acquisition date is probably best suited for wheat and barley separation, on which at least one of the target species already reached the phenological stage of heading.

In Germany, winter barley is usually sown in mid-September, shooting is observed in mid-April, heading in mid-May, yellow ripening in late June and harvest in mid-July. Winter wheat is sown between late September and mid-October and harvested in early August. Shooting is usually observed in late April, heading in late May or early June, and yellow ripening in the second half of July.

The phenological development of the selected field crops differs between locations, years and species. Within the study site, three phenological stations exist that observe phases for winter wheat. For winter barley also information observed on three stations were available, except for 2014, when only one station observed winter barley phases. The mean observed day as well as the variation within the site are listed in Table 3. In general, winter wheat reaches a phenological stage approximately 10–20 days after winter barley reached the same stage. The variations within the site are usually smaller than 5 days, which indicates a homogeneous plant development in the study region.
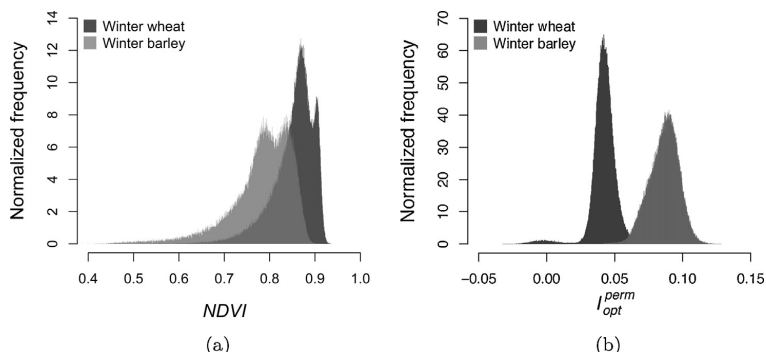
**Fig. 3.** Per-class index distributions of *NDVI* (a) and $I_{opt}^{perm}$ (b) on 4th June 2014.

### 4.1.2. Parameter determination for formulation of the optimized index

Considering the phenological behavior of winter wheat and winter barley, the RapidEye image acquired on 4th June 2014 (*DOY* = 184) is selected to demonstrate the optimization approach derived by the index permutation, because on this date the two crops are most likely at different growth stages (Table 3) and the differences in their phenotypes (see Section 4.1.1) are expected to be visible. According to Table 3, winter barley already reached the phenological stage of heading on that date, while winter wheat did not yet.

The resulting optimal vegetation index $I_{opt}^{perm}$ (see Eq. (1)) is characterized by $\eta^2 = 0.82$ and calculated as following:

$$I_{opt}^{perm} = \frac{R_{RE} - R_B}{R_{RE} + 7.5 \times R_B + R_{NIR} + 1} \tag{2}$$

*NDVI*$^{RE}$ has been found to be the best standard vegetation index $I_{opt}^{s}$ with $\eta^2 = 0.56$. In contrast, the traditional *NDVI* only shows a $\eta^2$ value of 0.26. Finally, the comparison of $I_{opt}^{perm}$ and *NDVI* distributions illustrates the effect of index optimization (Fig. 3). While the *NDVI* distributions of winter wheat and winter barley are characterized by a high degree of overlap, the $I_{opt}^{perm}$ distributions are clearly distinguishable from each other.

### 4.1.3. Classification and validation

*NDVI* and $I_{opt}^{perm}$ were calculated for all test and validation fields. A *k-means* classification was performed, and the classification
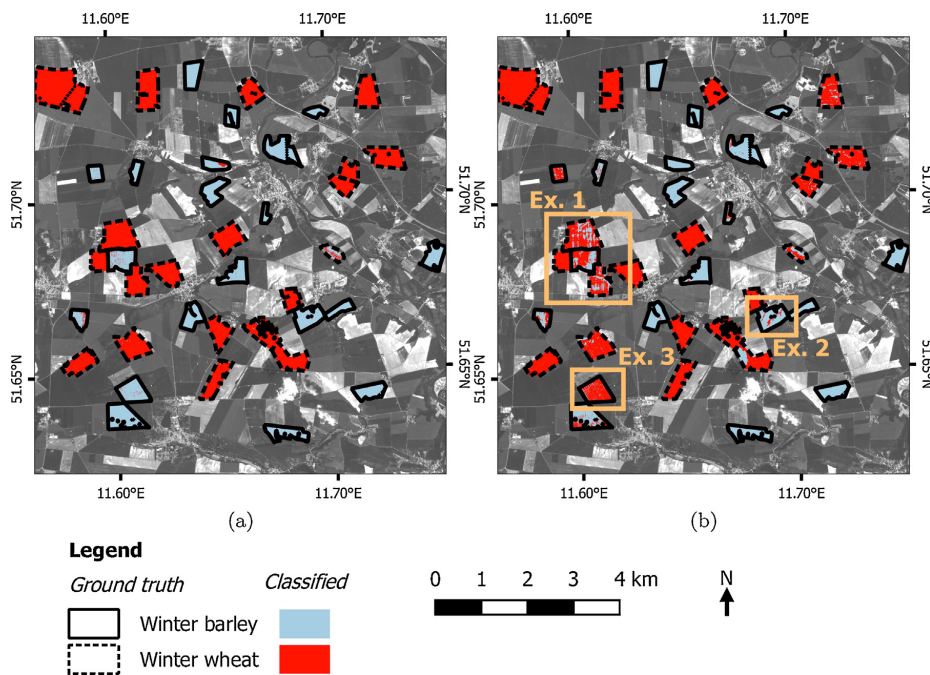


**Fig. 4.** *K-means* classifications based on $I_{opt}^{perm}$ (a) and $I_{opt}^{s}$ (b) for 4th June 2014 and actually cultivated crops. The orange boxes indicate parcels where major differences are visible (b). Projection: EPSG 4258 (Spatialreference, 2016). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
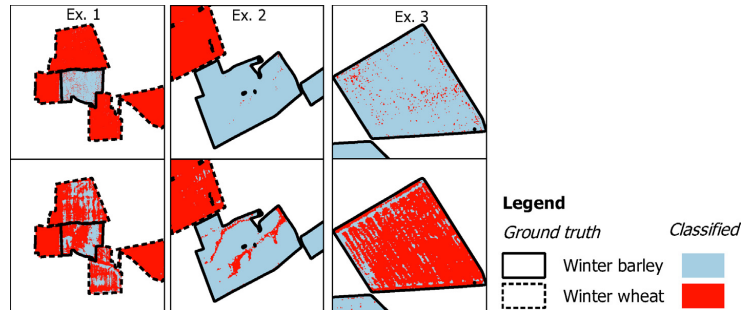
**Fig. 5.** Example fields for different classification results (see Fig. 4) between $I_{opt}^{perm}$ (top row) and $I_s^{perm}$ (bottom row).

accuracy was assessed (see Section 3.3). The classification results are shown in Fig. 4.

The classification based on $I_{opt}^{perm}$ reaches an accuracy $A = 0.97$, while for $I_{opt}^s$ it is only $A = 0.87$. In Fig. 5, the differences between the two index classification results are visualized on the example of three subsets (see Fig. 4):

- Example 1 shows that the classification based on $I_{opt}^{perm}$ leads to more homogeneous parcel-specific classes. In contrast, the $I_{opt}^s$ classification tends to a mixture of both crop types in one parcel (Fig. 5, left).
- In the case of example 2 (Fig. 5, center), a clearly recognizable river channel was false classified as winter wheat using $NDVI^{RE}$, while $I_{opt}^{perm}$ is robust against phenomena associated with riverbeds such as higher soil water content, less vegetation coverage, etc. Almost the complete field was classified correctly as winter barley.
- While the classified index $I_{opt}^{perm}$ is almost exactly corresponding to the ground truth data, some fields are nearly completely classified wrong when $NDVI^{RE}$ was used. In example 3 (Fig. 5, right), the pixels within the example winter barley field are false classified as winter wheat by about 90%.

### 4.2. Multi-annual separability profiles

#### 4.2.1. Parameter determination for optimized indices

After applying the workflow on the five-year time series, a spectral separability profile has been computed which summarizes the maximal $\eta^2$ values for each acquisition date. Fig. 6 shows the results of the similarity assessment for the complete time series. The threshold of $\eta^2 > 0.26$ (see Section 3.2) is exceeded on 20 of the 43 acquisition dates. Twelve of these dates are between May and August of the respective seasons, while eight are between October and March. The summer maxima show significantly higher $\eta^2$ values. An exception is the 2012 season, when no images between mid-May and mid-August were available.

The configurations of $I_{opt}^{perm}$ as well as $I_{opt}^s$ are listed in Table 4 for all acquisition dates on which $\eta^2(I_{opt}^{perm})$ exceeds the 0.26 threshold. The three scenes acquired on 16th July 2010, 22nd July 2013 and 26th July 2014 have been excluded since the phenological observations (Table 3) indicate that barley is most likely already harvested on these dates while winter wheat is still unharvested. Visual inspections of the satellite images support these findings.

The results confirm the significance of the spectral region between red and near infrared for the detection of differences in vegetation status in general as well as for the classification of winter wheat and winter barley in particular. Consequently, the vast majority of $I_{opt}^{perm}$ combinations include $R_{NIR}$ (16 occurrences), followed by the $R_R$ and $R_{RE}$ (10 occurrences, respectively). Indices

computed only of two spectral bands were selected on two dates. Only on four dates, an $I_{opt}^s$ was selected that does not include the red edge band, which underlines the importance of the red edge spectral region. The bands, that were selected most frequently are $R_{NIR}$, $R_{RE}$ and $R_B$ (five occurrences) and $R_{NIR}$, $R_{RE}$ and $R_R$ with four occurrences, with different assignment to $B_1$, $B_2$ and $B_3$ and different scaling parameters.

The $\eta^2$ differences between $I_{opt}^{perm}$ and $I_{opt}^s$ vary throughout the time series. Especially in 2014 the optimal band combinations outperform $I_{opt}^s$, while in 2010 only small differences were observed. On three dates, $I_{opt}^s$ is far below the separability threshold of 0.26 while $I_{opt}^{perm}$ exceeds this threshold (1st May 2012, 10th and 29th March 2014). On these days, $I_{opt}^{perm}$ includes the green band which indicates an specific importance of this spectral region for discrimination of winter wheat and winter barley that is poorly addressed by the most standard vegetation indices.

#### 4.2.2. Classification and validation

All possible index combinations have been calculated and clustered for the training and validation fields for all of the available RapidEye images. Confusion matrices have been calculated and the overall classification accuracy was assessed. Here, $A$ values close to 1 represent high accordance between the validation classes and the *k-means* clusters. For each element of $I_i^{perm}$, $d^{KS}$ was also calculated.

In Table 5, the classification accuracies achieved by classification of $I_{opt}^{perm}$ and $I_{opt}^s$ as well as the classification accuracies using the index with $d_{max}^{KS}$ are listed for the dates on which $\eta^2$ exceeds the significance threshold of 0.26.

For most of the acquisition dates, the accuracy derived by classification of the optimized index using $\eta^2$ as separability metric is higher than if other metrics or standard indices were used, except for five acquisition dates:

- On two dates (4th June 2011 and 14th November 2012), the $A(\eta_{max}^2)$ is lower than the classification accuracies of the best standard indices and the optimized index using $d_{max}^{KS}$. This difference is very small (1%) on 14th November 2012 but higher, 5% compared to $A(d_{max}^{KS})$ and 9% compared to $I_{opt}^s$, on 4th June 2011.
- On 29th March 2014, $d_{max}^{KS}$ finds an index that leads to slightly (2%) higher classification accuracy than $\eta_{max}^2$.
- On 23th October 2010 and 2nd June 2011, the $I_{opt}^s$ outperforms $I_{opt}^{perm}$ by 2 and 3%.

Despite these exceptions, which are mostly characterized by only very small accuracy differences, the results underline that $\eta^2$ finds a better suited index with higher probability than $d^{KS}$ and the optimized indices are usually better suited for the tested classification scenario than the established standard indices.
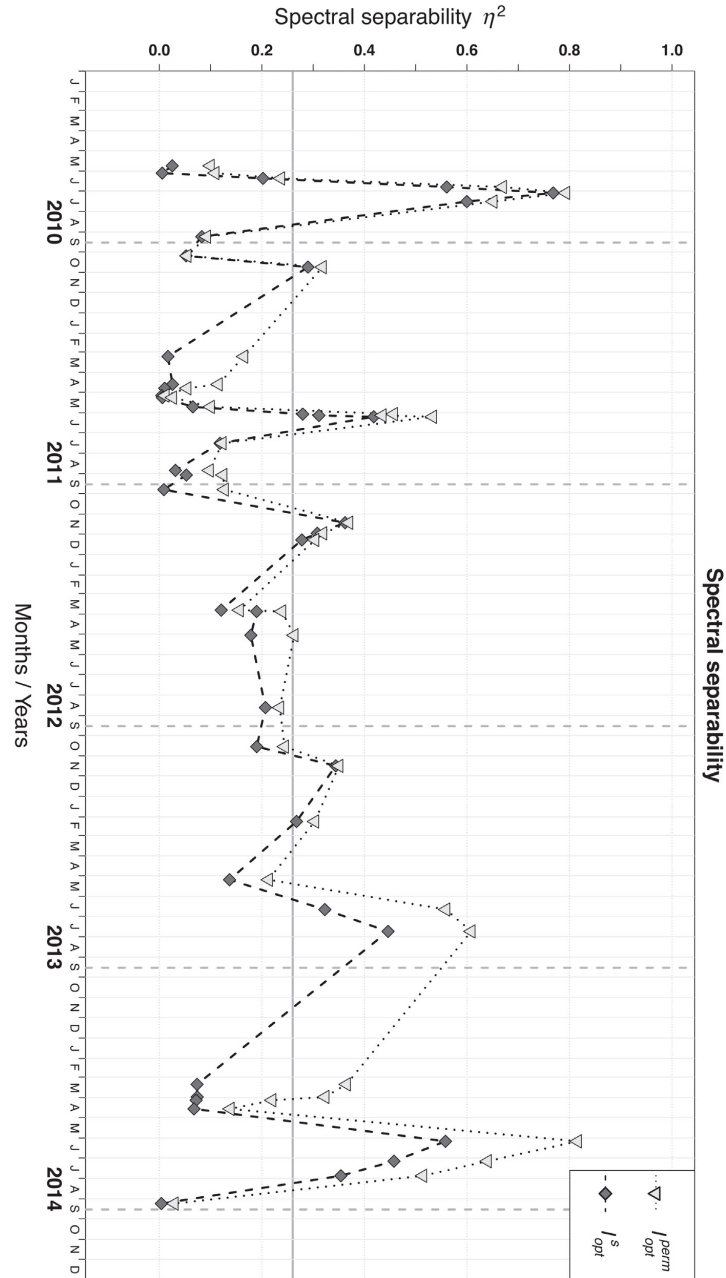
**Fig. 6.** Spectral separability profiles for the cropping seasons from 2010 to 2014. Points and triangles represent the acquisition dates of the RapidEye image. The bold vertical line indicates the threshold for a high association between class membership and index values, the dashed horizontal lines are the approximate sowing dates.

To underline these findings, the maximal classification accuracy of all index variants ($A_{max}$) was compared to the accuracy of the clustered $I_{opt}^{perm}$ for both spectral separability measures, $A(\eta_{max}^2)$ and $A(d_{max}^{KS})$, as shown in Fig. 7. $A_{max}$ ranges between 0.974 on 3rd Jul 2010 and 0.61 on 9th September 2014. The accuracy of the clustered $I_{opt}^{perm}$ indicated by $\eta^2$ ($A(\eta_{max}^2)$), ranges from 0.974 again on 3rd July 2010 and 0.435 on 3rd June 2010. Indices determined as optimal using $\eta^2$ show a significantly higher correlation ($r = 0.90$) to the $A(d_{max}^{KS})$ values ($r = 0.77$), which illustrated the capability of $\eta^2$ for index selection for a dichotomous classification problem.

**Table 4**
Optimal bands and parameter values ($I^{perm}_{opt}$), best standard vegetation indices ($I^{s}_{opt}$) and corresponding $\eta^2(I^{perm}_{opt})$ and $\eta^2(I^{s}_{opt})$ values for acquisition dates of high separability.

| Date | $B_1$ | $B_2$ | $B_3$ | $c_1$ | $c_2$ | $L$ | $G$ | $I^{s}_{opt}$ | $\eta^2(I^{perm}_{opt})$ | $\eta^2(I^{s}_{opt})$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 24 June 2010 | $R_{RE}$ | $R_R$ | $R_{NIR}$ | 0 | 6 | −1 | 1 | SRR | 0.67 | 0.56 |
| 3 July 2010 | $R_B$ | $R_{RE}$ | $R_{NIR}$ | −1 | 1 | 0 | 1 | $NDVI^{RE}$ | 0.79 | 0.77 |
| 23 October 2010 | $R_R$ | $R_{NIR}$ | – | −1 | −1 | −1 | 1 | SAVI | 0.32 | 0.29 |
| 2 June 2011 | $R_B$ | $R_{NIR}$ | $R_{RE}$ | 1 | 1 | 1 | 1 | $EVI^{RE}$ | 0.46 | 0.28 |
| 4 June 2011 | $R_{RE}$ | $R_B$ | $R_{NIR}$ | 7.5 | −1 | 1 | 1 | $EVI^{RE}$ | 0.43 | 0.31 |
| 6 June 2011 | $R_R$ | $R_{RE}$ | $R_{NIR}$ | 2.4 | 6 | −1 | 1 | $EVI^{RE}$ | 0.53 | 0.42 |
| 13 November 2011 | $R_R$ | $R_{NIR}$ | – | 0 | −1 | 1 | 1 | $EVI^{RE}$ | 0.37 | 0.36 |
| 29 November 2011 | $R_{NIR}$ | $R_R$ | $R_{RE}$ | −1 | 6 | −1 | 1 | SRR | 0.32 | 0.31 |
| 9 December 2011 | $R_{NIR}$ | $R_R$ | $R_{RE}$ | −1 | 6 | −1 | 1 | SRR | 0.30 | 0.28 |
| 1 May 2012 | $R_G$ | $R_{NIR}$ | $R_B$ | 0 | 2.4 | −1 | 1 | SRR | 0.26 | 0.18 |
| 14 November 2012 | $R_{NIR}$ | $R_R$ | $R_G$ | 2.4 | 1 | 1 | 1 | SAVI | 0.35 | 0.34 |
| 6 February 2013 | $R_{NIR}$ | $R_R$ | $R_B$ | 6 | 6 | 1 | 1 | WDRVI | 0.30 | 0.27 |
| 19 June 2013 | $R_{RE}$ | $R_B$ | $R_{NIR}$ | 6 | −1 | 1 | 1 | $EVI^{RE}$ | 0.56 | 0.32 |
| 10 March 2014 | $R_G$ | $R_B$ | $R_R$ | 2.4 | 2.4 | 1 | 1 | $EVI^{RE}$ | 0.37 | 0.07 |
| 29 March 2014 | $R_G$ | $R_{RE}$ | $R_{NIR}$ | 1 | −1 | −1 | 1 | $EVI^{RE}$ | 0.32 | 0.07 |
| 4 June 2014 | $R_{RE}$ | $R_B$ | $R_{NIR}$ | 7.5 | −1 | 1 | 1 | $NDVI^{RE}$ | 0.82 | 0.56 |
| 4 July 2014 | $R_R$ | $R_{NIR}$ | $R_B$ | 2.4 | 7.5 | −1 | 1 | SAVI | 0.64 | 0.46 |

**Table 5**
K-means classification accuracies A for optimized indices $I^{perm}_{opt}$ and standard indices $I^{s}_{opt}$ detected by using $\eta^2_{max}$ and $d^{KS}_{max}$.

| Date | $I^{perm}_{opt}$ | | | | $I^{s}_{opt}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | $\eta^2_{max}$ | $A(\eta^2_{max})$ | $d^{KS}_{max}$ | $A(d^{KS}_{max})$ | $\eta^2_{max}$ | $A(\eta^2_{max})$ | $d^{KS}_{max}$ | $A(d^{KS}_{max})$ |
| 24 June 2010 | 0.67 | 0.94 | 0.87 | 0.93 | 0.56 | 0.92 | 0.85 | 0.93 |
| 3 July 2010 | 0.79 | 0.97 | 0.95 | 0.89 | 0.77 | 0.96 | 0.94 | 0.96 |
| 23 October 2010 | 0.32 | 0.72 | 0.62 | 0.72 | 0.29 | 0.75 | 0.55 | 0.75 |
| 2 June 2011 | 0.46 | 0.70 | 0.67 | 0.70 | 0.28 | 0.72 | 0.51 | 0.72 |
| 4 June 2011 | 0.43 | 0.66 | 0.65 | 0.71 | 0.31 | 0.74 | 0.55 | 0.74 |
| 6 June 2011 | 0.53 | 0.88 | 0.75 | 0.67 | 0.42 | 0.73 | 0.64 | 0.73 |
| 13 November 2011 | 0.37 | 0.83 | 0.57 | 0.70 | 0.36 | 0.82 | 0.50 | 0.79 |
| 29 November 2011 | 0.32 | 0.82 | 0.60 | 0.70 | 0.31 | 0.80 | 0.48 | 0.80 |
| 9 December 2011 | 0.30 | 0.81 | 0.59 | 0.71 | 0.28 | 0.78 | 0.50 | 0.80 |
| 1 May 2012 | 0.26 | 0.79 | 0.59 | 0.72 | 0.18 | 0.71 | 0.49 | 0.42 |
| 14 November 2012 | 0.35 | 0.85 | 0.57 | 0.86 | 0.34 | 0.86 | 0.57 | 0.86 |
| 6 February 2013 | 0.30 | 0.83 | 0.51 | 0.67 | 0.27 | 0.81 | 0.47 | 0.82 |
| 19 June 2013 | 0.56 | 0.90 | 0.80 | 0.89 | 0.32 | 0.86 | 0.77 | 0.86 |
| 10 March 2014 | 0.37 | 0.84 | 0.69 | 0.61 | 0.07 | 0.59 | 0.28 | 0.59 |
| 29 March 2014 | 0.32 | 0.73 | 0.54 | 0.75 | 0.07 | 0.58 | 0.23 | 0.43 |
| 4 June 2014 | 0.82 | 0.97 | 0.97 | 0.97 | 0.56 | 0.87 | 0.93 | 0.95 |
| 4 July 2014 | 0.64 | 0.87 | 0.81 | 0.77 | 0.46 | 0.74 | 0.80 | 0.69 |
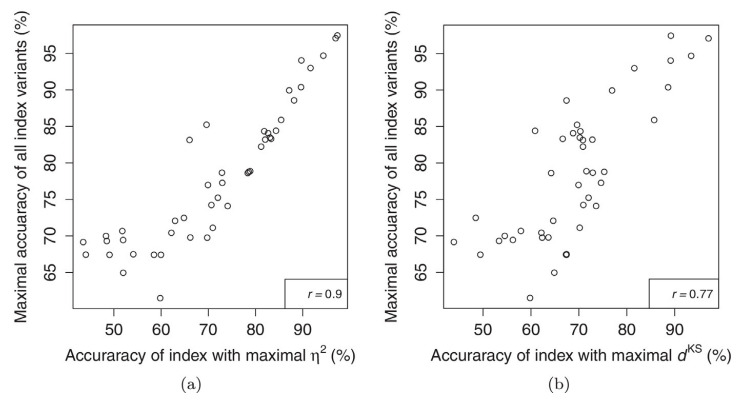


**Fig. 7.** Pearson's correlation coefficients and correlations between the maximum possible classification accuracy ($A_{max}$) and the accuracies using the $I^{perm}_{opt}$ determined by $\eta^2$ ($A(\eta^2_{max})$, sub-figure a) and Kolmogorov–Smirnov distance ($A(d^{KS}_{max})$, sub-figure b).

## 5. Discussion

### 5.1. Index permutation

In principle, the presented spectral optimization procedure aims at the automatic selection of a spectral index. In doing so, we calculate numerous index variants of a *EVI*-based formulation. This includes 2-band normalized difference indices but also more complex formulations including three bands and empirical parameters that are reported to be effective in land cover classification (Henebry et al., 2004; Viña et al., 2004). This is in contrast to the approach by Rivera et al. (2014) which enables the analysis of much more complex indices computed of up to 10 bands, but does not allow the integration of empirical weighting parameters. For the crop classifications based on multi-spectral imagery, we consider indices of higher complexity as inappropriate due to two practical reasons:

1 The majority of multi-spectral satellite sensors only have less than 10 bands, and those that have more, like Sentinel-2, contain spectral bands that are not meaningful for vegetation analysis (e.g., bands for coastal applications, aerosol, cirrus and water vapour retrieval).
2 For the index optimization for classification purposes, much larger sample sizes are required to balance intra-class variations due to growth status, soil properties, fractional vegetation cover and others. In our case, we analyzed approximately 500,000 pixels/samples per data set compared to 108 ground measurements used by Rivera et al. (2014). This leads to much higher computational requirements.

### 5.2. Index optimization for a selected satellite image

The index optimization applied on the data set acquired on 4th June 2014 improved the classification accuracies compared to $NDVI^{RE}$. As expected, the formulation of the optimized index is based on the red edge reflectance. The optimized index is calculated using the same bands as the also tested $EVI^{RE}$ (see Table 1). However, the weighting parameter configuration is different, and the $NDVI^{RE}$ was determined as best-suited standard index. Within the optimized index, the blue reflectance is highly accentuated by the weighting factor $c_1 = 7.5$, and the near infrared reflectance is suppressed compared to $EVI^{RE}$. This indicates significant differences in the blue band between barley and wheat that are most likely determined by grain and awn formation of barley as well as by chlorophyll degradation during ripening (see Section 4.1).

For the presented data set and test field selection, the optimization results in an index which is robust against within-field reflectance differences due to variations e.g., in soil water content and exposition. However, it has to be investigated in which extent the optimized index is fitted to the test data or whether classifications of larger areas can also be improved using the same number of test fields. If so, only a small number of test fields would be required to optimize the index and classify for example the complete Rapid-Eye tile or a composite of neighboring tiles.

### 5.3. Multi-annual separability profiles

### 5.3.1. Time frame detection

Although the minor separability maximum between November and December was detected in all five years, this time frame is more uncertain than the summer peak of separability. This is mainly due to the low sun angles during this time and potentially due to anthropogenic influences like the timing of tilling and sowing.

It is worth mentioning, that the blue band was included within the optimal band combination of nine acquisition dates (see Table 4), mostly in early summer. As ripening of barley proceeded further during that time than ripening of wheat, chlorophyll decomposition as consequence of ripening also proceeded. Consequently, chlorophyll absorption in the blue spectral region is lower for winter barley than for winter wheat which makes the blue band significant for the index optimization. However, atmospheric scattering is most relevant in the blue parts of the electromagnetic spectrum. For RapidEye data of farmland in the US, Moufid (2014) found that signal-to-noise ratio for the blue band is significantly lower than for the other spectral bands. Similar to Green et al. (1988) and Dwivedi and Rao (1992), we have tested the replacement of the blue band with other bands. To estimate the loss of information by replacing the blue band with others, inter-band correlations were computed. Pearson's correlation coefficients within the test fields computed for the data set from 4th June 2014 are 0.69 and 0.78 for $R_B$ vs. $R_G$ and $R_B$ vs. $R_R$, respectively. Consequently, clustering of the indices with band replacement leads to lowered accuracies of 0.85 and 0.82 for green and red, respectively, while the accuracy for the clustered $I_{opt}^{perm}$ is 0.97. Taking these results into account, a replacement appears to be not appropriate.

### 5.3.2. Influence of varying phenological development

Since differences in the timing of spring phenological phases influence the plant's reflectance profile (e.g., Förster et al., 2012; Boschetti et al., 2009; Duveiller et al., 2012; White et al., 2009), phenological variations are probably accounting for the inter-annual shifts in the timing of the summer separability peaks. The optimization for each data set instead of a single vegetation index allows the usage of an index that is sensitive to specific reflectance characteristics of crops during a specific growth stage. This becomes obvious by comparison of years with almost identical acquisition dates available, which is the case for 2010 (6th June and 3rd July) and 2014 (4th June and 4th July).

In 2010, the July image provides the maximal $\eta^2$, while in 2014 this is the case with the image acquired in June. This shift is also visible within the phenological observations data set (see Table 3). In 2014, beginning of heading of winter barley was observed 32 days earlier than in 2010 (23rd April vs. 25th May). For winter wheat, this difference is approximately 22 days. Yellow ripening was observed approximately 13 days earlier for both crops.

So the phase combinations are different between the years on the acquisition date in June. On 6th June 2010, winter wheat did not reach the stage of heading while for winter barley heading was observed before the image acquisition. During the heading of barley, long awns are built which lead in a brighter appearance that is also visible in the reflectance spectrum resulting most likely to the separability maximum that was observed in 2010. On 4th June 2014, for both crops heading was already observed, so winter wheat also shows a more yellowish color.

Yellow ripening of winter barley was observed 17 days before the image acquisition in 2014 but only four days before the 2010 image was acquired. This indicates, that ripening proceeded much further in 2014. The proceeded ripening is effective for the reflectance spectrum since it is characterized by declining chlorophyll and water content. Simultaneously, winter wheat did not reach the stage of yellow ripening in both years on the image acquisition dates. These two findings are probably accounting for the observed higher separability in July 2014 compared to July 2010.

### 5.4. Classification and validation

The classification of the optimized index was demonstrated to be more accurate than for standard vegetation indices. For the acquisition date of the highest spectral separability, the optimization derives an index, which is fitted to a subset of the RapidEye tile,

while the standard indices have been designed to be applicable on up to continental scales.

The validation framework tends to indicate higher classification accuracies than applied on natural scenarios of classification. This is mainly due to the fact that within the presented dichotomous classification scenario the class number is known and no pixels exist within the image data set which do not belong to one of the target classes. In more realistic classification scenarios the possible class number is much higher or even unknown which most likely increases classification uncertainty. However, for demonstration of the algorithm the laboratory-like conditions are constructive for the comparison of the permuted vegetation indices. The transferability of the approach to natural classification schemes has to be evaluated separately.

## 6. Conclusions and outlook

In this study we presented a vegetation index optimization methodology to find the best index for a dichotomous classification scenario. The index, of which the class-specific histograms of two crop types are as non-similar as possible, is detected out of a huge number of systematically calculated vegetation indices. The $\eta^2$ metric was introduced as spectral separability measure which is non-parametric and thus better suitable in selecting the optimal spectral bands for species discrimination than other separability measures which require normal distributions.

The $\eta^2$ statistic was successfully applied for the detection of the vegetation index that allows the most accurate separation of two crop types. Mathematically, $\eta^2$ can also be used as spectral separability indicator for a $n$-class classification scenario, but its performance for such a scenario has to be tested in the future.

The optimization approach was successfully applied on a Rapid-Eye image for spectral separation of winter wheat and winter barley. The values of the optimized index for the target crop types are highly species-dependent, which was demonstrated on classification results and histogram distribution of the index. We have also shown that the selected indices are more significant for discrimination of winter wheat and barley than widely used vegetation indices like *NDVI*, *EVI* and *SAVI*. The most significant spectral features are located between the red, red edge and near infrared parts of the spectrum.

Furthermore, we applied this optimization procedure in order to perform a multi-annual spectral separability analysis of winter wheat and winter barley. We detected two time frames, that lead to highest classification accuracies. These time frames are located in summer between June and July, and, with minor significance and higher uncertainty, between November and December. Consequently, a satellite image must be selected from the time series that was acquired during early summer to allow a most accurate classification. For winter cereals, a point in time on which at least one species already entered the phenological stage of heading is most promising but these findings are to be validated in future studies in areas of different climatic conditions. Further, the approach has to be tested for classification of other crop species.

Since Sentinel-2 will provide data of three red edge channels additional to the red and near infrared spectral regions, the application of the optimal index selection algorithm on Sentinel-2 time series has the capability to select sub-parts of the red edge region that are most significant for accurate discrimination of field crops.

If accurate phenological information, either spatio-temporal modeled using ground observations or derived from satellite data, is used as auxiliary variable, shifts in separability peaks can be explained exactly and typical growth stages, which allow most accurate classifications, can be determined.

## References

Anderson, G.P., Felde, G.W., Hoke, M.L., Ratkowski, A.J., Cooley, T.W., Chetwynd Jr., J.H., Gardner, J.A., Adler-Golden, S.M., Matthew, M.W., Berk, A., Bernstein, L.S., Acharya, P.K., Miller, D.P., Lewis, P.E., 2002. MODTRAN4-based atmospheric correction algorithm: FLAASH (fast line-of-sight atmospheric analysis of spectral hypercubes). In: AeroSense 2002. International Society for Optics and Photonics., pp. 65–71.

Bannari, A., Morin, D., Bonn, F., Huete, A.R., 1995. A review of vegetation indices. Remote Sens. Rev. 13 (1–2), 95–120.

Barnes, E.M., Clarke, T.R., Richards, S.E., Colaizzi, P.D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R.J., Li, H., Moran, M.S., 2000. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In: Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, MN, pp. 16–19.

Boschetti, M., Stroppiana, D., Brivio, P.A., Bocchi, S., 2009. Multi-year monitoring of rice crop phenology through time series analysis of MODIS images. Int. J. Remote Sens. 30 (18), 4643–4662.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. L. Erlbaum Associates.

Deutscher Wetterdienst, 2015. Vorschriften und Betriebsunterlagen für die phänologischen Beobachter des Deutschen Wetterdienstes. Offenbach, Germany.

Dorigo, W., Zurita-Milla, R., de Wit, A., Brazile, J., Singh, R., Schaepman, M., 2007. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. Int. J. Appl. Earth Obs. Geoinf. 9 (2), 165–193.

Duveiller, G., Baret, F., Defourny, P., 2012. Remotely sensed green area index for winter wheat crop monitoring: 10-year assessment at regional scale over a fragmented landscape. Agric. For. Meteorol. 166, 156–168.

Dwivedi, R.S., Rao, B.R.M., 1992. The selection of the best possible Landsat TM band combination for delineating salt-affected soils. Int. J. Remote Sens. 13 (11), 2051–2058.

Eitel, J.U.H., Vierling, L.A., Litvak, M.E., Long, D.S., Schulthess, U., Ager, A.A., Krofcheck, D.J., Stoscheck, L., 2011. Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. Remote Sens. Environ. 115 (12), 3640–3646.

Förster, S., Kaden, K., Förster, M., Itzerott, S., 2012. Crop type mapping using spectral–temporal profiles and phenological information. Comput. Electron. Agric. 89 (0), 30–40.

Fox, J., Friendly, M., Monette, G., 2015. heplots: Visualizing Tests in Multivariate Linear Models. R Package Version 1.0-16. http://CRAN.R-project.org/package=heplots.

Gitelson, A.A., 2004. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. J. Plant Physiol. 161 (2), 165–173.

Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sens. Environ. 58 (3), 289–298.

Green, A.A., Berman, M., Switzer, P., Craig, M.D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Trans. Geosci. Remote Sens. 26 (1), 65–74.

Guerschman, J.P., Paruelo, J.M., Bella, C.D., Giallorenzi, M.C., Pacin, F., 2003. Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. Int. J. Remote Sens. 24 (17), 3381–3402.

Henebry, G.M., Viña, A., Gitelson, A.A., 2004. The wide dynamic range vegetation index and its potential utility for gap analysis. GAP Anal. Progr. Bull. 12, 50–56.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25 (3), 295–309.

Huete, A.R., Justice, C.O., Van Leeuwen, W.J.D., 1999. MODIS Vegetation Index (MOD 13). Version 3. Algorithm Theoretical Basis Document. NASA, Goddard Space Flight Center, Greenbelt, MD (accessed May 07.05.11).

Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sens. Environ. 112 (10), 3833–3845.

Kolmogorov, A.N., 1933. Sulla determinazione empirica di una legge di distribuzione. Giorn. Istit. Ital. Attuari 4, 92–99.

Le Maire, G., Francois, C., Soudani, K., Berveiller, D., Pontaillier, J., Breda, N., Genet, H., Davi, H., Dufrene, E., 2008. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. Remote Sens. Environ. 112 (10), 3846–3864.

Liu, H.Q., Huete, A.R., 1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE Trans. Geosci. Remote Sens. 33 (2), 457–465.

Löw, F., Knöfel, P., Conrad, C., 2015. Analysis of uncertainty in multi-temporal object-based classification. ISPRS J. Photogramm. Remote Sens. 105, 91–106.

Massey, F.J., 1951. The Kolmogorov–Smirnov test for goodness of fit. J. Am. Stat. Assoc. 46 (253), 68–78.

Möller, M., Birger, J., Gidudu, A., Gläßer, C., 2013. A framework for the geometric accuracy assessment of classified objects. Int. J. Remote Sens. 34 (24), 8685–8698.

Möller, M., Müller, S., Doktor, D., Gläßer, C., 2012. Phenological structuring of multi-temporal RapidEye imagery. In: 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 4934–4937.

Moufid, T., 2014. SNR characterization in RapidEye satellite images (Master's thesis). Luleå University of Technology, Luleå, Sweden.

Murakami, T., Ogawa, S., Ishitsuka, N., Kumagai, K., Saito, G., 2001. Crop discrimination with multitemporal spot/HRV data in the Saga Plains, Japan. Int. J. Remote Sens. 22 (7), 1335–1348.

Nieweglowski, L., 2013. CLV: Cluster Validation Techniques. R Package Version 0.3-2.1. http://CRAN.R-project.org/package=clv.

Prishchepov, A.V., Radeloff, V.C., Dubinin, M., Alcantara, C., 2012. The effect of Landsat ETM/ETM+ image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. Remote Sens. Environ. 126, 195–209.

R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org/.

Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., Heitkönig, I.M.A., 2012. Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne RapidEye sensor. Int. J. Appl. Earth Obs. Geoinf. 19, 151–162.

Richards, J., 2012. Remote Sensing Digital Image Analysis: An Introduction. Springer Berlin Heidelberg.

Rivera, J., Verrelst, J., Delegido, J., Veroustraete, F., Moreno, J., 2014. On the semi-automatic retrieval of biophysical parameters based on spectral index optimization. Remote Sens. 6 (6), 4927–4951.

Rounds, E.M., 1980. A combined nonparametric approach to feature selection and binary decision tree design. Pattern Recognit. 12 (5), 313–317.

Rouse Jr., J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring Vegetation Systems in the Great Plains with ERTS., pp. 309, NASA special publication.

Schuster, C., Förster, M., Kleinschmit, B., 2012. Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. Int. J. Remote Sens. 33 (17), 5583–5599.

Simmons, A., Fellous, J.-L., Ramaswamy, V., Trenberth, K., Asrar, G., Balmaseda, M., Burrows, J.P., Ciais, P., Drinkwater, M., Friedlingstein, P., Gobron, N., Guilyardi, E., Halpern, D., Heimann, M., Johannessen, J., Levelt, P.F., Lopez-Baeza, E., Penner, J., Scholes, R., Shepherd, T., 2016. Observation and integrated earth-system science: a roadmap for 2016–2025. Adv. Space Res. 57 (10), 2037–2103.

Spatialreference, 2016. Catalogs of Spatial Reference Systems. http://spatialreference.org (accessed 22.02.16).

Swain, P.H., Davis, S.M. (Eds.), 1978. Remote Sensing: The Quantitative Approach, vol. 1. McGraw-ill, New York City, NY, USA.

Tang, Y., Zhang, L., Huang, X., 2011. Object-oriented change detection based on the Kolmogorov–Smirnov test using high-resolution multispectral imagery. Int. J. Remote Sens. 32 (20), 5719–5740.

Tyc, G., Tulip, J., Schulten, D., Krischke, M., Oxfort, M., 2005. The RapidEye mission design. Acta Astronaut. 56 (1–2), 213–219.

Vaiphasa, C., Ongsomwang, S., Vaiphasa, T., Skidmore, A.K., 2005. Tropical mangrove species discrimination using hyperspectral data: a laboratory study. Estuar. Coast. Shelf Sci. 65 (1–2), 371–379.

Van Niel, T.G., McVicar, T.R., 2004. Determining temporal windows for crop discrimination with remote sensing: a case study in south-eastern Australia. Comput. Electron. Agric. 45 (1–3), 91–108.

Viña, A., Henebry, G.M., Gitelson, A.A., 2004. Satellite monitoring of vegetation dynamics: sensitivity enhancement by the wide dynamic range vegetation index. Geophys. Res. Lett. 31 (4).

White, M.A., de Beurs, K.M., Didan, K., Inoue, D.Y., Richardson, A.D., Jensen, O.P., O'Keefe, J., Zhang, G., Nemani, R., van Leeuwen, W.J.D., Brown, J., de Wit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A., Kimball, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., Lauenroth, W.K., 2009. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. Glob. Change Biol. 15 (10), 2335–2359.

Wu, C., Niu, Z., Gao, S., 2012. The potential of the satellite derived green chlorophyll index for estimating midday light use efficiency in maize, coniferous forest and grassland. Ecol. Indic. 14 (1), 66–73.

## 4.2 CONCLUSIONS FROM PAPER 1

In Paper 1, it could be demonstrated that spectral features, which are employed as variables for classification, can be successfully optimised regarding their intended application. The systematic optimisation detects features that provide more species-specific separability information than a set of vegetation indices defined prior to the experiment.

As one major outcome it could be shown, that temporal windows of spectral separability rather depend on plant phenological phases than on calendar dates. The interpretation was of limited thematic depth due to shortcomings of the phenological observation design. Significant phenological differences were observed between short spatial differences, which reveals the need for more detailed phenological information for the area between phenological stations. A phenological model that is based on plant physiological responses to climatic differences to estimate a growth stage for regions between phenological stations is suggested to address these shortcomings.

<div style="text-align: right; font-size: 3em;">5</div>

# PHASE: A GEOSTATISTICAL MODEL FOR THE KRIGING-BASED SPATIAL PREDICTION OF CROP PHENOLOGY USING PUBLIC PHENOLOGICAL AND CLIMATOLOGICAL OBSERVATIONS

In Chapter 4, Paper 1 revealed the need for modelled phenological information to explain temporal differences in spectral response of spectrally similar land use classes. Consequently, a phenological model was developed and is presented in Paper 2, that utilises the dense data base of phenological information provided by the DWD. The presented model, called *Phenological model for Application in Spatial and Environmental sciences* (PHASE), is based on the established growing-degree-days approach (Réaumur 1735). All daily mean temperatures, that have been accumulated between the date of sowing of the respective crop type until a phenological phase was observed, are extracted from spatially interpolated observations. An indicator temperature sum is calculated from the resulting distribution and the Day of Year (DOY), on which this indicator sum is first exceeded during a growing season, is calculated for each raster cell of an underlying DEM. Here, the universal kriging geostatistical interpolation technique is used. The model results are compared to the simpler interpolation techniques ordinary kriging and inverse distance weighted

interpolation. Finally, on the example of soil moisture, the capability of the model to include additional explaining variables which also influence the phenological development of plants, was demonstrated.

CrossMark

Original papers

# PHASE: A geostatistical model for the Kriging-based spatial prediction of crop phenology using public phenological and climatological observations

H. Gerstmann [a,*], D. Doktor [b], C. Gläßer [a], M. Möller [c,a]

[a] Martin Luther University Halle-Wittenberg, Department for Remote Sensing and Cartography, Von-Seckendorff-Platz 4, 06120 Halle (Saale), Germany
[b] Helmholtz Centre for Environmental Research, Department for Computational Landscape Ecology, Permoserstraße 15, 04318 Leipzig, Germany
[c] Martin Luther University Halle-Wittenberg, Institute of Agriculture and Nutrition Science, Department of Farm Management, Karl-Freiherr-von-Fritsch-Straße 4, 06120 Halle (Saale), Germany

## ARTICLE INFO

## ABSTRACT

Detailed information on plant developmental stages, referred as phenological phases, can assist research, applications and synergies e.g., in land use, climate science and remote sensing. Usually, detailed ground information about phenological phases is only available as point observations. However, in most application scenarios of spatially interpolated phenological information is required. In this article, we present an approach for modeling and interpolation of crop phenological phases in temperate climates on the example of the total area of Germany using statistical analysis and a Kriging prediction process. The presented model consists of two major parts. First, daily temperature observations are spatially interpolated to retrieve a countrywide temperature data set. Second, this temperature information is linked to the day of year on which a phenological event was observed by a governmental observation network. The accumulated temperature sum between sowing and observed phenological events is calculated. The day on which the temperature sum on any location exceeds a phase-specific critical temperature sum, which indicates the day of entry of the modeled phase, is finally interpolated to retrieve a countrywide data set of a specific phenological phase. The model was applied on the example of eight agricultural species including cereals, maize and root crops and 37 corresponding phases in 2011. The results for most of the tested crops and phases show significantly lower *root mean squared errors* (*RMSE*) values and higher *goodness of fit* ($R^2$) values compared to results computed using Ordinary Kriging (OK) and Inverse Distance Weighting (IDW). The modeling accuracy varies between 2.14 days and 11.45 days for heading and emergence of winter wheat, respectively. The uncertainty of the majority of the modeled phases is less than a week. The model is universally applicable due to automatic parametrization, but model accuracies depend on the crop type and increase during a growing season. The possibility to enhance the model by additional explaining variables is demonstrated by consideration of soil moisture within an extended model setting.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Phenology studies periodic events in plant development and their dependence on shifting environmental factors such as temperature, day length and precipitation (Kirby et al., 1987; McMaster et al., 2009). Such events and phases are clearly visible developmental stages like blossoming or ripening (Schwartz, 2006).

The main climatic drivers of plant phenology vary in different ecoregions. Temperature is the main driving factor for intra-seasonal timing of phenological events in temperate regions like Central Europe (Chmielewski et al., 2004; Menzel, 2007). Many studies observed that in temperate climates the timing of phenological events is relatively stable and independent of other environmental factors than temperature (e.g., McMaster et al., 2009). Other factors influencing plant phenological development are photoperiod (Masle et al., 1989), daily temperature amplitude (Solantie, 2004), water availability and soil moisture especially in arid and semi-arid climates (McMaster and Wilhelm, 2003; Idso et al., 1978), solar radiation, distance to coasts and settlements, soil

properties (Zhao et al., 2013) and management factors like date of planting or fertilization practices (Nellis et al., 2009).

Crop phenology in Germany follows several spatial trends. Due to Germany's temperate climatic conditions, phenology is predominantly determined by temperature. Other factors influencing phenology in temperate regions to a lesser extent are precipitation and soil moisture, especially for autumn phases (Menzel, 2007), elevation, sea proximity and population density (Hense and Müller, 2007). Thus plant development in Germany is delayed in coastal and mountainous regions compared to the favored regions in south-western and central lowland regions (Siebert and Ewert, 2012).

Knowledge about plant phenological phases and their timing is of interest for wide application scenarios. Since plants react to changing temperatures and carbon dioxide content, long-term phenological time series can be used to monitor responses of plant phenology to global and regional warming (Estrella et al., 2007). Prevailing phenological information is also required for the assessment of famine risks and food production problems (Vrieling et al., 2011).

Several studies have also shown the potential of phenological information to support land cover classification of remote sensing images, models for crop yield estimation and precision farming on regional and continental scales (Van Niel and McVicar, 2004; van Bussel et al., 2011; Möller et al., 2012; Foerster et al., 2012; Prishchepov et al., 2012). Furthermore, phenology information has the potential to provide valuable input to soil erosion monitoring (Möller et al., 2015), mapping of biodiversity (Turner et al., 2003) or monitoring of invasive plant species (Bradley and Mustard, 2006; Huang and Asner, 2009).

Such support can be expected to continue to gain importance since the temporal availability of medium or high spatial resolution satellite sensors will considerably increase once the Sentinel-2 satellite constellation is working operational (Berger et al., 2012; Drusch et al., 2012). This requires reliable algorithms for data set selection in which phenology can play a major role to detect the most significant data sets for an image classification problem (Möller et al., 2012). In doing so, the required data amount is reduced with minimal loss in accuracy and thus enables an operational use of these data amounts both in environmental and agricultural sciences as well as in policy and decision making.

Detailed phenological data are mostly available as point observations of irregular spatial distribution which represent phenological phases in standardized numeric codes. Spatial information about phenological phases of crops can be also extracted from satellite images of high temporal resolution and corresponding vegetation indices provided for instance by Meteosat (Sobrino et al., 2013) and MODIS (e.g. Zhang et al., 2003; Lunetta et al., 2006; Jönsson et al., 2010; Xiao et al., 2013). However, these methods are mostly applied on only a few clearly visible phases like green-up or onset (Hird and McDermid, 2009).

The mentioned application scenarios require operationally effective and detailed phenological information. To produce such data, point observations have to be spatially interpolated using phenological models. Menzel (2007) and Zhao et al. (2013) distinguish three main types of phenological models:

1. Statistical fitting models which relate climatic variables to phenological development phases (e.g. McMaster and Wilhelm, 1997; Picard et al., 2005).
2. Mechanistic models that are based on cause-effect-relationships (Jamieson et al., 1998; Kramer et al., 2000; Ewert et al., 2002; Hänninen and Kramer, 2007).
3. Theoretical models which focus on plant physiological processes (Kaduk and Heimann, 1996; Schaber and Badeck, 2003; Peng et al., 2011).

Mechanistic and theoretical approaches require a large number of parameters and experimental effort. Statistical fitting methods only require a few input data sets, are of lower complexity and thus more frequently applied. One of the most often applied statistical fitting approach is based on the relation between the observation day of year of a phenological event ($DOY_{obs}$) and the corresponding accumulated effective temperature (Chuine et al., 2003; Hänninen and Kramer, 2007). This phase- and plant-specific temperature sum is usually referred as *growing degree days* (*GDD*), *heat units*, or *thermal time* (Zhao et al., 2013).

The majority of studies focused on either a region of limited extent or differences in plant parameters, mainly base temperature (Holen and Dexter, 1996; McMaster and Wilhelm, 2003), for different cultivars or cultivation sites of one crop type and between phenological phases (e.g. Wang and Engel, 1998; Ewert et al., 2002; Salazar-Gutierrez et al., 2013). A common problem is that the optimal starting day for *GDD* summation is difficult to determine (Wielgolaski, 1999). Furthermore, most of these studies do not combine phenological models and spatial interpolation since they often refer to pre-defined reference units (e.g. van Bussel et al., 2011; Siebert and Ewert, 2012).

To address these disadvantages, we present a framework which combines a geostatistical method and the *GDD* concept. In doing so, all critical parameters are extracted automatically and dynamically from the input data. After the geostatistical interpolation of daily mean temperatures, temperature sums and observed phenological phases are empirically related in order to extract the entry date of a specific phenological phase. These entry dates are again geostatistically interpolated to obtain area-wide predictions. The model has been designed to be easy-to-use, independent of expert knowledge, extendable, and transferable to any region of temperate climate where phenological observations and temperature measurements are available. The framework consisting of the combined model and the geostatistical interpolation was named PHASE (**PH**enological model for **A**pplication in **S**patial and **E**nvironmental sciences).

In this article, we describe the model structure, its underlying algorithms and methodological background (Section 3.2). We demonstrate its application on a selection of frequently grown crop types with special focus on winter wheat (*Triticum aestivum L.*) for the entire area of Germany (Section 3.2.3) using temperature data and phenological information provided on-demand for free by the German Weather Service.[1] The possibility to enhance the model by additional explaining variables is demonstrated by consideration of soil moisture within an extended model setting.

## 2. Materials and data

### 2.1. Phenological data

In Germany, the data base for phenological and meteorological observations is of unique density and quality and thus well-suited for model development. The German Weather Service (*German: Deutscher Wetterdienst* – DWD) operates a phenological monitoring network consisting of about 1200 active stations spread over Germany which report the Julian day of entry (day of year – *DOY*) for numerous phenological phases of agricultural crops, wine and natural plants at the end of each year (Hense and Müller, 2007). Each plant is observed on a different number of stations, depending on the abundance and agrometeorological relevance of the respective crop type. The observations are recorded by volunteers following standardized criteria, and a numeric code is assigned for each phase (Table 1).

---

[1] http://www.dwd.de.

**Table 1**

Numeric codes for observed phenological phases of agricultural crops (Deutscher Wetterdienst, 2015a).

| Phenological phase | Numeric code (phase ID) |
|---|---|
| Beginning of flowering | 5 |
| Full flowering | 6 |
| Beginning of sowing | 10 |
| Emergence | 12 |
| Closed stand | 13 |
| 4th leaf unfolded | 14 |
| Beginning of shooting/stem elongation | 15/67 |
| Beginning of bud formation | 17 |
| Beginning of heading | 18/66 |
| Beginning of milk ripening | 19 |
| Early dough ripening | 20 |
| Beginning of yellow ripening | 21 |
| Beginning of full ripening | 22 |
| Harvest | 24 |
| Beginning of tassel emergence | 65 |

A certain degree of quality control is already applied to these data, which is based on a multi-step cross validation procedure using ±25 days observation error as first threshold and ±15 days error in the second step (Hense and Müller, 2007). The most frequently cultivated crops in Germany (Statistisches Bundesamt, 2011) and the related, continuously reported phases are listed in Table 2. The annually collected phenological observations of these predominantly grown crops can be obtained via FTP server (Deutscher Wetterdienst, 2015a).

The distribution of the phenological stations is irregular both between regions of different environmental conditions and within regions of similar suitability for cultivation of a specific crop. For instance, in regions where winter wheat is cultivated frequently (Central Germany, and southern Germany apart from the mountainous regions) the point density varies between 1.5 and 3 stations per 1000 km², but no regional trend is visible. In mountainous and coastal regions, demanding crops cannot or only rarely be grown so that observation density is much lower. Additionally, the number and location of observations vary between different years due to crop rotation practices. Since the observers define the extent of their observation area according to the current cultivation status, positional accuracy of the phenological stations is approximately 2–5 km in position and 50 m in elevation (Deutscher Wetterdienst, 2015b). This limits the applicability for modeling phenology on medium spatial resolution, because it is impossible to consider temperature differences on mesoclimatic scale. Despite quality control is operationally applied to the data by the DWD, there are still outliers in the final product. In this study, a statistical filter is therefore applied which only considers observations within an interval of 1.5 standard deviations around the mean of the total data set. Exemplary, all stations that reported the phenological phase *beginning of yellow ripening* in 2011 are shown in Fig. 1b.

### 2.2. Temperature data

Daily mean temperatures were obtained for 503 stations per day on which the DWD is running daily automatic quality control procedures to detect and remove measurement errors. This error removal results in slight daily differences of the actual number of temperature observations per day. The positional uncertainty of the stationary temperature stations is assumed to be not significant. The temperature stations are not regularly distributed (Fig. 1b). Population density determines the network density as well as meteorological criteria. For instance, in mountainous areas where more short-term weather changes occur, the density is higher than in lowland areas of more constant weather conditions.

### 2.3. Elevation data

Since temperature is strongly dependent on elevation and topography, a digital elevation model (DEM) is used as explanatory variable for temperature interpolation. For this study, the *SRTM DEM* (USGS, 2004, see Fig. 1a) was selected. The data set has been produced by the space-borne STS-99 Shuttle Radar Topography Mission (SRTM) sensor and is freely available outside the United States at 90 m resolution with a vertical accuracy of 20 m and a horizontal accuracy of 16 m (Rabus et al., 2003). The DEM has been filtered following Lee (1980) to reduce signal noise. Elevations below sea level were set to 0 m.
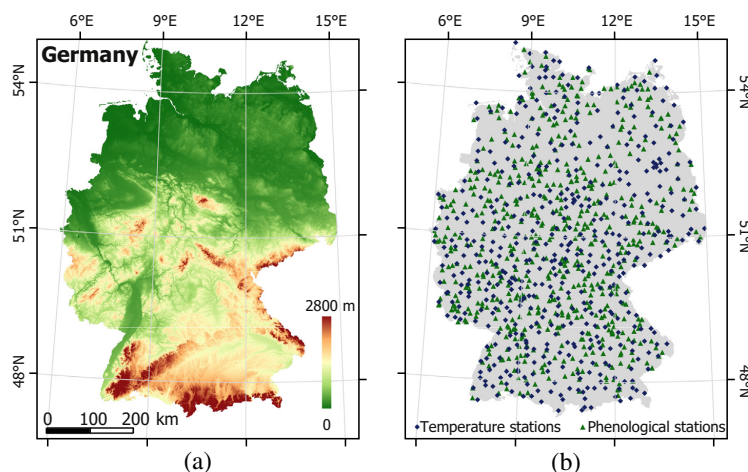


**Fig. 1.** Germany-wide input data sets. 1 km digital elevation model (a); all phenological and meteorological DWD stations that observed the phase *beginning of yellow ripening* for winter wheat in 2011 and all stations that provided daily temperature measurements on 1st January 2011 (b). Projection: EPSG code 25632 (Spatialreference, 2015). Data sources: USGS (2004) and Deutscher Wetterdienst (2015a).

The DEM raster cells were aggregated to $1 \times 1$ km pixel size with a total number of 358,320 pixels. The 1 km$^2$ resolution was chosen to consider the uncertainty in position and altitude of the phenological observations.

### 2.4. Soil moisture data

The DWD also provides daily raster data which represent the soil moisture in per cent of accessible field capacity for grass covering entire Germany (Deutscher Wetterdienst, 2015a). Spatial resolution of the raster data sets is $1 \times 1$ km. According to the data set publisher (DWD Climate Data Center (CDC), 2016), the data set is produced using observational data of approximately 280 synoptic climate stations. The soil moisture is calculated using a modified Penman–Monteith rationale for determination of evapotranspiration (Löpmeier and Deutscher Wetterdienst, 1983). The station-specific soil moisture values are interpolated using multiple linear regression and triangulation. No quality assessment routines are applied by the DWD, hence the data uncertainty cannot be quantified. Apart of the algorithm-dependent uncertainty, two further restrictions decrease the usability for the phenological model. First, sandy loam is used as soil type for entire Germany without taking regional differences into account. Second, the interpolated values for field capacity are only valid for unspecified grass vegetation.

## 3. Methods

### 3.1. Methodological background

#### 3.1.1. Phenological modeling

Numerous studies have shown that in temperate climates phenological development correlates with accumulated heat sums (Sitch et al., 2003; Miller et al., 2001; McMaster and Wilhelm, 1997; Russelle et al., 1984). They are calculated from the measured daily mean temperatures above a base temperature $T_B$, which is the lower threshold for photosynthetic activity. The underlying concept has been described by Réaumur (1735) and applied for prediction of crop phenology for centuries. This relation is often used for modeling plant phenology and is mostly referred as *growing degree day* (*GDD*):

$$GDD = 0.5 \times (T_{max} - T_{min}) - T_B \tag{1}$$

According to Eq. (1), a *GDD* is the daily contribution to the accumulated heat sum, $T_{max}$ and $T_{min}$ are the daily observed maximum and minimum temperatures, and $T_B$ is the base temperature. $T_B$ can be either species-specific or individual for each phase of a crop species (Slafer and Savin, 1991). The daily *GDD* – beginning at a starting point in time ($DOY_{start}$) – is then accumulated over a vegetation cycle. Because winter crops are sown during autumn of the previous year, the often-used 1st January is not necessarily the optimal starting day. This day does not take environmental differences into account which are expected to lead to inaccurate extraction of the critical temperature sum. Station-specific starting days like the day of sowing or the start of snow melt can be chosen instead to respect topographic or climatic differences. Since other factors than climate are also influencing the day of sowing, using snow melt as starting point is also inappropriate. Hence we expect the reported day of sowing (phase ID: 10, see Table 1) as most reliable starting day.

On the day on which the accumulated temperature sum exceeds a phase-specific threshold, the plant reaches the next stage of its phenological cycle. However, this relation is only applicable for phenological phases with a physiological background and not for the phases *sowing and tilling* or *harvest*.

Photosynthetic activity can start even on days with a negative contribution to *GDD*, if $T_{max}$ is higher than $T_B$ and $T_{min}$ is lower than $T_B$ resulting in a mean temperature below 0 °C (McMaster and Wilhelm, 1997). This can be especially the case during the early months of a growing season. To determine the portion of the day with temperatures higher than $T_B$, hourly temperature measurements are required. Since the use of hourly observations would increase the amount of required data rapidly and we assume the effect of on the modeling results to be negligible, we used observed daily mean temperatures.

#### 3.1.2. Kriging

Kriging is a widely-applied geostatistical method for interpolation of spatial data that was first presented by Krige (1951) for improving ore reserve estimations. It is an algorithm that uses the decreasing autocorrelation between two sample points with increasing spatial distance to predict intermediate values. The algorithm is based on an empirical variogram and a fitted variogram model to predict a variable at a location where no sample exist. Several theoretical variogram models exist, but the Matérn (Matérn, 1960) model is strongly suggested for the interpolation of spatial data (Stein, 1999) and can be used universally both for short and long distance variation models (Hengl, 2009). This variogram model includes a smoothing parameter *kappa* and a Gaussian model as limiting case as well as an exponential model as special case and thus is more flexible to local behavior of the observations than other models.

Kriging allows the consideration of possible correlations of the predicted variable with various explanatory variables (Hengl et al., 2007). For instance, this is the case for air temperature, which shows a dependency to elevation and surface topography. Kriging using explanatory variables is referred as Universal Kriging, Kriging with external drift or Regression Kriging. Other methods like Inverse Distance Weighting interpolation, averaging values per polygon and Ordinary Kriging can be understood as special cases of the Regression Kriging method (Hengl, 2009), which hence is understood as best linear unbiased predictor (Stein, 1999; Hengl, 2009).

#### 3.1.3. Cross validation

Cross validation is an often applied method for accuracy assessment for interpolations of environmental variables (Kuhn and Johnson, 2013). It compares predicted values on a location with the observed values on the same location. Leave-one-out cross validation excludes one sample from the input data set and predicts this sample using the surrounding observation points. Afterwards, the difference between the predicted value and the observed value is used as estimate of the prediction performance. 10-fold cross validation is an alternative technique, which is a variant of bootstrapping and is insensitive against outliers (Hengl, 2009). Here, the data set is split into 10 parts (folds) of equal size and each fold is used for cross validation and calculation of the parameters *Root Mean Squared Error* (*RMSE*) and $R^2$ as metrics for the model's *goodness of fit*. *RMSE* is a measure to describe the difference between observed and modeled values and is calculated according to Eq. (2).

$$RMSE = \sqrt{\frac{(y_{obs} - y_{pred})^2}{N_{obs}}} \tag{2}$$

Here, $y_{obs}$ and $y_{pred}$ represent observed and modeled values and $N_{obs}$ is the number of observations. Cross validation provides global accuracy measures as well as the residuals for each observation which can be used to detect observations of high uncertainty that in turn can affect the entire prediction result negatively.

## 3.2. Basic model conception

The core PHASE model follows the growing degree days approach (see Eq. (1)) but includes day length as proxy for latitude influences. It determines the number of required heat units to reach a phenological phase, which is representative for the complete study area. First, a station- and phase-specific threshold of required heat units is determined (Eq. (3)).

$$T_{sum}^{eff}[j] = \sum_{i=DOY_{start}}^{DOY_{obs}} \left( (\overline{T}_{i,j} - T_B) \times \frac{DL_i}{24} \right) \tag{3}$$

$T_{sum}^{eff}[j]$ in heat units (HU) represents the accumulated effective temperatures on a specific phenological station $j$ between a fixed starting day $DOY_{start}$ and the observed phenological event $DOY_{obs}$. $\overline{T}$ is the daily mean temperature and $DL$ the day length at the phenological stations, $T_B$ is the base temperature.

Next, a critical HU value $T_P$ is determined as a quantile of the distribution of all values of $T_{sum}^{eff}$ (Eq. (4)).

$$T_P = Q_{opt}(T_{sum}^{eff}) \tag{4}$$

$T_P$ is the critical temperature sum required for a specific phase and $Q_{opt}$ is the optimal quantile of $T_{sum}^{eff}$ of all phenological stations. Finally, the first day $DOY_P$ on which the cummulated heat units are equal or exceed the critical temperature sum $T_P$ is calculated for each station.

The work flow in Fig. 2 can be distinguished in three parts:

1. Data pre-processing is performed which includes the interpolation of daily temperature observations to retrieve a daily mean temperature raster data set covering entire Germany ($\overline{T}_{i,j}$ in Eq. (3); see Section 3.2.1) and the outlier detection of the phenological observations,
2. Fitted values of the parameters of Eqs. (3) and (4) are estimated by finding the configuration that leads to the lowest possible *RMSE*, resulting in the calculation of accumulated daily effective temperatures ($T_{sum}^{eff}$).
3. $DOY_P$ is determined and interpolated to produce a raster data set representing the day of entry of the target phenological phase. Additionally, cross-validation is performed to assess the prediction accuracy (Section 3.3).

Due to the automatic fitting procedures, the model is completely data-driven and no prior assumptions on the data distribution are required for modeling.

The model has been implemented within the statistical computing environment *R* (R Core Team, 2015). All geodata are imported using the *R*-packages *rdgdal* (Bivand et al., 2015) and *raster* (Hijmans, 2015).

### 3.2.1. Temperature interpolation

The daily mean temperatures are interpolated using local Universal Kriging (see Section 3.1.2) which is implemented within the *gstat* package (Pebesma, 2004). Elevation is used as independent variable to respect topography influences in the prediction. A Matérn variogram model is automatically fitted to the empirical semivariogram using the function *fit.variogram()* which is also included in the *gstat* package. A crucial parameter is the number of stations to include in the temperature prediction at a certain location (*nmax*). In principle, the higher *nmax* the more accurate predictions can be assumed. On the other hand, Kriging is based on the assumption that more distant observations are almost ineffective for the prediction result (Hengl, 2009). Since computation time increases strongly with increasing number of considered
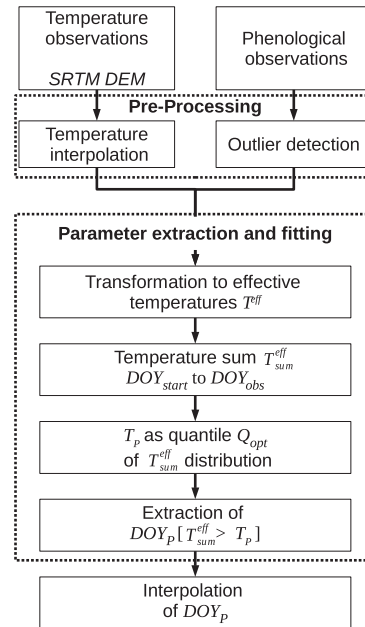


**Fig. 2.** Structure of the basic model including preprocessing, fitting and spatial interpolation. *SRTM DEM* - SRTM digital elevation model; $T_{sum}^{eff}$ - accumulated temperatures from $DOY_{start}$ to observed day $DOY_{obs}$; $T_P$ - required heat units; $Q_{opt}$ - Quantile of $T_{sum}^{eff}$ distribution that leads to lowest RMSE values; $DOY_P$ - day on which accumulated temperature sum exceeds $T_P$.

stations, the number of the included stations should be reduced to limit processing time.

### 3.2.2. Calculation of effective temperatures

The interpolated mean temperatures $\overline{T}$ are converted to station-specific daily effective temperatures $T^{eff}$ (Eq. (3), comparable with *GDD* in Eq. (1)). First, the base temperature is subtracted from the mean temperature of each day. Next, the resulting temperatures are adjusted to take the dependence of photosynthesis on available sunlight (photoperiod) into account. Here, the observed mean temperature are multiplied by the ratio of day length (*DL* [h]) to the whole day.

*DL* is calculated using the *R* function *daylength()* that is included in the package *geosphere* (Hijmans et al., 2014). $T_B$ is the base temperature below which no growth occurs. Although different values for $T_B$ exist, the most often applied default values which are used independently of crop type and region are 0 °C, 5 °C or 10 °C. The actual base temperature for a certain plant or cultivar and phase can be also calculated by statistical formulas (Yang et al., 1995) or selected using standard look-up-tables. In this study, a sequence of $T_B$ values is tested ranging from 0 °C to 10 °C.

### 3.2.3. Extraction of $DOY_P$

The plant- and station-specific critical temperature sum ($T_{sum}^{eff}$) for a phenological phase on a specific station is calculated by summing up all daily temperatures from a starting point $DOY_{start}$ to the day on which the phenological phase has been observed ($DOY_{obs}$). In this study, $DOY_{start}$ is defined by the reported sowing day. All negative effective temperatures are considered as not significant for photosynthetic activity.

The day $DOY_P$, on which the cumulative sum exceeds the indicator temperature sum for the entry of a target phase $T_P$, that is used as indicator for the entry of a target phase, is calculated for each station. $T_P$ is automatically detected by testing 19 quantiles ($Q \in [0.05, 0.95]$) in steps of 5% of the $T^{eff}_{sum}$ distribution. The quantile $Q_{opt}$ which leads to the lowest $RMSE$ between observed and estimated $DOY$ is used for determination of $T_P$.

### 3.2.4. Spatial interpolation of $DOY_P$

The modeled days are passed to a Kriging interpolation with elevation as independent variable using an automatically fitted Matérn variogram model (see Section 3.2.1). The initial values for fitting were set according to Hiemstra et al. (2009), the smoothing parameter *kappa* was set to 1. Here, the optimal value for *nmax* is determined using a 10-fold cross validation technique, and the *nmax* value with the lowest *RMSE* is used for the final prediction.

### 3.3. Accuracy assessment

An estimate of uncertainty is provided as part of the Kriging procedure. This estimate is referred as Kriging variance and represents the variance of the predicted result on each location in comparison to the optimum. The Kriging variance can be used to assess spatial patterns in uncertainty of the prediction.

*RMSE* and $R^2$ were computed using 10-fold cross validation (see Section 3.1.3). For comparison, *RMSE* and $R^2$ achieved by Inverse Distance Weighting interpolation (IDW) and Ordinary Kriging (OK) applied on the raw phenological observation data using a Matérn variogram model and the *fit.variogram()* function were computed.

### 3.4. Model extension using soil moisture

The model can be extended by further parameters that influence plant phenological development, e.g., soil moisture. Although soil moisture is less effective for modeling plant development in temperate climates than temperature (Fu et al., 2014), it is of special importance for the model's transferability to other, especially arid and semi-arid climates where soil water availability more influences plant development. Independently of the data limitations mentioned in Section 2.4 and the, for the test site with its temperate climate, expected low significance for the modeling results, a second normalization of the daily effective temperatures ($T^{eff}$) was tested (Eq. (5)).

$$T^{eff} = (\overline{T} - T_B) \times \frac{DL}{24} \times \frac{1 - \cos(\pi \times \theta_{fc})}{2} \qquad (5)$$

Here, $T^{eff}$ is the daily contribution of a day to $T^{eff}_{sum}$. $\overline{T}$ and $T_B$ are the mean and base temperature, $DL$ is the length of daylight and $\theta_{fc}$ is the field capacity representing the soil moisture accessible by plants. The additional normalization term is formulated to equal 1 when the plant available field capacity is 100% and declines in conditions of stress due to soil wetness and drought, so that water deficits and surplus reduce the daily contribution of $T^{eff}$ to $T^{eff}_{sum}$. In doing so, a 10% deficit as well as 10% surplus, for instance, result in the weighting factor of 0.976. Consequently, warm days with many day light hours and 100% plant available field capacity are understood as the most effective days for plant development.

## 4. Results and discussion

The model has been applied for all crop types and corresponding phenological phases with a plant physiological background (see Table 2) that were observed in 2011. This includes the spring and summer phases of crops that were harvested in 2011, and the

**Table 2**
Reported phenological phases for the eight typical field crops in Germany (DWD, 2015). See Table 1 for full phase names.

| Crop | Crop (scientific name) | Reported phases |
| --- | --- | --- |
| Winter wheat | *Triticum aestivum* L. | 10, 12, 15, 18,19, 21, 24 |
| Winter rye | *Secale cereale* L. | 10, 12, 15, 5, 6, 18, 21, 24 |
| Winter barley | *Hordeum vulgare* L. | 10, 12, 15, 18, 21, 24 |
| Oilseed rape | *Brassica napus* L. | 10, 12, 14, 67, 17, 5, 22, 24 |
| Oat | *Avena sativa* L. | 10, 12, 15, 66, 19, 21, 24 |
| Maize | *Zea mays* L. | 10, 12, 67, 5, 65, 19, 20, 21, 24 |
| Potato | *Solanum tuberosum* L. | 10, 12, 5, 24 |
| Sugar beet | *Beta vulgaris* subsp. *vulgaris* | 10, 12, 13, 24 |

autumn phases of the next vegetation cycle that was terminated in 2012. The results are discussed with special focus on the phases *beginning of shooting* and *beginning of yellow ripening* of winter wheat, which is the most frequently grown crop type in Germany, with the exception of perennial grasslands. In 2011, winter wheat was grown on 19% of the total area under cultivation (Statistisches Bundesamt, 2011), followed by maize (13%), winter barley (9%) and winter oilseed rape (8%). Fig. 3 shows the corresponding *DOY* distributions for the total area of Germany whose outliers were removed by statistical filtering (see Section 2.1).

### 4.1. Temperature interpolation

Daily temperatures between 2010 and 2011 were interpolated covering all days for which observations exist. According to Hengl (2009), 30–60 included observations result in sufficient accuracy of the final prediction, while Webster and Oliver (2007) stated that 100–150 stations are required at least. A 10-fold cross validation has been applied on all days to determine the number of stations to include in the temperature prediction at a certain location (*nmax*) by subsequent testing of values from 10 to 20, 25, 30, 40, 50, 60, 100 and 200 (Section 3.2.1). The processing times for the predictions of one single day are displayed in Fig. 4. In order to balance prediction accuracy and computation time, the *nmax*

Mean = 111    Mean = 195
Min = 94      Min = 180
Max = 130     Max = 211

*Shooting*    *Yellow ripening*
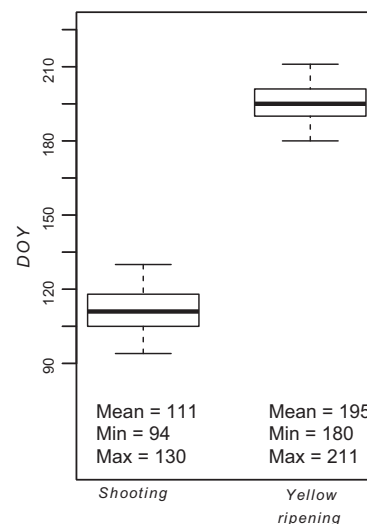
**Fig. 3.** Boxplots of the observation days for winter wheat phases *beginning of shooting* and *beginning of yellow ripening* in 2011 for the total area of Germany. *DOY* = 111 represents 21st April and *DOY* = 195 corresponds to 14th July 2011.
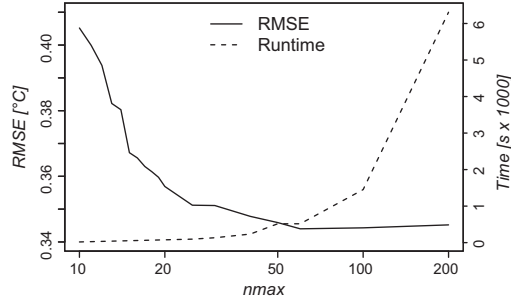
**Fig. 4.** Mean *RMSE* values of all temperature interpolations derived by 10-fold cross validation using different values for *nmax* (solid line) and processing time of corresponding Kriging interpolation (dashed line).

value was set to 30 for all daily interpolations which provides only slightly higher *RMSE* values than the optimum in most of the test cases while it needs significantly less processing time of about 134 s on a standard PC (4 GB RAM, 2.7 GHz, Microsoft Windows 7OS 64bit). The calculated *RMSE* values for different values for *nmax* vary about 0.06 °C with a decreasing trend to higher *nmax* values. The *RMSE* values for all temperature interpolations are below 1 °C on 89% of the days.

### 4.2. Extraction of $DOY_P$ and model calibration

#### 4.2.1. Base temperature

Daily temperatures for 2011 were transformed to $T^{eff}$ by normalization regarding photoperiod and subtraction of the base temperatures that have been determined iteratively by testing all values between 0 °C and 10 °C. As Slafer and Rawson (1995) stated for winter wheat, both base temperature as well as sensitivity to photoperiod can vary between different cultivars of a species. The base temperature that leads to the lowest errors in the final prediction was used. Despite $T_B = 0$ °C was used in most of the cases (see Table 3), significant differences exist between both crops and also phases of an individual crop. According to e.g. Salazar-Gutierrez et al. (2013), $T_B$ can vary between 1.6 and 8.4 °C for different winter wheat cultivars. Nevertheless, $T_B$ can be considered as a statistical variable in this study which enables the optimization of model results.

In Fig. 5, the mean errors of the predictions are plotted in relation to the tested base temperatures. The errors for *shooting* clearly exceed the errors for *yellow ripening*. For *yellow ripening*, the base temperature is less decisive for values between 0 °C and 6 °C than for *shooting*, for which the errors increase almost linearly with increasing base temperatures. The stable errors for *yellow ripening* for $T_B$ between 0 °C and 6 °C correspond to the most often reported $T_B$ values for winter wheat, which are between 0 °C and 5.5 °C (McMaster and Wilhelm, 2003; Barrett, 2013). The almost linear

**Table 3**
Optimal model parameters and accuracy metrics for crop phenological phases observed in 2011.

| Plant | Phase | $T_B$ [°C] | $Q_{opt}$ | $T_P$ [HU] | nmax | RMSE | | | $R^2$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Mod | Raw | IDW | Mod | Raw | IDW |
| Winter wheat | Emergence | 0 | 0.35 | 58 | 60 | 11.45 | 9.33 | 9.60 | 0.89 | 0.15 | 0.11 |
| | Shooting | 0 | 0.50 | 423 | 40 | 5.75 | 8.87 | 9.07 | 0.65 | 0.01 | 0.02 |
| | Heading | 10 | 0.50 | 118 | 40 | 2.14 | 4.96 | 5.05 | 0.72 | 0.21 | 0.18 |
| | Milk ripening | 5 | 0.45 | 588 | 50 | 3.57 | 8.80 | 8.92 | 0.65 | 0.03 | 0.04 |
| | Yellow ripening | 0 | 0.45 | 1265 | 14 | 3.83 | 7.11 | 7.04 | 0.70 | 0.08 | 0.10 |
| Winter rye | Emergence | 0 | 0.40 | 55 | 60 | 8.58 | 8.24 | 8.28 | 0.82 | 0.03 | 0.04 |
| | Shooting | 0 | 0.55 | 413 | 18 | 5.92 | 6.98 | 7.14 | 0.79 | 0.01 | 0.02 |
| | Beg. of flowering | 10 | 0.50 | 116 | 30 | 2.75 | 5.53 | 5.61 | 0.69 | 0.18 | 0.16 |
| | Full flowering | 10 | 0.50 | 135 | 60 | 3.09 | 6.63 | 6.88 | 0.62 | 0.09 | 0.05 |
| | Heading | 8 | 0.55 | 141 | 50 | 3.66 | 5.05 | 5.10 | 0.70 | 0.14 | 0.13 |
| | Yellow ripening | 0 | 0.45 | 1283 | 50 | 4.51 | 8.47 | 8.64 | 0.61 | 0.11 | 0.08 |
| Winter barley | Emergence | 0 | 0.45 | 73 | 40 | 5.04 | 5.64 | 5.87 | 0.81 | 0.09 | 0.05 |
| | Shooting | 0 | 0.55 | 451 | 50 | 5.19 | 7.36 | 7.45 | 0.70 | 0.05 | 0.05 |
| | Heading | 0 | 0.50 | 632 | 60 | 3.59 | 4.45 | 4.66 | 0.68 | 0.25 | 0.18 |
| | Yellow ripening | 6 | 0.50 | 1064 | 50 | 3.88 | 7.79 | 7.96 | 0.70 | 0.03 | 0.02 |
| Oilseed rape | Emergence | 0 | 0.45 | 107 | 60 | 7.16 | 7.10 | 7.30 | 0.81 | 0.04 | 0.02 |
| | 4th leaf unfolded | 0 | 0.40 | 260 | 60 | 7.98 | 8.88 | 9.15 | 0.67 | 0.01 | 0.01 |
| | Shooting | 0 | 0.75 | 603 | 25 | 4.90 | 5.16 | 5.46 | 0.72 | 0.22 | 0.10 |
| | Bud formation | 0 | 0.70 | 633 | 60 | 5.90 | 5.45 | 5.56 | 0.71 | 0.19 | 0.16 |
| | Full flowering | 0 | 0.55 | 694 | 50 | 4.06 | 4.23 | 4.30 | 0.67 | 0.37 | 0.35 |
| | Full ripening | 0 | 0.45 | 1484 | 20 | 4.90 | 8.01 | 8.33 | 0.64 | 0.12 | 0.07 |
| Oats | Emergence | 0 | 0.40 | 66 | 60 | 5.04 | 6.00 | 6.14 | 0.81 | 0.15 | 0.13 |
| | Shooting | 1 | 0.50 | 284 | 60 | 4.12 | 8.66 | 9.11 | 0.63 | 0.02 | 0.00 |
| | Heading | 0 | 0.45 | 541 | 50 | 3.23 | 6.24 | 6.57 | 0.66 | 0.03 | 0.01 |
| | Milk ripening | 0 | 0.45 | 855 | 18 | 3.29 | 8.78 | 9.26 | 0.57 | 0.02 | 0.01 |
| | Yellow ripening | 0 | 0.45 | 1059 | 25 | 3.43 | 8.31 | 8.62 | 0.65 | 0.08 | 0.04 |
| Maize | Emergence | 3 | 0.45 | 79 | 30 | 5.06 | 5.14 | 5.25 | 0.82 | 0.18 | 0.14 |
| | Shooting | 8 | 0.45 | 160 | 40 | 3.24 | 9.19 | 9.54 | 0.40 | 0.01 | 0.00 |
| | Tassel emergence | 3 | 0.45 | 625 | 60 | 3.20 | 5.72 | 5.86 | 0.65 | 0.16 | 0.12 |
| | Beg. of flowering | 0 | 0.45 | 902 | 60 | 3.68 | 7.10 | 7.18 | 0.63 | 0.11 | 0.10 |
| | Milk ripening | 0 | 0.45 | 1211 | 60 | 3.96 | 7.64 | 7.81 | 0.67 | 0.09 | 0.06 |
| | Early dough ripening | 0 | 0.45 | 1364 | 30 | 4.30 | 8.24 | 8.59 | 0.68 | 0.13 | 0.07 |
| | Yellow ripening | 0 | 0.40 | 1468 | 19 | 5.32 | 8.33 | 8.55 | 0.70 | 0.05 | 0.04 |
| Potato | Emergence | 0 | 0.45 | 172 | 30 | 5.75 | 7.35 | 7.23 | 0.73 | 0.07 | 0.09 |
| | Closed stand | 4 | 0.45 | 295 | 20 | 3.88 | 7.05 | 7.15 | 0.64 | 0.10 | 0.08 |
| Sugar beet | Emergence | 3 | 0.45 | 73 | 30 | 5.35 | 10.63 | 7.61 | 0.71 | 0.00 | 0.02 |
| | Closed stand | 3 | 0.50 | 421 | 40 | 3.21 | 7.43 | 7.94 | 0.47 | 0.06 | 0.01 |
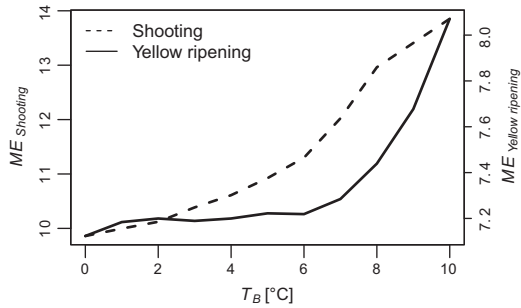
**Fig. 5.** Mean errors between $DOY_P$ and $DOY_{obs}$ for different base temperatures.
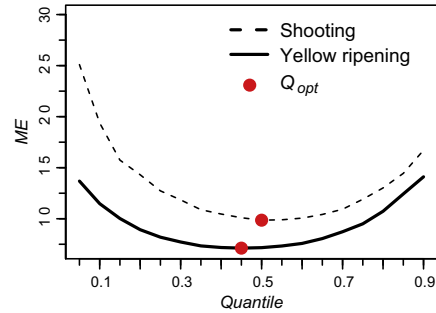


**Fig. 7.** Mean errors between $DOY_{obs}$ and $DOY_P$ and the optimal quantiles of the $T_{sum}^{eff}$ distributions.

increase of the mean error for *shooting* is probably caused by the exclusion of days with a mean temperature below $T_B$ but photosynthetically effective hours (McMaster and Wilhelm, 1997). This effect decreases for summer phases like *yellow ripening*, but is visible in all further modeling steps (Table 3).

*4.2.2. Temperature sums and optimal quantiles*

For each phenological station and day, $T^{eff}$ for *shooting* and *yellow ripening* of winter wheat were calculated to retrieve the accumulated temperature sum $T_{sum}^{eff}$ between $DOY_{start}$ and $DOY_{obs}$ by using $T_B = 0\,°C$. According to Fig. 6, the median of $T_{sum}^{eff}$ for *shooting* and *yellow ripening* are 423 HU and 1273 HU with standard deviations of 84.12 and 102.07, respectively.

The $DOY_P$ value on which the temperature sum exceeds each of the tested quantiles was determined. The mean errors of the observed and modeled values and corresponding $Q$ values are shown in Fig. 7. Accordingly, $Q_{opt}$ is equal to the expected median for *shooting* and slightly below for *yellow ripening* where $Q_{opt}$ equals the 45%-quantile.

All data points where the residual of $DOY_P$ and $DOY_{obs}$ is larger than the standard deviation $\sigma$ of $DOY_{obs}$ were excluded from the data set. This filtering process reduces the considered stations by about 49% and 41% respectively (see Fig. 8). As a result, the model fitting increases significantly. $R^2$ and $RMSE$ for the filtered point data set are 0.65 and 5.75 for *shooting* and 0.70 and 3.83 for *yellow ripening*. Point scattering is higher for *shooting* than for *yellow*
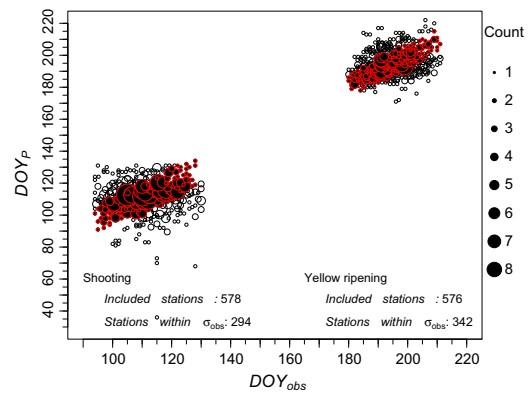


**Fig. 8.** Density plot of $DOY_P$ and $DOY_{obs}$ for *shooting* (left) and *yellow ripening* (right) of winter wheat in 2011. The diameter of the points indicates the frequency of $DOY_{obs}$-$DOY_P$-combinations. All data points within the standard deviation $\sigma$ are highlighted in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

*ripening*, which indicates a better model performance for the prediction of summer phases than on spring phases due to less uncertainty in the calculation of $T^{eff}$ mentioned in Section 4.2.



**Fig. 6.** Distribution of $T_{sum}^{eff}$ calculated for all phenological stations from sowing to the observation date of *shooting* (a) and *yellow ripening* (b) for winter wheat in 2011.

For all other phases and plants, the optimal quantiles as well as $T_B$ and *RMSE* values are listed in Table 3.

### 4.3. Interpolation of $DOY_P$

The model results and Kriging uncertainties for the phenological phases *shooting* and *yellow ripening* of winter wheat are shown in Fig. 9. To retrieve an uncertainty measure in days, the square root of the Kriging variance (*Kriging standard de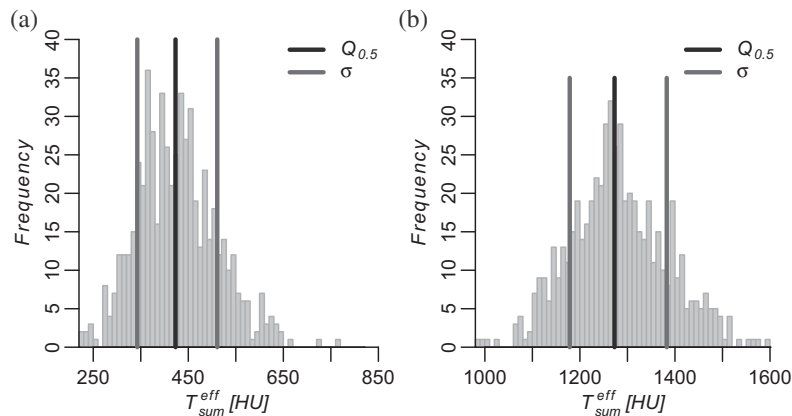viation*) is plotted. Relatively early dates were modeled for the Rhine valley and the favored regions in eastern Germany, while in mountainous and coastal regions the plant development is delayed. In some regions, the distribution of the observations is relatively coarse (see Fig. 1 and Section 2.1), which causes higher uncertainties (>6 days) compared to regions of dense observations (≈4 days). The prediction in mountainous regions is of high uncertainty (>8 days) for both phases.

The prediction is performed even in areas where a particular crop cannot be grown due to unsuitable environmental conditions. For instance, winter wheat can only be grown in regions with sufficient precipitation, mild winters and warm summers as well as on fertile soils which are not present in entire Germany. However, the model produces information independently of these requirements. This problem is particularly obvious in north-eastern Germany. There, the soils are very sandy and thus not very fertile, and winter wheat is consequently only rarely cultivated and almost no observations exist from this area. However, the Kriging algorithm predicts the phase using the closest locations, in this case from the very fertile regions in Central Germany. Consequently, regions of actually poor environmental conditions appear to be favored for wheat cultivation. The integration of a mask derived from soil types or detailed land use information that margins the prediction on potential cultivation areas could solve this problem in the future.

### 4.4. Optimal parametrization and model accuracy

The results of the applied 10-fold cross validation (see Section 3.3) are listed in Table 3, where the determined optimal model parameters *nmax*, $Q_{opt}$ and $T_P$ and the accuracy metrics $RMSE_{mod}$ and $R^2_{mod}$ of the interpolated model output for all tested crops and phases are combined. The metrics $RMSE_{raw}$ and $R^2_{raw}$,
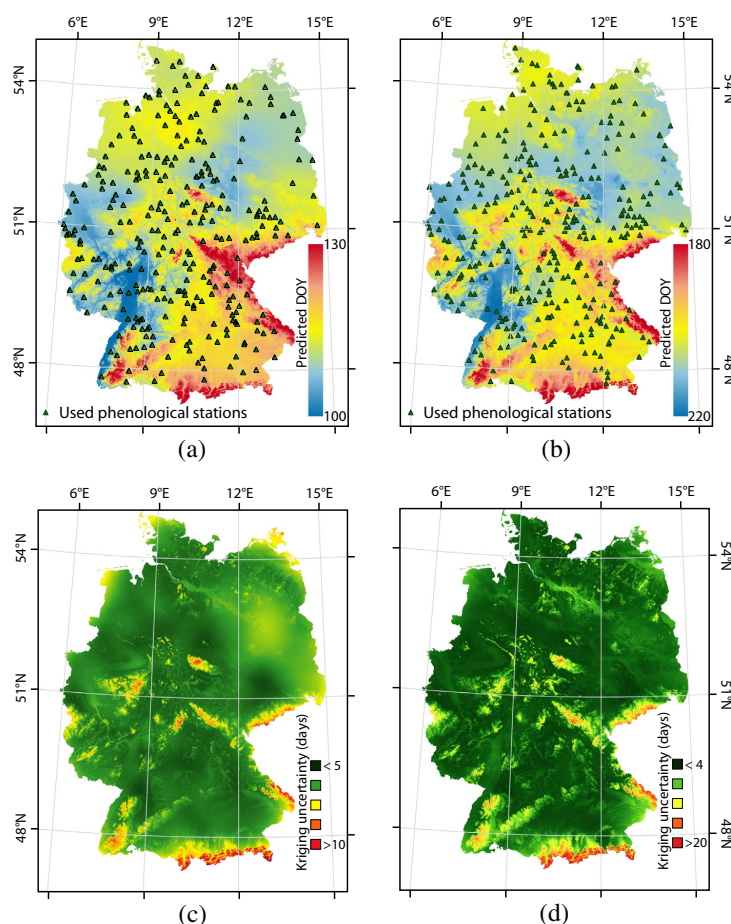


**Fig. 9.** Interpolated modeling results for the phenological phases *beginning of shooting* (a) and *yellow ripening* (b) of winter wheat in 2011 and corresponding Kriging standard deviations (c and d). Projection: EPSG code 25632 (Spatialreference, 2015).

derived by Ordinary Kriging of the raw phenological observations and $RMSE_{IDW}$ and $R^2_{IDW}$, calculated using the widely applied Inverse Distance Weighting interpolation for the original and unfiltered observations, are added for comparison.

$Q_{opt}$ is mostly ±5% around the median, with the exceptions of *emergence* of winter wheat, winter rye and oats. In case that $T_B$ is differing between two subsequent phases, $T_P$ can be lower for the later than for earlier phase. This is the case for the phases *shooting* and *emergence* of winter wheat, for which $T_B$ differs by 10 °C. However, the determined $T_P$ in general correlates to *GDD* values published in the literature (e.g., Miller et al., 2001). The modeling of the early phases shows higher $RMSE_{mod}$ values and smaller differences to $RMSE_{raw}$ values compared to later phases, while no declining trend is visible for $RMSE_{raw}$ values. Since *emergence* occurs only a few days after sowing, this phase depends highly on management practices and less on temperature, and the model accuracy is consequently reduced for this phase. $RMSE_{mod}$ is higher than $RMSE_{raw}$ for some phases e.g., for *emergence* of winter wheat, winter rye, oilseed rape and potatoes, which indicates deficits of the model for early phases.

Oilseed rape phases are modeled with lower accuracy than other crops. For two phases, *shooting* and *bud formation*, the optimal quantile is significantly above the median (75% and 70%, respectively). Additionally, $RMSE_{mod}$ for *bud formation* is higher than $RMSE_{raw}$ and $RMSE_{IDW}$. This higher uncertainty indicates a stronger dependency of their phenological development on factors which are not considered within the model like comparatively high requirements to nitrogen, potassium and sulfur nutrition (Grant and Bailey, 1993) or fertilization.

### 4.5. Integration of soil moisture

For each of the 1 × 1 km DEM and temperature raster cell, the corresponding soil moisture value was extracted for every day and used for calculation of the modified effective temperature term (see Eq. (5)). The consideration of the soil moisture normalization term for the calculation of $T^{eff}$ resulted in reduced *RMSE* for 18 of the modeled phases, while the remaining 19 showed no effect or slightly increased *RMSE*. The improvement is highest for *emergence* of winter rye, with *RMSE* decreasing by more than 4 days. The

improvement for *early dough ripening* of maize is close to one day, while the other performance gains are less than 0.5 days. In general, the effect of incorporation of soil moisture on the model accuracy is low, only for two phases it is close to one day or higher (see Table 4).

### 4.6. Correlation between date of sowing and phenological events

The date on which a plant was sown is an additional, anthropogenic influencing factor on plant phenological development and modeling. To quantify this influence, Pearson's correlation coefficients $r$ were calculated between the date of sowing and the $DOY_P$ variants, with and without consideration of soil moisture, as well as between the date of sowing and $DOY_{obs}$ (Table 5).

As expected, $DOY_{obs}$ and the date of sowing are almost linear correlated for *emergence* and show high significance. For *shooting*, $r$ is moderate with very high significance, while for all later phases $r$ is low. For *heading*, all $r$ values are close to zero and not significant, while for the two ripening phases they are slightly higher. In general, correlation between the raw observations is lower than the correlation of the modeled phases and the observed day of sowing. Incorporation of soil moisture slightly increases the correlation coefficient.

### 4.7. Error sources

Volunteered Geographic Information in general are of limited reliability (Flanagin and Metzger, 2008). Here, for the phenological observations data set, this concerns especially locational uncertainties, cultivar differences and the lack of information regarding fertilization and other management practices (see Sections 2.1 and 4.4) which is one source for model uncertainties.

Due to the spatial interpolation of the observations, uncertainties of the used temperatures are unavoidable. The average uncertainty and an approach to balance accuracy and processing time were discussed in Section 4.1. Other sources for mean temperatures are satellite data, which provide spatially explicit values. However, these values are only valid under cloud-free conditions.

The resolution of the underlying DEM suppresses small-scale variations in elevation and thus reduces interpolation accuracies. The usage of a higher resolution DEM from satellite data e.g. native the *SRTM DEM, WorldDEM* (90 m and 12 m resolution, respectively)

**Table 4**
Difference of *RMSE* between modeling results with and without consideration of soil moisture. Negative values indicate accuracy improvement when soil moisture is considered.

|  | W. wheat | W. rye | W. barley | Oil. rape | Oats | Maize | Potato | Sug. beet |
|---|---|---|---|---|---|---|---|---|
| Emergence | 0.55 | −4.38 | 0.49 | 0.22 | 0.02 | −0.01 | −0.18 | −0.12 |
| Shooting | −0.03 | −0.05 | 0.03 | 0.12 | 0.00 | −0.05 | – | – |
| Heading | 0.27 | 0.23 | 0.02 | – | −0.11 | – | – | – |
| Milk ripening | 0.09 | – | – | – | −0.01 | −0.05 | – | – |
| Yellow ripening | 0.07 | 0.1 | 0.27 | – | 0.08 | −0.35 | – | – |
| Beg. of flowering | – | 0.03 | – | 0.09 | – | −0.05 | – | – |
| Full flowering | – | −0.09 | – | – | – | – | – | – |
| 4th leaf unfolded | – | – | – | 0.88 | – | – | – | – |
| Bud formation | – | – | – | −0.04 | – | – | – | – |
| Full ripening | – | – | – | −0.08 | – | – | – | – |
| Early dough ripening | – | – | – | – | – | −0.94 | – | – |
| Tassel emergence | – | – | – | – | – | 0.04 | – | – |
| Closed stand | – | – | – | – | – | – | −0.45 | 0.14 |

**Table 5**
Correlation coefficients $r$ and $p$-values for $DOY_P^{\theta}, DOY_P$ and $DOY_{obs}$ against the observed date of sowing for winter wheat.

|  | Emergence | | Shooting | | Heading | | Milk ripening | | Yellow ripening | |
|---|---|---|---|---|---|---|---|---|---|---|
|  | r | p | r | p | r | p | r | p | r | p |
| $DOY_P$ | 0.98 | <2.2e−16 | 0.49 | <2.2e−16 | −0.04 | 0.38 | 0.28 | 7.1e−7 | 0.25 | 2.8e−6 |
| $DOY_P^{\theta}$ | 0.98 | <2.2e−16 | 0.50 | <2.2e−16 | −0.03 | 0.56 | 0.29 | 1.9e−7 | 0.26 | 1.4e−6 |
| $DOY_{obs}$ | 0.94 | <2.2e−16 | 0.31 | 8.6e−8 | −0.05 | 0.30 | 0.21 | 1.3e−4 | 0.15 | 4.1e−3 |

or DEM derived from airborne laser scanning with resolutions down to a few centimeters could increase interpolation accuracy, but due to computational limitations study site size and DEM resolution must be balanced.

Further model uncertainties are consequences of the model's intended simplicity. A more sophisticated model that includes parameters like other cardinal temperatures (e.g., upper thresholds, optimum temperatures), which can be freely obtained for many WMO-listed weather stations worldwide, could increase the model accuracy. However, the higher the input data requirements are, the smaller are the application scenarios for the model due to different data availability between regions, countries, phenological observations, etc.

*4.8. Comparison with other models*

The PHASE model is based on only free and operationally available Germany-wide information on temperature, phenological events and elevation in order to create on-demand maps of up-to-date or past phenological conditions at any location in Germany. On the one hand, the resulting phenological maps represent spatially explicit interpolation results of a specific and adaptable geometric resolution (here: $1 \times 1$ km$^2$). This is in contrast to comparable *GDD* approaches applied in Germany which aggregate phenological observation values to reference units like eco-regions (Siebert and Ewert, 2012) or large-scaled grid cells of $50 \times 50$ or $100 \times 100$ km$^2$ (van Bussel et al., 2011). On the other hand, the thematic depth of the PHASE interpolation results is superior and more crop-specific compared to spatially explicit and satellite-based parameters (e.g. Hird and McDermid, 2009).

Unlike other statistical phenological models, PHASE does not include features and capabilities which might lead to more accurate predictions. Some statistical phenological models include species-specific chilling requirements during the period of dormancy (Luedeling et al., 2009). Consideration of chilling requirements could further improve modeling accuracies for some species, since chilling effects are mostly effective for winter crops but less for summer cultivars.

Furthermore, other models allow the prediction of phenological responses to climate projections into the future (Schröder et al., 2014). This is a very useful capability for a large variety of climate-related sciences, but it is beyond the scope of the PHASE model. One example is the *Promotor Inhibitor Model* (PIM; Schaber and Badeck, 2003) which has recently been applied to investigate the impact of future climate(s) on tree phenology (Lange et al., 2016). PIM is a generalized physiology-based model and fitted on long-term phenological observations, e.g. DWD observational network. Consequently, PIM can assess and explain the importance of underlying drivers like temperature in the sense of chilling requirements or day length which may potentially vary between phenological phases and species. Once robustly fitted, it can also be used for comparable climates or species even when no phenological observations are available. If the underlying mechanism – as exhibited in the long-term observational set – changes, PIM's transferability can be substantially restricted. PHASE, in contrast, uses (geo-) statistics of a single phenological cycle. This ensures for a single year and phase a better predictive performance (2–3 days less MAE compared to PIM).

## 5. Conclusions and outlook

In this article, we presented the geo-statistical model PHASE which makes the automatic, Germany-wide and spatially explicit prediction of phenological phases possible. The model uses publicly available input data and is characterized by the following features:

- PHASE enables an automatic optimization of all critical parameters which makes the model independent of user-specific parametrization. Thus, the model allows the spatial interpolation of any phase for which phenological observations exist. Each modeling result is characterized by accuracy metrics.
- Only three types of input data are needed including daily provided point data on temperature and phenological events and digital elevation model.
- As demonstrated for the example of soil moisture, additional explaining variables can be integrated in the model to possibly increase the results' accuracy.
- The predictions' geometric resolution is free adaptable. In the presented model setting, the spatial resolution of the resulting phenological raster data set is $1 \times 1$ km. Currently, no publicly available phenological model of equivalent performance regarding spatial and thematic resolution as well as processing time is known.

Although the introduced workflow refers to German data situation, the approach is, however, transferable to other temperate regions where a sufficient data base of phenological observations and daily temperature measurements exist or are continuously reported. These requirements are fulfilled in some other European countries, e.g. the Netherlands, Belgium and Great Britain (Rodriguez-Galiano et al., 2015) and the US (Rosemartin et al., 2014). Apart from networks driven by state authorities, the amount of volunteered geographic information including phenological observations has increased in recent years due to the improvement in online communication and positional tracking techniques (Mehdipoor et al., 2015). In addition to human-based observations, stationary near-surface cameras (phenocams) will increasingly provide a permanent visual record of phenological developments (Richardson et al., 2013).

A model transferability is also to be expected to other crop types like soybeans whose growth and development is influenced by heat units and by photoperiod (Setiyono et al., 2007), but which are rarely or almost not cultivated in Central Europe. Depending on the region-specific availability of relevant open-source input data sets, the model can be extended or modified accordingly.

The modeling results are of interest for remote sensing applications, especially for crop classification which has been one of the key applications for remote sensing data over decades. However, the number of distinct cultivated crops that can be discriminated accurately is mostly limited on a few and partly aggregated classes. Class separation on species level is often hampered by high spectral similarity of closely related species e.g., cereal species like wheat, rye and barley. Phenological information can work as indicators for the selection of optimal satellite images to achieve more accurate classification results on species level (Gerstmann et al., 2016).

A further information gain by the usage of PHASE outputs could be achieved in mapping of invasive species. Within the study by Bradley and Mustard (2006), invasive cheatgrass was mapped in northern Nevada, USA, using phenology as indicator for classification and for derivation of invasion risk. For this purpose, Landsat images were selected arbitrary under the knowledge, that cheatgrass has its highest greenness in mid-May, slightly earlier than the grass species which are native in northern Nevada steppes. This arbitrary selection can be assisted and potentially improved by modeled phenology as provided by PHASE.

Turner et al. (2003) stated, that remote sensing based information on plant phenological development are important variables for mapping of biodiversity and determination of vegetation patterns down to species levels. However, remote sensing methods provide a less detailed thematic resolution than phenological information modeled using ground observations. Adapted to the

specific application, PHASE thus has the potential to improve results of biodiversity studies using remote sensing data.

In addition, the coupling of up-to-date phenological predictions with corresponding remotely sensed imagery or simulated times series opens up opportunities for the monitoring of ecosystem services like soil erosion control (Guerra et al., 2014). In doing so, temporal windows and specific reference units like parcels can be detected where soils are potentially covered by sparse or dense vegetation, crop residues or are free of coverage, which enables dynamic soil erosion modeling (Möller et al., 2015).

It is planned to enable public access of operationally produced modeling results within a WebGIS environment. Therefore, for time-consuming parts of the PHASE model, like the prediction of daily temperatures or the optimization of base temperatures (Section 3.2.1) and accumulated effective temperatures (Section 3.2.3), appropriate parallelization techniques (Schiele et al., 2012) are currently tested to enable computational efficient processing.

### Acknowledgement

### References

Barrett, E.C., 2013. Introduction to Environmental Remote Sensing, fourth ed. Routledge.

Berger, M., Moreno, J., Johannessen, J.A., Levelt, P.F., Hanssen, R.F., 2012. ESA's Sentinel missions in support of Earth system science. Rem. Sens. Environ. 120, 84–90.

Bivand, R., Keitt, T., Rowlingson, B., 2015. rgdal: Bindings for the Geospatial Data Abstraction Library. R Package Version 1.0-7.

Bradley, B.A., Mustard, J.F., 2006. Characterizing the landscape dynamics of an invasive plant and risk of invasion using remote sensing. Ecol. Appl. 16 (3), 1132–1147.

Chmielewski, F.-M., Müller, A., Bruns, E., 2004. Climate changes and trends in phenology of fruit trees and field crops in Germany, 1961–2000. Agric. For. Meteorol. 121 (1-2), 69–78.

Chuine, I., Kramer, K., Hänninen, H., 2003. Plant development models. In: Schwartz, M.D. (Ed.), Phenology: An Integrative Environmental Science, Tasks for Vegetation Science, vol. 39. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 217–235.

Deutscher Wetterdienst, 2015b. Vorschriften und Betriebsunterlagen für die phänologischen Beobachter des Deutschen Wetterdienstes. Offenbach, Germany.

Drusch, M., Bello, U.D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Rem. Sens. Environ. 120, 25–36.

Deutscher Wetterdienst, 2015a. Climate Data Center. URL <ftp://ftp-cdc.dwd.de/pub/CDC/> (last access: 15-03-2015).

DWD Climate Data Center (CDC), 2016. Daily Grids of Soil Moisture under Grass and Sandy Loam, Version 0.x. Tech. Rep.

Estrella, N., Sparks, T.H., Menzel, A., 2007. Trends and temperature response in the phenology of crops in Germany. Glob. Change Biol. 13 (8), 1737–1747.

Ewert, F., Rodriguez, D., Jamieson, P., Semenov, M., Mitchell, R., Goudriaan, J., Porter, J., Kimball, B., Pinter Jr., P., Manderscheid, R., et al., 2002. Effects of elevated $CO_2$ and drought on wheat: testing crop simulation models for different experimental and climatic conditions. Agric. Ecosyst. Environ. 93 (1), 249–266.

Flanagin, A.J., Metzger, M.J., 2008. The credibility of volunteered geographic information. GeoJournal 72 (3–4), 137–148.

Foerster, S., Kaden, K., Foerster, M., Itzerott, S., 2012. Crop type mapping using spectral-temporal profiles and phenological information. Comput. Electron. Agric. 89 (0), 30–40.

Fu, Y., Zhang, H., Dong, W., Yuan, W., 2014. Comparison of phenology models for predicting the onset of growing season over the northern hemisphere. PLoS ONE 9 (10), e109544.

Gerstmann, H., Möller, M., Gläßer, C., 2016. Optimization of spectral indices and long-term separability analysis for classification of cereal crops using multi-spectral RapidEye imagery. Int. J. Appl. Earth Obs. Geoinf. 52, 115–125.

Grant, C.A., Bailey, L.D., 1993. Fertility management in canola production. Can. J. Plant Sci. 73 (3), 651–670.

Guerra, C.A., Pinto-Correia, T., Metzger, M.J., 2014. Mapping soil erosion prevention using an ecosystem service modeling framework for integrated land management and policy. Ecosystems 17 (5), 878–889.

Hänninen, H., Kramer, K., 2007. A framework for modelling the annual cycle of trees in boreal and temperate regions. Silva Fenn. 41 (1), 167–205.

Hengl, T. 2009. A Practical Guide to Geostatistical Mapping of Environmental Variables. Tech. Rep. European Commission, Joint Research Centre, Institute for Environment and Sustainability.

Hengl, T., Heuvelink, G.B., Rossiter, D.G., 2007. About regression-kriging: from equations to case studies. Comput. Geosci. 33 (10), 1301–1315.

Hense, A., Müller, M., 2007. Geostatistische Modellierung und Qualitätskontrolle von phänologischen Beobachtungen. Promet 33 (1/2), 7–13.

Hiemstra, P., Pebesma, E., Twenhöfel, C., Heuvelink, G., 2009. Real-time automatic interpolation of ambient gamma dose rates from the dutch radioactivity monitoring network. Comput. Geosci. 35 (8), 1711–1721.

Hijmans, R.J., 2015. raster: Geographic Data Analysis and Modeling. R Package Version 2.4-20. URL <http://CRAN.R-project.org/package=raster>.

Hijmans, R.J., Williams, E., Vennes, C., 2014. Geosphere: Spherical Trigonometry for Geographic Applications.

Hird, J., McDermid, G., 2009. Noise reduction of NDVI time series: an empirical comparison of selected techniques. Rem. Sens. Environ. 113 (1), 248–258.

Holen, C., Dexter, A., 1996. A growing degree day equation for early sugarbeet leaf stages. Sugarbeet Res. Ext. Rep. 27, 152–157.

Huang, C.-y., Asner, G.P., 2009. Applications of remote sensing to alien invasive plant studies. Sensors 9 (6), 4869–4889.

Idso, S.B., Jackson, R.D., Reginato, R.J., 1978. Extending the degree day concept of plant phenological development to include water stress effects. Ecology 59 (3), 431–433.

Jamieson, P., Porter, J., Goudriaan, J., Ritchie, J., Van Keulen, H., Stol, W., 1998. A comparison of the models AFRCWHEAT2, CERES-wheat, sirius, SUCROS2 and SWHEAT with measurements from wheat grown under drought. Field Crops Res. 55 (1), 23–44.

Jönsson, A., Eklundh, L., Hellström, M., Bärring, L., Jönsson, P., 2010. Annual changes in MODIS vegetation indices of swedish coniferous forests in relation to snow dynamics and tree phenology. Rem. Sens. Environ. 114 (11), 2719–2730.

Kaduk, J., Heimann, M., 1996. Assessing the climate sensitivity of the global terrestrial carbon cycle model SILVAN. Phys. Chem. Earth 21 (5–6), 529–535.

Kirby, E., Porter, J., Day, W., Adam, J.S., Appleyard, M., Ayling, S., Baker, C., Belford, R., Biscoe, P., Chapman, A., et al., 1987. An analysis of primordium initiation in avalon winter wheat crops with different sowing dates and at nine sites in England and Scotland. J. Agric. Sci. 109 (01), 123–134.

Kramer, K., Leinonen, I., Loustau, D., 2000. The importance of phenology for the evaluation of impact of climate change on growth of boreal, temperate and mediterranean forests ecosystems: an overview. Int. J. Biometeorol. 44 (2), 67–75.

Krige, D.G., 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. J. Chem. Metall. Min. Soc. S. Afr. 52 (6), 119–139.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer.

Lange, M., Schaber, J., Marx, A., Jäckel, G., Badeck, F.-W., Seppelt, R., Doktor, D., 2016. Simulation of forest tree species bud burst dates for different climate scenarios: chilling requirements and photo-period may limit bud burst advancement. Int. J. Biometeorol.

Lee, J.S., 1980. Digital image enhancement and noise filtering by use of local statistics. IEEE Trans. Pattern Anal. Mach. Intell. 2, 165–168.

Löpmeier, F.-J.Deutscher Wetterdienst, 1983. Agrarmeteorologisches Modell zur Berechnung der aktuellen Verdunstung (AMBAV). Dt. Wetterdienst (DWD), Zentrale Agrarmeteorologische Forschungsstelle Braunschweig.

Luedeling, E., Zhang, M., McGranahan, G., Leslie, C., 2009. Validation of winter chill models using historic records of walnut phenology. Agric. For. Meteorol. 149 (11), 1854–1864.

Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G., Worthy, L.D., 2006. Land-cover change detection using multi-temporal MODIS NDVI data. Rem. Sens. Environ. 105 (2), 142–154.

Masle, J., Doussinault, G., Farquhar, G.D., Sun, B., 1989. Foliar stage in wheat correlates better to photothermal time than to thermal time. Plant Cell Environ. 12 (3), 235–247.

Matérn, B., 1960. Spatial variation, . Lecture Notes in Statistics, second ed., vol. 36. Springer, Berlin.

McMaster, G.S., White, J.W., Weiss, A., Baenziger, P.S., Wilhelm, W.W., Porter, J.R., Jamieson, P.D., 2009. Simulating crop phenological responses to water deficits. In: Ahuja, L.R., Reddy, V.R., Anapalli, S.A., Yu, Q. (Eds.), Modeling the Response of Crops to Limited Water: Recent Advances in Understanding and Modeling Water Stress Effects on Plant Growth Processes. ASA-SSACSSA, pp. 277–300.

McMaster, G.S., Wilhelm, W.W., 1997. Growing degree-days: one equation, two interpretations. Agric. For. Meteorol. 87 (4), 291–300.

McMaster, G.S., Wilhelm, W.W., 2003. Phenological responses of wheat and barley to water and temperature: improving simulation models. J. Agric. Sci. 141 (2), 129–147.

Mehdipoor, H., Zurita-Milla, R., Rosemartin, A., Gerst, K., Weltzin, J., 2015. Developing a workflow to identify inconsistencies in volunteered geographic information: a phenological case study. PLoS ONE 10 (10), 1–14.

Menzel, A., 2007. Phänologische Modelle – phenological models. Promet 33 (1/2), 20–27.

Miller, P., Lanier, W., Brandt, S., 2001. Using Growing Degree Days to Predict Plant Stages. Montana State University, USA (Extension Service).

Möller, M., Gerstmann, H., Thürkow, D., Gao, F., Förster, M., 2015. Coupling of phenological information and synthetically generated time-series for crop types as indicator for vegetation coverage information. In: 8th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp).

Möller, M., Müller, S., Doktor, D., Gläßer, C., 2012. Phenological structuring of multi-temporal RapidEye imagery. In: Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, pp. 4934–4937.

Nellis, M., Price, K.P., Rundquist, D., 2009. Remote sensing of cropland agriculture. In: Warner, T.A., Nellis, M.D., Foody, G.M. (Eds.), The SAGE Handbook of Remote Sensing, vol. 1. SAGE Publications, London, UK, pp. 368–380.

Pebesma, E.J., 2004. Multivariable geostatistics in s: the gstat package. Comput. Geosci. 30, 683–691.

Peng, C., Guiot, J., Wu, H., Jiang, H., Luo, Y., 2011. Integrating models with data in ecology and palaeoecology: advances towards a model–data fusion approach. Ecol. Lett. 14 (5), 522–536.

Picard, G., Quegan, S., Delbart, N., Lomas, M.R., Toan, T., Woodward, F., 2005. Bud-burst modelling in Siberia and its impact on quantifying the carbon budget. Glob. Change Biol. 11 (12), 2164–2176.

Prishchepov, A.V., Radeloff, V.C., Dubinin, M., Alcantara, C., 2012. The effect of Landsat ETM/ETM+ image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. Rem. Sens. Environ. 126, 195–209.

R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org/>.

Rabus, B., Eineder, M., Roth, A., Bamler, R., 2003. The shuttle radar topography mission–a new class of digital elevation models acquired by spaceborne radar. ISPRS J. Photogramm. Remote Sens. 57 (4), 241–262.

Réaumur, M., 1735. Observations du thermométre faites á Paris pendant l'année 1735 comparées avec celles qui ont été faites sous la ligne á l'Ile de France, á Alger et en quelques-unes de nos îles de l'Amérique. Académie Royale Sci., 545–576

Richardson, A., Keenan, T., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agric. For. Meteorol. 169, 156–173.

Rodriguez-Galiano, V., Dash, J., Atkinson, P., 2015. Characterising the land surface phenology of Europe using decadal MERIS data. Remote Sens. 7 (7), 9390–9409.

Rosemartin, A., Crimmins, T., Enquist, C., Gerst, K., Kellermann, J., Posthumus, E., Denny, E., Guertin, P., Marsh, L., Weltzin, J., 2014. Organizing phenological data resources to inform natural resource conservation. Biol. Conserv. 173, 90–97.

Russelle, M., Wilhelm, W., Olson, R., Power, J.F., 1984. Growth analysis based on degree days. Crop Sci. 24 (1), 28–32.

Salazar-Gutierrez, M., Johnson, J., Chaves-Cordoba, B., Hoogenboom, G., 2013. Relationship of base temperature to development of winter wheat. Int. J. Plant Prod. 7 (4).

Schaber, J., Badeck, F.-W., 2003. Physiology-based phenology models for forest tree species in Germany. Int. J. Biometeorol. 47 (4), 193–201.

Schiele, S., Möller, M., Blaar, H., Thürkow, D., Müller-Hannemann, M., 2012. Parallelization strategies to deal with non-localities in the calculation of regional land-surface parameters. Comput. Geosci. 44, 1–9.

Schröder, W., Schmidt, G., Schönrock, S., 2014. Modelling and mapping of plant phenological stages as bio-meteorological indicators for climate change. Environ. Sci. Eur. 26 (5), 2–13.

Schwartz, M.D. (Ed.), 2006. Phenology: an integrative environmental science. Tasks for Vegetation Science, vol. 39. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Setiyono, T., Weiss, A., Specht, J., Bastidas, A., Cassman, K., Dobermann, A., 2007. Understanding and modeling the effect of temperature and daylength on soybean phenology under high-yield conditions. Field Crops Res. 100, 257–271.

Siebert, S., Ewert, F., 2012. Spatio-temporal patterns of phenological development in Germany in relation to temperature and day length. Agric. For. Meteorol. 152, 44–57.

Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J.O., Levis, S., Lucht, W., Sykes, M.T., et al., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Glob. Change Biol. 9 (2), 161–185.

Slafer, G., Rawson, H., 1995. Base and optimum temperatures vary with genotype and stage of development in wheat. Plant Cell Environ. 18 (6), 671–679.

Slafer, G., Savin, R., 1991. Developmental base temperature in different phenological phases of wheat (Triticum aestivum). J. Exp. Bot. 42 (8), 1077–1082.

Sobrino, J.A., Julien, Y., Soria, G., 2013. Phenology estimation from Meteosat second generation data. IEEE J. Sel. Topics Appl. Earth Observations Remote Sens. 6 (3), 1653–1659.

Solantie, R., 2004. Daytime temperature sum – a new thermal variable describing growing season characteristics and explaining evapotranspiration. Boreal Environ. Res. 9, 319–333.

Spatialreference, 2015. Catalogs of Spatial Reference Systems. URL <http://spatialreference.org> (accessed on 22nd February 2015)

Statistisches Bundesamt, 2011. Land- und Forstwirtschaft, Fischerei. Bodennutzung der Betriebe (landwirtschaftlich genutzte Flächen). Agrarstrukturerhebung. online. URL <https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00004301/2030312117004.pdf> (2016-04-22).

Stein, M., 1999. Interpolation of Spatial Data: Some Theory for Kriging. Springer Series in Statistics. Springer, New York.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., Steininger, M., 2003. Remote sensing for biodiversity science and conservation. Trends Eecol. Evol. 18 (6), 306–314.

USGS, 2004. Shuttle Radar Topography Mission.

van Bussel, L.G.J., Ewert, F., Leffelaar, P.A., 2011. Effects of data aggregation on simulations of crop phenology. Agric. Ecosyst. Environ. 142 (1-2), 75–84.

Van Niel, T.G., McVicar, T.R., 2004. Determining temporal windows for crop discrimination with remote sensing: a case study in south-eastern australia. Comput. Electr. Agric. 45 (1-3), 91–108.

Vrieling, A., de Beurs, K.M., Brown, M.E., 2011. Variability of African farming systems from phenological analysis of NDVI time series. Climatic Change 109 (3–4), 455–477.

Wang, E., Engel, T., 1998. Simulation of phenological development of wheat crops. Agric. Syst. 58 (1), 1–24.

Webster, R., Oliver, M., 2007. Geostatistics for Environmental Scientists. Statistics in Practice. Wiley.

Wielgolaski, F.-E., 1999. Starting dates and basic temperatures in phenological observations of plants. Int. J. Biometeorol. 42 (3), 158–168.

Xiao, W., Sun, Z., Wang, Q., Yang, Y., 2013. Evaluating MODIS phenology product for rotating croplands through ground observations. J. Appl. Remote Sens 7 (1), 073562.

Yang, S., Logan, J., Coffey, D.L., 1995. Mathematical formulae for calculating the base temperature for growing degree days. Agric. For. Meteorol. 74 (1-2), 61–74.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens. Environ. 84 (3), 471–475.

Zhao, M., Peng, C., Xiang, W., Deng, X., Tian, D., Zhou, X., Yu, G., He, H., Zhao, Z., 2013. Plant phenological modeling and its application in global climate change research: overview and future challenges. Environ. Rev. 21 (1), 1–14.

## 5.2 CONCLUSIONS FROM PAPER 2

Paper 2 demonstrated, that a concise phenological model based on only three types of input data provides more accurate phenological information for regions without observation data than simple geostatistical interpolation methods.

Another finding is, that volunteered geographic information is of enough quality to achieve results of sufficient accuracy, but also is a source of uncertainty that cannot be definitely quantified.

It could also be proven that the flexibility of the model design allows the consideration of additional auxiliary input data sets but demonstrates the strong dependence of plant phenology on temperature as main driver in temperate climates. The developed phenological model addresses exactly the conclusions of Paper 1, enabling the combination of both methods to relate separability patterns to accurate, spatially explicit phenological data.

# 6

## DETECTION OF PHENOLOGY-DEFINED DATA ACQUISITION TIME FRAMES FOR CROP TYPE MAPPING

In Paper 3, the conclusions of the findings of Chapter 4 are addressed by the developed phenological model PHASE that was described in Chapter 5. The shortcomings of raw observational phenological data to explain spatial and temporal differences in spectral separability of cereal crops are approached by utilisation of the operationally applied PHASE model.

Furthermore, it could be again demonstrated that a single vegetation index cannot be used alone for spectral separation, since during different growth stages the different plant physiological processes result in varying spectral response. The approach was applied on established variants of NDVI and single-band spectral reflectances. For demonstrative purpose, the separability analyses were performed for the three dominating crop types winter wheat, winter barley and winter rapeseed.

The resulting separability profiles are linked to phenological phases modelled using PHASE. The phenological stage that was most frequently present during the separability peak is understood as indicator phase that can be used for the selection of the best-suited data sets of a satellite image time series. The indicator phases were defined for the Harz study sites (see Section 3.1.1) and applied for transferability assessment to the DEMMIN site (see Section 3.1.2). The phenological modelling procedure is automatised to be updated daily. An infrastructure is presented to distribute the modelled phases to the public for free.

DGPF · CrossMark

ORIGINAL ARTICLE

# Detection of Phenology-Defined Data Acquisition Time Frames For Crop Type Mapping

Henning Gerstmann[1] · Cornelia Gläßer[1] · Detlef Thürkow[1] · Markus Möller[2]

**Abstract** Agricultural monitoring and assessment based on satellite data increasingly gains importance due to the growing number of available satellite sensors with high geometric and temporal resolution. Such tasks often require multiple images acquired on specific dates that among others account for inter-annual phenological variations to provide accurate results. This contribution presents an approach that links peaks of spectral separability profiles to crop phenological phases. The phases are spatially interpolated using a phenological model and ground observations. The profiles show the respective temporal development of the $F$-measure which is used as indicator for class-wise separability. It originates from binary classifications of vegetation indices computed for each set of a satellite data archive covering multiple years. Acquisition dates, which repeatedly show a separability maximum define phenological indicator phases. Potential alternative phases can be also defined. Experiments based on multi-temporal RapidEye satellite imagery were performed for three crops at two German test sites under different environmental conditions. The results showed that the phases yellow ripeness, heading and flowering can function as indicator phases for high spectral separability of winter barley, winter wheat and winter rapeseed. We could identify at least two identical, stable indicator phases per crop type for both test sites, which suggests the transferability and robustness of the presented approach.

**Zusammenfassung** *Detektion von phänologisch definierten Datenaufnahmezeiträumen für die Klassifikation von Feldfrüchten*
Auf Satellitendaten basierendes landwirtschaftliches Monitoring gewinnt durch die wachsende Anzahl verfügbarer Sensoren mit hoher zeitlicher und geometrischer Auflösung zunehmend an Bedeutung. Für solche Anwendungen werden oftmals Satellitendaten von verschiedenen Aufnahmetagen benötigt, deren Auswahl inter-annuelle phänologische Variationen berücksichtigen muss, um exakte Ergebnisse zu liefern. Dieser Beitrag präsentiert einen Ansatz, um Maxima von spektralen Trennbarkeitsprofilen mit phänologischen Phasen von Feldfrüchten zu verbinden. Diese Phasen werden unter Nutzung eines phänologischen Modelles und Beobachtungsdaten räumlich interpoliert. Die Trennbarkeitsprofile zeigen den zeitlichen Verlauf des $F$-Maß, das als Indikator für klassenspezifische Trennbarkeit genutzt wird. Dieses stammt von binär klassifizierten Vegetationsindizes, die für jeden Datensatz einer mehrjährigen, multi-temporalen Zeitserie von Satellitenbilddatensätzen berechnet wurden. Zeitpunkte, während denen wiederholt das Trennbarkeitsmaximum beobachtet werden konnte, weisen die Indikatorphasen aus. Potentielle Alternativphasen können ebenso bestimmt werden. Die Untersuchungen wurden für drei Fruchtarten in zwei Untersuchungsgebieten in Deutschland unter verschiedenen Umweltbedingungen auf Basis von RapidEye-Satellitendaten durchgeführt. Die Ergebnisse zeigen, dass die Phasen Gelbreife, Ährenschieben und Blüte als Indikatoren für hohe spektrale Trennbarkeit von Wintergerste, Winterweizen und Winterraps dienen können. Für jede untersuchte Fruchtart konnten wenigstens zwei, für beide Untersuchungsgebiete übereinstimmende, stabile Indi-

Henning Gerstmann
henning.gerstmann@geo.uni-halle.de

[1]   Institute for Geosciences and Geography, Martin Luther University Halle-Wittenberg, Von-Seckendorff-Platz 4, 06120 Halle (Saale), Germany

[2]   Institute for Strategies and Technology Assessment, Julius Kühn-Institute, Federal Research Centre for Cultivated Plants, Stahnsdorfer Damm 81, 14532 Kleinmachnow, Germany

🙂 Springer

katorphasen ausgewiesen werden, was die Übertragbarkeit und Robustheit des gezeigten Verfahrens belegt.

**Schlagwörter** Spektrale Trennbarkeit · Phänologische Phasen · Multitemporal · RapidEye · Deutschland

## 1 Introduction

The accuracy and efficiency of remote sensing applications in agriculture based on optical satellite imagery is strongly affected by the variability of crop type phenology. The phenology of crops, describing the timing and the sequence of developmental stages, varies significantly between crop types as well as under different agronomic practices and climates. This has to be considered when remote sensing methods are applied for crop and land use mapping (e.g., Schmidt et al. 2014), in-season yield estimations for food security assessment (e.g., Meroni et al. 2014) or efficient image compositing (Frantz et al. 2017). All these applications require an exactly timed data set selection to provide the most accurate results.

Crop type mapping is one key application in agricultural remote sensing. To derive accurate land use maps that are essential for many purposes, crop types can be efficiently distinguished during specific phenological phases. These phenological windows require dynamic adjustments for annual variability in crop development, which has been continuously studied during the last few decades (Murakami et al. 2001; van Niel and McVicar 2004; Förster et al. 2012; Conrad et al. 2014; Schmidt et al. 2014; Azar et al. 2016; Möller et al. 2017).

The phenology-aided selection of the time steps based on expert knowledge is widely studied (Guerschman et al. 2003; Peña-Barragàn et al. 2011; Conrad et al. 2014; Schmidt et al. 2014). However, these studies are often based on static crop calendars that do not reflect weather-induced annual shifts in phenology (Meroni et al. 2014). Crop-specific responses to such variations account for inter-annual differences between the times of highest separability among crop types (Nitze et al. 2015), which points at the necessity for a dynamic derivation of spatial phenological information. For this purpose, the analysis of temporal profiles that consist of satellite-based vegetation indices (e.g., Frantz et al. 2017) proved to be a successful strategy. For instance, phenological metrics like maximum and minimum vitality derived from such temporal profiles enabled the direct extraction of a small number of key growth stages such as green up, heading and senescence (Xu et al. 2017). Phenological in situ observations, which usually provide a much higher number of individual stages, can be alternatively used to interpret variations of separability profiles of crops (Förster et al. 2012). Furthermore, they can be used to validate phenological stages

derived from vegetation index profiles (Xu et al. 2017) which are often highly affected by noise. The spatial modelling of point observations of the crops' growth status (Gerstmann et al. 2016a) combines the advantages of spatially explicit and in situ phenological data sets and can thus improve the knowledge about critical days for crop type classification.

Currently, the majority of modelling and classification approaches aggregates crop types of similar spectral and phenological behaviour to wider classes such as winter cereals or root crops. While these aggregations are sufficient for most applications, information on crop species level is essential for detailed yield predictions (Nitze et al. 2015), water management (Conrad et al. 2013), or subsidy control. However, this class aggregation is often unavoidable, since data gaps caused by cloud coverage or sensor-specific characteristics limit the capability to separate spectrally similar species.

These data gaps are globally an issue for almost all applications based on optical remote sensing imagery. Hence, the consideration of broader time frames increases the chance of having usable data available during high separability periods. Narrow alternative time frames can also be found by inter-annual analyses of species-specific phenological behaviour, but these time frames have only limited relevance due to the mentioned high probability of data gaps.

The main objective of this study is to identify phenological phases that suit for optimal class separation in crop mapping. Therefore, a framework was developed that systematically combines interpolated phenological ground observations with satellite image acquisition dates and spectral separability patterns. In doing so, phenological phases are analysed in terms of their suitability as indicators for optimal acquisition time frames.

The approach is applied to multi-temporal RapidEye data acquired for an agriculturally used test site in Central Germany. At this site, indicator phases are defined based on separability profile and phenological phases. The transferability of these indicator phases is evaluated on a validation site characterised by different growing conditions.

Finally, a web-based tool is presented that provides almost real-time phenological raster data covering Germany to apply the framework in practise.

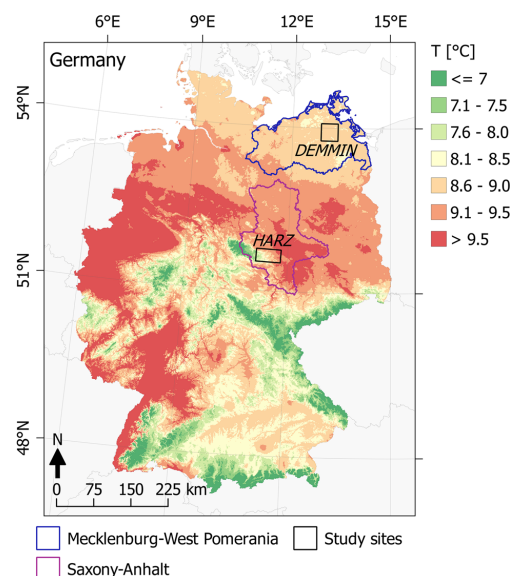## 2 Study Sites and Data

### 2.1 Site Descriptions

The two study sites have been intensively investigated by the Terrestrial Environmental Observatories (TERENO) research network that focuses on observations of long-term climate change impacts on regional scales (Bogena 2016). They were used for method development and testing, respectively.

17

The training site is located in Central Germany (Fig. 1), approximately 30 km north of the city of Halle (Saale), with the town Hettstedt being its centre (Fig. 2, left). The site consists of two sub-sites: a western sub-site covering the eastern parts of the Harz mountain range (HM) and an eastern sub-site representing the adjacent lowlands (HL). Both sub-sites



**Fig. 1** Locations of the study sites in Germany and annual mean temperature (DWD Climate Data Center 2017b). Projection: WGS84/UTM Zone 33N

cover an area of 25 km × 25 km each. Annual mean temperatures and precipitation are 8.5 °C and 630 mm for HM and 9.2 °C and 540 mm for HL, respectively. The area is dominated by agricultural land use on fertile chernozemic soils that allow the cultivation of demanding species. The dominating crop types are winter wheat, cropped on ≈ 45% of the agricultural area, winter rapeseed (≈ 20%), perennial grassland (≈ 10%) and winter barley (≈ 10%).

The second site used for the transferability study is located in the federal state of Mecklenburg-West Pomerania, in the surrounding of the city Demmin (DM, Fig. 2, right). The site covers an area of 35 km × 35 km. Annual mean temperature and precipitation in DM are 8.7 °C and 590 mm, respectively. Here, sandy soils of reduced fertility require more frequent cultivation of less demanding crops, which is represented by the lower percentage of winter wheat (≈ 30%) and higher percentages of the area covered by perennial grassland (≈ 15%) and by less-demanding winter rye (≈ 5%). Gradients between comparatively warm temperatures in the HL and relatively cool temperatures in HM are apparent in Fig. 1. Contrary to this strong gradient, the DM site is characterized by a homogeneous annual mean temperature which ranges between the two extremes of the Harz sub-sites.

Two aspects determined the selection of the study sites. First, natural conditions, e.g., soil fertility, should differ among the regions to assess the robustness of the presented framework. For instance, precipitation sums also varied between the sites, especially in the strong easterly gradient of decreasing precipitation between the Harz sub-sites (> 700 mm in the west and < 500 mm in the east). Second, crop types of the training and validation sites should show
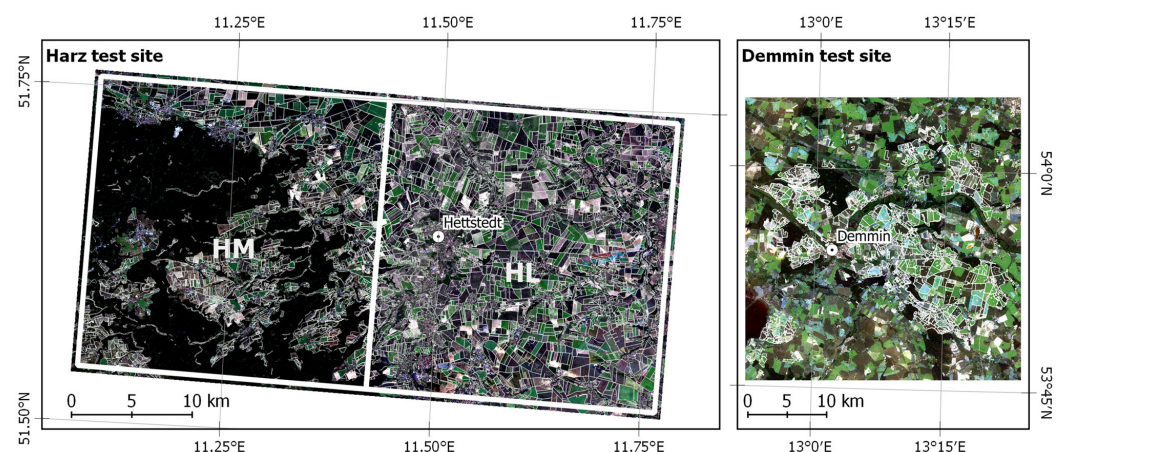


**Fig. 2** RapidEye images of the study sites Harz (left) and Demmin (right) acquired on June 6, 2011 (HL), June 2, 2011 (HM) and July 3, 2010 (DM). Band combination RGB 3-2-1 (true colour). Fields with crop cultivation data are accentuated. Projection: WGS84/UTM Zone 33N

⚊ Springer

some similarities to sustain the comparability of the results. So, the field sizes should ensure that a sufficient number of pure (not mixed) pixels is available for the calculation of the reflectance means per field. Both study sites fulfilled these requirements.

### 2.2 Satellite Data

Multi-temporal RapidEye data sets covering the HM and HL sites between 2010 and 2015 were made available through the RapidEye Science Archive (RESA; Borg et al. 2013, grant no: 653). The RapidEye satellite constellation collects imagery of the Earth's surface in five spectral bands (blue: 440–510 nm, green: 520–590 nm, red: 530–685 nm, red edge: 690–730 nm, near infra-red/NIR: 760–850 nm, Tyc et al. 2005). The data were obtained at processing level 3A with a spatial resolution of 5 m × 5 m. Atmospheric correction and cloud masking were performed using the software ATCOR 2© (Richter and Schläpfer 2015). For the DM site, RapidEye data sets for the years 2010–2012 were available at processing level 1B with 6.5 m × 6.5 m resolution. Preprocessing of these data sets was performed using the software CATENA© (Krauß et al. 2013) that also includes ATCOR 2©. The effects of differing sensor viewing angles (Nagol et al. 2015) on vegetation reflectance are eliminated by application of the ATCOR processing chain. The temporal distribution of the data sets is displayed in Fig. 3 and highlights inter-annual differences in temporal coverage and density, which are typical for optical imagery of high temporal and geometric resolution (Whitcraft et al. 2015).

### 2.3 Auxiliary Data

Parcel-based cropping information for 2010–2015 for the Harz sites were provided by the Ministry of Agriculture and Environment of Saxony-Anhalt and for DM via the TERENO long-term research programme. As part of the European

**Table 1** Observed phenological phases for winter wheat (WW), winter barley (WB) and winter rapeseed (WR) and corresponding phase IDs

| Name | ID | Crop |
| --- | --- | --- |
| Sowing | 10 | WW, WB, WR |
| Emergence | 12 | WW, WB, WR |
| Fourth leaf unfolded | 14 | WR |
| Shooting | 15 | WW, WB |
| Heading | 18 | WW, WB |
| Stem elongation | 67 | WR |
| Bud formation | 17 | WR |
| Beginning of flowering | 5 | WR |
| End of flowering | 7 | WR |
| Milk ripening | 19 | WW, WB |
| Yellow ripening | 21 | WR |
| Full ripening | 22 | WR |
| Harvest | 24 | WW, WB, WR |

Commission's Land Parcel Identification System (LPIS), the parcels are determined by ownership structure and homogeneity regarding the cultivated crop type (Inan et al. 2010). Germany-wide volunteer-based phenological point observations and meteorological measurements were obtained from the German Weather Service (DWD Climate Data Center 2017a, c) via FTP.[1] The names of the available phenological phases for winter wheat, winter rapeseed and winter barley are listed in Table 1 along with their numeric code according to the DWD observation programme (Kaspar et al. 2014).

Furthermore, a digital elevation model of 1 km × 1 km resolution was generated from the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM, Rabus et al. 2003).

## 3 Methods

The approach is based on crop-specific spectral separability profiles for six subsequent years. These profiles were computed from the RapidEye data sets by applying separability analysis on spectral features (Gerstmann et al. 2016b) calculated from parcel averages of reflectances. The profiles were, analogous to Möller et al. (2017), coupled with modelled phenological phases. The separability maxima over time exhibit the indicator phases, i.e., phenological phases that optimally suit for crop separation. However, before finally assigning the indicator phases, a reliability check was included, which among others targets at the question, if the maxima of separability occur at the same position over several years (here:
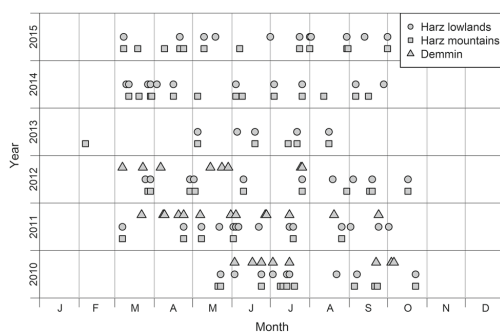


**Fig. 3** Temporal distribution of the RapidEye data sets

---

[1] ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany.
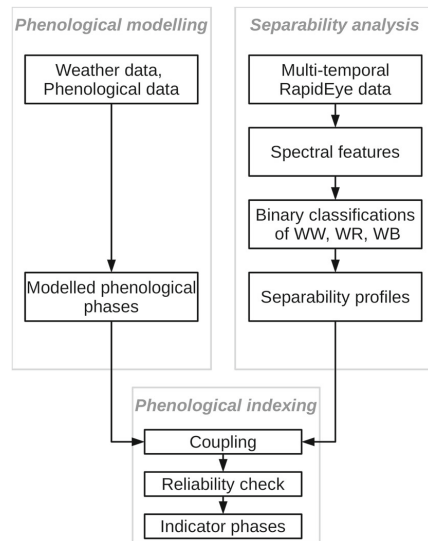
**Fig. 4** Workflow for the detection of indicator phases for the optimal separability of the crop types winter wheat (WW), winter barley (WB) and winter rapeseed (WR)

2010–2015). The workflow is summarized in Fig. 4. The components of the framework are described in detail in the following sections.

### 3.1 Modelling of Phenological Phases

Phenological modelling was performed using the PHASE model according to Gerstmann et al. (2016a). PHASE is a statistical model based on the growing-degree-days approach that relates phenological events to accumulated temperatures from a defined starting point. The model sums up accumulated daily mean temperatures between the day of sowing and the observation of the phenological phase. The temperatures have been interpolated from meteorological measurements on the $1\,km \times 1\,km$ DEM, for the period between sowing of the crop and the observed date of a phenological event on a phenological station. The temperatures are adjusted for photoperiod on the specific location, which is dependent on latitude and day of year. A specific temperature sum is defined as the 45% quantile of the distribution of the temperature sums for all phenological stations in Germany that reported the target phase. Finally, the day on which this temperature sum is exceeded at each station is determined and interpolated on the Germany-wide grid.

The PHASE model as implemented in the *PHASE Analyser*[2] web application was utilized to extract all DWD-observed phenological phases of winter wheat, winter barley and winter rapeseed (see Table 1). This service is a *GeoServer*[3]-based open source map server application that provides daily updated phenological raster data. These data sets were created using the PHASE model that was implemented as a daily modelling routine. The results are distributed under strict application of *Open Geospatial Consortium* (OGC) standards via a *Drupal*[4]-based web interface. In Fig. 5, the subsequent processing steps are displayed, beginning with the automatic modelling, followed by the creation of the web services and visualisation. The underlying infrastructure utilises *php*-based parser applications to integrate the required input data automatically into a *PostgreSQL/PostGIS* database. These back-end procedures are required to call the *PHASE* model at a daily temporal interval ("Cronjob").

### 3.2 Spectral Separability Assessment

The separability analysis was exemplarily implemented for the three crop types winter wheat, winter barley and winter rapeseed, which dominate the Harz study site. These species are most relevant for regional modelling approaches or yield estimations, as they cover around 60% of the total agricultural areas in the study sites. The high number of available classes ($> 60$) would lead to inaccurate results, because they include numerous crop types that are only cultivated on a small number of fields. Thus, the analyses were performed only on fields that were covered by the target crops and other crop types cultivated at least on 25 fields on average over the study period, which includes summer-cropped wheat, durum wheat, barley and oats, as well as winter rye, maize, sugar beet, potatoes and perennial grassland. Consequently, the used fields covered more than 75% of all fields in the study area, but the number of classes was reduced by more than 80%.

Vegetation index profiles show specific (inter-) annual patterns that can be traced back to plant phenological phases (Förster et al. 2012) and are, along with single band reflectance values, powerful features for crop classification (Löw et al. 2013). Thus, a setting similar to the configuration presented by Löw et al. (2013) was implemented. It combined single-band reflectances and four well-established vegetation indices as spectral features (see Table 2). The Normalized Difference Vegetation Index (NDVI, Rouse et al. 1974), which is the most frequently applied vegetation index for agricultural remote sensing applications, shows a variety of issues that can be solved by including of other spectral
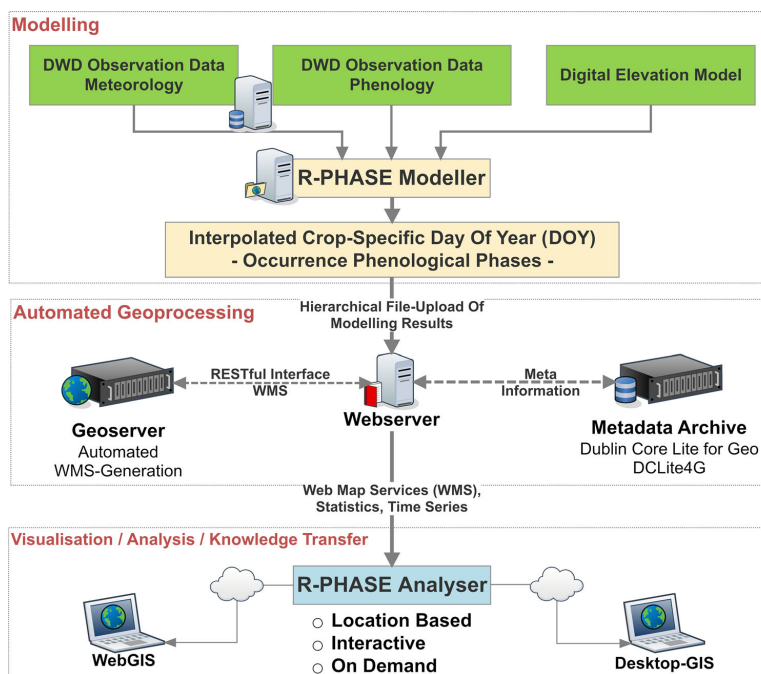
---

[2] http://phase.geo.uni-halle.de/phase-wms-dienste.

[3] http://www.geoserver.org.

[4] http://www.drupal.org.

**Fig. 5** Workflow of the automatic generation of modelled phenological phases



**Table 2** Vegetation indices and other spectral features used for the calculation of separability profiles

| Index | Abbreviation | Formula | References |
|---|---|---|---|
| Normalized Difference Vegetation Index | NDVI | $\frac{\rho_{NIR} - \rho red}{\rho_{NIR} + \rho red}$ | Rouse et al. (1974) |
| Green NDVI | GNDVI | $\frac{\rho_{NIR} - \rho green}{\rho_{NIR} + \rho green}$ | Buschmann and Nagel (1993) |
| Red edge NDVI | RE_NDVI | $\frac{\rho_{NIR} - \rho red\,edge}{\rho_{NIR} + \rho red\,edge}$ | Gitelson and Merzlyak (1994) |
| Wide-Dynamic Range Vegetation Index | WDRVI | $\frac{0.15 \times \rho_{NIR} - \rho red}{0.15 \times \rho_{NIR} + \rho red}$ | Gitelson (2004) |
| Blue band | B1 | $\rho_{blue}$ | Tyc et al. (2005) |
| Green band | B2 | $\rho_{green}$ | |
| Red band | B3 | $\rho_{red}$ | |
| Red edge band | B4 | $\rho_{red\,edge}$ | |
| Near infrared band | B5 | $\rho_{NIR}$ | |

$\rho_{region}$ reflectance in the specified spectral region

bands or mathematical modifications. Three of these modifications (Table 2) were incorporated into the classification scheme. Each of the modifications is more sensitive to a specific shortcoming of the NDVI, specifically the Red Edge NDVI (RE_NDVI) for the red edge region as a vitality indicator, the Green NDVI (GNDVI) for green chlorophyll content and the Wide-Dynamic Range Vegetation Index (WDRVI) for saturation of the band reflectances.

Classifications were applied at the field level, because object-based classifications are reported to be superior to pixel-based approaches when high resolution imagery is used

(De Wit and Clevers 2004; Blaschke 2010). A single-date classification approach was chosen to reflect the high probability that in some parts of the world, e.g., cloud cover limits the availability of multiple data sets useful for classification (Nitze et al. 2015; Frantz et al. 2017). For each acquisition date and parcel, mean values of the spectral bands and indices extracted from the RapidEye data were calculated.

Next, a classification scheme was set up that performs the widely used supervised random forest classification (Breiman 2001) implemented by Liaw and Wiener (2002) in the statistical computing environment *R* (R Core Team 2016).

🌀 Springer

The critical parameters node size and number of trees was set to 2 and 1000, respectively. This ensures stable results and reduces the chance of building over-fitted trees (Kuhn and Johnson 2013).

Three feature evaluations were performed per acquisition date, to find the best-suited feature for the separation of each target crop. Here, each spectral feature was classified sequentially. Evaluations were performed as binary classifications, i.e., features of fields covered by one out of the three target crops, winter wheat, winter barley and winter rapeseed, were classified against those from all other fields (see right box in Fig. 4). Finally, the $F1$- or $F$-measure, introduced by van Rijsbergen (1979), was calculated as a metric for the assessment of the class-wise classification accuracy. The $F$-measure is defined as the harmonic mean of the positive predictive value (precision) and the true positive rate (recall) of the classification confusion matrix. This statistic is especially meaningful for non-parametric classifiers and proved to be a good accuracy measure for random forest-based crop classifications (Löw et al. 2015). Precision, recall and $F$-measure were calculated based on the contingency table produced by the random forest implementation.

The specific optimal index per crop and acquisition date was then defined as the spectral feature with the highest $F$-measure. Based on the $F$-measure of all acquisition dates, a temporal profile of separability was computed (see Sect. 3.3).

### 3.3 Phenological Indexing of Inter-Annual Separability Patterns

All available modelled phases (see Table 1) for the three target crops for the years 2010–2015 were processed. Following Möller et al. (2017), a test site-specific phenological phase was considered as the period between the crop-specific medians of two subsequently beginning phases. The period between the first DOY and the first observed phase of a year was named after the last observed phase of the previous year.

The phenological indexing of separability patterns was started with the coupling of the separability profiles and the corresponding phenological phases. The duration of the last observed phase before harvest was prolonged by 1 week to address the stronger small-scale variations in harvesting due to the sequence of operations within farming cooperatives. The shift of the modelled harvesting ensured that the majority of all fields were actually harvested on the acquisition date. Furthermore, since the phases were linked to spectral response of the crops, the prolongation also addressed the fact that freshly harvested fields are often hard to distinguish from fields with fully ripened cereals using the spectral regions recorded by RapidEye. This similarity is due to the fact that

crop residues remaining on the fields spectrally resemble ripened crops before harvest.

A score value $R^P$ was calculated to assess the reliability of a phase to be optimal for separation of the corresponding crop type according to the following equation.

$$R^P = N_{obs}^P \times (F_{max}^P + F_{mean}^P) \tag{1}$$

$N_{obs}^P$ describes how often a phase has been identified as optimal for a year and study site. For these observations, $F_{mean}^P$ and $F_{max}^P$ represent the maximum and mean of the $F$-measures. This score value was required to account for differences of the temporal distribution of the satellite images. Such temporal gaps might possibly lead to undetected periods of high separability (see Fig. 3 and Sects. 1 and 2.2).

## 4 Results

### 4.1 Phenological Development

The phenological differences for all phases of the target crops between 2010 and 2015 vary among the sites, which is exemplarily shown for winter wheat in Fig. 6. For winter barley and winter rapeseed, these patterns are similar. In general, it could be observed that the growing season in DM starts with a delay of 1–4 days compared to HM and 5–12 days compared to HL, with the exception of 2015 when shooting started earliest in DM. The end of the growing season showed an opposite tendency, because the model predicted the start of the harvest period in DM 3–12 days earlier than in HL.

### 4.2 Spectral Separability Profiles

For each test site and each of the three investigated crop types, spectral separability profiles were computed from the
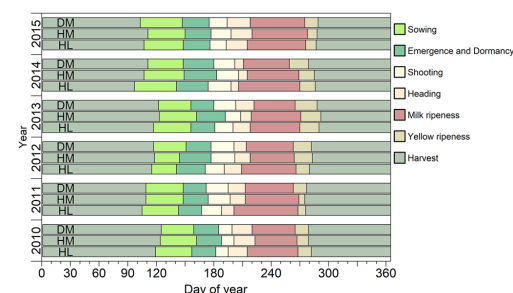


**Fig. 6** Modelled phenological phases of winter wheat in the three study sites for the years 2010 to 2015. The beginning of a phase is defined by the modelled average day of year of phenological event in the study site

🗍 Springer

**Fig. 7** Spectral separability profiles based on RapidEye data and modelled phenological phases for winter wheat, winter barley and winter rapeseed for the HL (**a**) and the HM (**b**) sites between 2010 and 2015. The spectral feature with the highest $F$-measure per season is noted above the respective peak in the profiles. Valid acquisitions are timed between sowing and harvest. *NDVI* Normalised Difference Vegetation Index, *WDRVI* Wide-dynamic Range Vegetation Index, *RE_NDVI* Red Edge NDVI, *GNDVI* Green NDVI
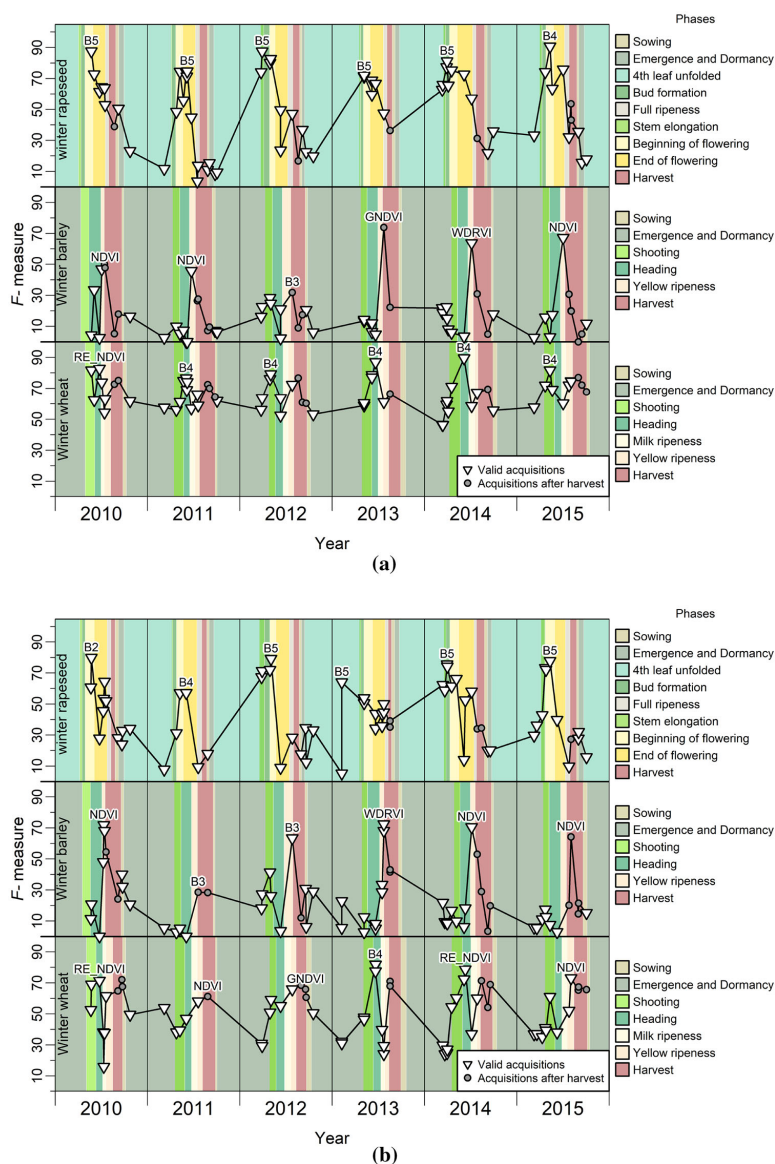


highest $F$-measure of all classified spectral features per random forest classification. The profiles are displayed as black lines in Fig. 7 separately for the two Harz sub-sites. Intra-annual patterns are visible that are characterized by peaks and depressions of the calculated $F$-measure profiles. For instance, winter wheat and winter rapeseed showed higher $F$-measures at their annual maximum and less intra-annual variations than winter barley.

**4.3 Phenological Indexing of High Separability Time Frames**

The RapidEye imagery of the HL and HM test sites, for which the annual $F$-measure was maximal, were coupled to the corresponding phenological phases (Fig. 7). The phase with the highest reliability value $R^P$ (Eq. 1) value is considered as the optimal phenological indicator phase. All other phases,

Ⓢ Springer

**Table 3** Indicator phases

| Crop | Phase | $N_{obs}^P$ | $DOY_{min}$ | $DOY_{max}$ | $F_{max}^P$ | $F_{mean}^P$ | $R^P$ | Spectral feature |
|---|---|---|---|---|---|---|---|---|
| Winter wheat | Heading | 7 | 153 | 175 | 0.89 | 0.81 | 11.94 | Band 4 |
| | Shooting | 4 | 122 | 130 | 0.82 | 0.80 | 6.48 | Band 4 |
| | Yellow ripening | 3 | 200 | 213 | 0.73 | 0.66 | 4.16 | GNDVI |
| Winter barley | Yellow ripening | 8 | 173 | 207 | 0.73 | 0.63 | 10.80 | WDRVI |
| | Shooting | 3 | 111 | 125 | 0.28 | 0.20 | 1.44 | NDVI |
| Winter rapeseed | Beginning of flowering | 6 | 123 | 143 | 0.91 | 0.81 | 10.31 | Band 4 |
| | Stem elongation | 2 | 88 | 88 | 0.88 | 0.88 | 3.50 | Band 5 |
| | Bud formation | 2 | 88 | 88 | 0.81 | 0.78 | 3.19 | Band 5 |
| | Fourth leaf unfolded | 3 | 274 | 37 | 0.64 | 0.33 | 2.90 | Band 5 |
| | End of flowering | 2 | 153 | 157 | 0.75 | 0.66 | 2.81 | Band 5 |

Number of observations ($N_{obs}^P$), earliest ($DOY_{min}$) and latest ($DOY_{max}$) observation, maximum ($F_{max}^P$) and average $F$-measure ($F_{mean}^P$), reliability ($R^P$) and spectral variable with highest $F$-measure (phases: see Table 1). The phases are ordered decreasingly according to their $R^P$ value

during which the annual separability reaches its maximum at least twice, are referred to as alternative phenological indicator phases. Table 3 summarizes the resulting indicator phases and corresponding separability metrics.

For winter barley, only the phases "yellow ripening" and "shooting" were detected to be optimal at least twice. "Yellow ripening" showed the highest $R^P$ value, since it was considered as optimal a total of eight times. In addition, both $F_{max}^P$ and $F_{mean}^P$ are significantly higher than for "shooting", which was selected three times. The WDRVI reached the highest $F$-measure during "yellow ripening", while during "shooting", the NDVI outperformed all other spectral features.

Three indicator phases could be found for winter wheat. "Heading" ($N_{obs}^P = 7$) outperformed "yellow ripening" ($N_{obs}^P = 3$) and "shooting" ($N_{obs}^P = 4$). $F_{max}^P$ and $F_{mean}^P$ are comparable for "shooting" and "heading", while for "yellow ripening", the $F$-measure metrics were significantly lower. The red edge band was the best-performing spectral feature both for "heading" and "shooting".

The phases usable as indicators for rapeseed classification are, as expected, dominated by "beginning of flowering", with a $R^P$ value of 10.30. During flowering, the NIR reflectance of rapeseed starts to decrease while the reflectance measured by the green spectral band, that is nearest to the yellow wavelength region, increases as a consequence of the intense yellow colour. These changes are visible in the reflectance spectrum and unique during that temporal period compared to the other crops that are still highly vital. Furthermore, four time frames are potential alternatives, namely the phases "stem elongation", "bud formation", "end of flowering" and "fourth leaf unfolded", which are characterized by very low reliabilities. The $F_{max}^P$ of 0.88 and 0.81 for "stem elongation" and "bud formation" almost reach the values for "beginning of flowering" (0.91). It is also noteworthy that "stem elongation" showed a higher $F_{mean}^P$ value than "beginning of flowering", but is observed as optimal much rarer.

## 4.4 Regional Transferability and Validation of Indicator Phases

The $F$-measure separability profiles of the DM study site, for which both environmental conditions and temporal distribution of the data sets are different compared to the two Harz sites (see Sect. 2), were also calculated and are visualised in Fig. 8.

The profiles clearly show a dependency on data availability, since not all indicator phases are represented at least by one data set in every year. Accordingly, the separability maximum for winter wheat was observed during "heading" in 2010 and 2011, which was found to be the optimal indicator phase for winter wheat separation based on the Harz sites analyses. Also, the annual maxima of the $F$-measure separability profiles for winter barley in 2010 and for winter rapeseed in 2012 exactly correspond to the detected optimal phases.

In 2012, no data set was available that represented "heading" for winter wheat. There, the separability maximum was observed during the alternative phase "shooting". However, this maximum is found six days before the optimal phase "heading" starts, which corresponds to the uncertainty caused by modelling, observation and small-scale differences between parcels (Gerstmann et al. 2016a).

The maximum separability for winter barley in 2011 was found during the alternative phase "shooting". In 2012, no data set was available during "yellow ripening" or "shooting". In general, the $F$-measures for winter barley are low ($< 0.5$) due to relatively small sample sizes ($< 20$) and thus the findings are only of limited reliability.

No data set was available during the optimal phase of winter rapeseed ("beginning of flowering") for 2010. The detected separability maximum occurred during the alternative phase "end of flowering". "End of flowering" was also detected as optimal for 2011, although a data set rep-
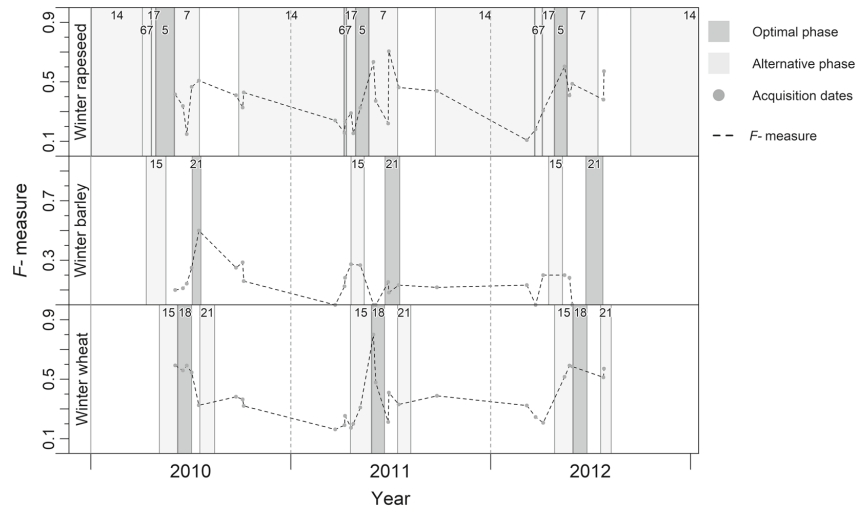
**Fig. 8** *F*-measure separability profiles for winter wheat, winter barley and winter rapeseed for the DM site and the phenological indicator phases detected through the analysis of the HL and HM test sites (see Fig. 7; Table 3). The numbers represent the phenological phases (see Table 1)

resenting the indicator phase "beginning of flowering" was available.

In summary, the separability maxima coincide with the defined phenological indicator phases (optimal or alternative) in eight out of nine crop-year-combinations. For three crop-year-combinations, no data set was available representing the optimal phase, which underlines the necessity for definition of alternative phases.

## 5 Discussion

### 5.1 Phenological Uncertainty

As the results showed, phenological phases can be used as indicators for acquisition time frames suitable for spectral discrimination of crops. However, as described in Sect. 4.4, sources of potential inaccuracy have to be considered when interpreting the results. In this study, this especially concerns the phenological interpolation results of the PHASE model, which also provides the kriging standard deviation as spatial inaccuracy metric (Gerstmann et al. 2016a). This metric can be used for additional interpretation of the separability profiles. Accordingly, the model uncertainty is expected to be between 3 and 8 days (Möller et al. 2017), with the majority of phases less than 5 days. Based on data of 2011 Gerstmann et al. (2016a) reported accuracies for winter wheat between < 4 and 11 days, for winter barley between 3.8 and 5 days and for winter rapeseed between 5 and 7 days. Modelling earlier phases was associated with higher uncertainty than that of

the later phases. Due to small-scale variations in plant phenology, the period of phase transition must be understood as a gradient rather than a sharp turnover. Phenological variations in turn are caused by management or cultivar-specific differences that cannot be observed by the monitoring network. Additionally, the reliability of volunteer-based observations is complicated to assess due to the subjectiveness of the observation design (Flanagin and Metzger 2008; Mehdipoor et al. 2015).

### 5.2 Sensor-borne Uncertainty

The defined phenological patterns also depend on the temporal composition of the satellite data sets. Time frames of high separability may remain undetected in situations of reduced image availability. Since the image acquisitions for the Harz data set are almost synchronous for both sub-sites, this factor is assumed to be negligible for this site. However, it becomes more important, e.g., for winter rapeseed at the Demmin site, for which the first acquisition in 2010 was available in the "end of flowering" phase only. In other words, this data set was acquired after the detected optimal phase "beginning of flowering", which underlines the necessity of having alternative indicator phases (see Sect. 5.4).

### 5.3 Separability Profiles

The separability profiles are based on a small set of spectral features which are well-established for a large variety of crop types. Other sensors such as Sentinel-2 provide additional

bands especially in the red edge and short-wave infra-red wavelength ranges useful for crop mapping (e.g., Immitzer et al. 2016). As the separability profiles are partly characterized by strong variations, optimized spectral indices, as proposed by Rivera et al. (2014) or Gerstmann et al. (2016b), could lead to more stable profiles.

The accuracies achieved by application of machine learning algorithms such as random forest generally tend to increase with the number of included predictor variables. Hence, this would be the preferred strategy instead of classifying each feature separately if the approach is applied in practise. The performance analysis of the spectral features could then be assessed by their variable importance, a parameter that ranks the predictor variables according to their contribution to the prediction accuracy. In doing so, both classification results and time frame reliability should increase. As a side-effect, the probability to produce an overfitted model would be reduced, although Breiman (2001) reported that random forest classifiers are relatively robust against overfitting to the training data compared to other classification approaches. However, in this study, we refrained from this option and applied a single-variable random forest approach instead to reduce the complexity for the demonstration of the approach and the results. Nevertheless, overfitting might be an issue when only one predictor variable is used as it is the case in this study, but several variations of the parametrisation (tree size, number of trees, number of terminal nodes) of the random forest classifier only slightly affected the results.

### 5.4 Alternative Indicator Phases

Alternative phases, that have been detected as optimal at least twice, can be valuable indicators for data set selection when no satellite data are available during the optimal phenological time frame. However, alternative phases can produce additional inaccuracies. For instance, the alternative phases "bud formation" and "stem elongation" for winter rapeseed showed comparably low $R^P$ values, since they represent short phases only. Thereby, exact start and end dates are complicated to define due to the modelling uncertainty of the beginning phenological phases (see Sects. 4.1 and 5.1).

From the botanical perspective, "bud formation" and "stem elongation" are actually more reliable than indicated by $R^P$. The selection of these early spring phases is due to shorter winter dormancy of rapeseed compared to the dormancy of cereal crops. This difference is measurable by vegetation indices as an earlier increase of photosynthetic activity. The alternative indicator phase "fourth leaf unfolded" occurs before winter dormancy and is characterized by significantly higher green vegetation coverage compared to winter cereals, which is also measurable by vegetation indices.

### 5.5 Validation and Regional Transferability

The regional transferability was proven for most of the investigated crops and years. However, as mentioned earlier in Sects. 4.4 and 5.2, no data set was available during the optimal phase for some crop-year combinations, which means that the expected separability maximum could not be found.

The bad performance ($F < 0.5$) for winter barley could be due to the fact that barley was cultivated on between 20 and 40 fields only. This number is remarkably lower than for winter wheat and winter rapeseed. Consequently, data gaps due to clouds or other factors had a stronger influence on statistical analysis than for wheat and barley, since sample sizes can easily fall below the lower threshold of approximately 30 samples for reliable statistical analyses.

### 5.6 Application Strategies

Both mono- and multi-temporal approaches can benefit from the presented methodology, once time frames have been defined also for other relevant crop types. For mono-temporal classification, a single data set should be selected for a point in time when the growth status of all target crops is within an optimal or alternative phenological indicator phase.

Multi-temporal classification should include at least one data set acquired during the optimal phase for all relevant crops at the study site. Thus, if the modelled indicator phases are available in the *PHASE Analyser* distribution system (see Sect. 3.1), they can be obtained and used for the user-specific application without the necessity of having a continuous temporal coverage of the satellite data over the complete growing cycle.

For multi-temporal classification, the approach would most likely lead to some congruent acquisition dates that are optimal for most of the crop types. For instance, the alternative phase "shooting" of winter wheat and winter barley is congruently timed with the optimal phase "beginning of flowering" for winter rapeseed. This supports the findings by, for instance, Schmidt et al. (2014) and Murakami et al. (2001). In their studies, a total of three to four images was found to be necessary to achieve a sufficient classification accuracy. However, the proof if the identified indicator phases are also valid in multi-temporal classifications is still an outstanding task.

The presented approach combined with the described web interface can assist data-intensive classification approaches based on satellite data of high spatial, spectral and temporal resolution like Landsat 8, Sentinel-2 or HJ-1, since it represents an effective strategy to reduce the calculation effort. This reduction is due to the focus on the most significant spectral features instead of all spectral bands provided by these sensors and on the reduction of acquisition dates to

⚫ Springer

be included in the classification during the most distinctive temporal periods.

Other applications which might benefit from phenological indicator phases and pre-assessment are hierarchical classification frameworks (e.g., Förster et al. 2012; Forkuor et al. 2015) and phenology-adaptive algorithms for compositing (Frantz et al. 2017). In the latter approach, the sub-images that are merged into phenology-adaptive, pixel-based image composites have to be selected to preferably represent an identical growth status of the land use classes. Especially in regions of persistent cloud coverage like Zambia, which Frantz et al. (2017) investigated in their study, the approach presented in this study can improve the timing of such target dates. Lastly, crop yield estimations for food security especially in regions of large inter-annual phenological variations (Meroni et al. 2014) might also be improved by the application of phenological indicator phases.

All these applications focus on the retrospective selection of already acquired remote sensing data. Beyond this, the spatially explicit knowledge of the phenological indicator phases can be applied to accurately schedule airborne flight campaigns or tasked satellite image acquisitions.

## 6 Conclusions and Outlook

This study presented a methodology that combines multi-annual crop separability profiles and automatically modelled phenological phases to derive indicator phases. These phases represent periods of time within a growing season during which spectral separability of a crop type is maximal, at least for the selected classification approach.

We tested the methodology for three frequently cultivated crops in Germany at two study sites under different environmental conditions and different data characteristics in terms of pre-processing and satellite image acquisition dates. The results showed that for each of the investigated crop types a minimum of two stable indicator phases exist.

The approach is currently spatially limited due to Germany's unique phenological observation network. However, the promising results suggest to put further research on its transferability to other regions of similar natural conditions.

Apart from crop mapping, the general framework of linking systematically collected phenological ground truth observations to vegetation reflectance patterns is ready to be tested for a large variety of other possible applications within an agricultural context. Thereby, the web interface providing phenological raster data for entire Germany enables various user groups to apply this approach of using phenological indicator phases to their own specific study regions.

## References

Azar R, Villa P, Stroppiana D, Crema A, Boschetti M, Brivio PA (2016) Assessing in-season crop classification performance using satellite data: a test case in Northern Italy. Eur J Remote Sens 49(1):361–380

Blaschke T (2010) Object based image analysis for remote sensing. ISPRS J Photogramm Remote Sens 65(1):2–16

Bogena HR (2016) Tereno: German network of terrestrial environmental observatories. J Large-scale Res Facil JLSRF 2:52

Borg E, Daedelow H, Apel M, Missling KD (2013) Rapideye science archive: Remote sensing data for the German scientific community. In: Borg E, Daedelow H, Johnson R (eds) RESA, GITO mbH Verlag, Berlin, RESA Workshop, Neustrelitz, vol 3, pp 5–20. http://elib.dlr.de/81718/

Breiman L (2001) Random forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/A:1010933404324

Buschmann C, Nagel E (1993) In vivo spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. Int J Remote Sens 14(4):711–722. https://doi.org/10.1080/01431169308904370

Conrad C, Rahmann M, Machwitz M, Stulina G, Paeth H, Dech S (2013) Satellite based calculation of spatially distributed crop water requirements for cotton and wheat cultivation in fergana valley, uzbekistan. Glob Planet Change 110:88–98

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

De Wit A, Clevers J (2004) Efficiency and accuracy of per-field classification for operational crop mapping. Int J Remote Sens 25(20):4091–4112

DWD Climate Data Center (2017a) Historical daily station observations (temperature, pressure, precipitation, sunshine duration, etc.) for Germany, version v005

DWD Climate Data Center (2017b) Multi-annual means of grids of air temperature (2m) over Germany, 1981-2010, version v1.0

DWD Climate Data Center (2017c) Phenological observations of crops from sowing to harvest (annual reporters, historical), Version v003

Flanagin AJ, Metzger MJ (2008) The credibility of volunteered geographic information. GeoJournal 72(3):137–148. https://doi.org/10.1007/s10708-008-9188-y

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(Suppl C):380–389. https://doi.org/10.1016/j.compag.2015.09.020

Förster S, Kaden K, Förster M, Itzerott S (2012) Crop type mapping using spectral-temporal profiles and phenological information. Comput Electron Agric 89:30–40

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

Gerstmann H, Doktor D, Gläßer C, Möller M (2016a) PHASE: a geo-statistical model for the kriging-based spatial prediction of crop phenology using public phenological and climatological observations. Comput Electron Agric 127:726–738

Gerstmann H, Möller M, Gläßer C (2016b) Optimization of spectral indices and long-term separability analysis for classification of cereal crops using multi-spectral RapidEye imagery. Int J Appl Earth Obs Geoinf 52:115–125

Gitelson AA (2004) Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. J Plant Physiol 161(2):165–173. https://doi.org/10.1078/0176-1617-01176

Gitelson AA, Merzlyak MN (1994) Quantitative estimation of chlorophyll-$a$ using reflectance spectra: experiments with autumn chestnut and maple leaves. J Photochem Photobiol B 22(3):247–252. https://doi.org/10.1016/1011-1344(93)06963-4

Guerschman J, Paruelo J, Bella CD, Giallorenzi M, Pacin F (2003) Land cover classification in the argentine pampas using multi-temporal landsat tm data. Int J Remote Sens 24(17):3381–3402

Immitzer M, Vuolo F, Atzberger C (2016) First experience with sentinel-2 data for crop and tree species classifications in central europe. Remote Sens 8(3):166

Inan H, Sagris V, Devos W, Milenov P, van Oosterom P, Zevenbergen J (2010) Data model for the collaboration between land administration systems and agricultural land parcel identification systems. J Environ Manag 91(12):2440–2454

Kaspar F, Zimmermann K, Polte-Rudolf C (2014) An overview of the phenological observation network and the phenological database of Germany's national meteorological service (Deutscher Wetterdienst). Adv Sci Res 11:93–99

Krauß T, d'Angelo P, Schneider M, Gstaiger V (2013) The fully automatic optical processing system CATENA at DLR. ISPRS Hannover Workshop 1:177–181

Kuhn M, Johnson K (2013) Applied predictive modeling. Springer, New York

Liaw A, Wiener M (2002) Classification and regression by randomforest. R News 2(3):18–22. http://CRAN.R-project.org/doc/Rnews/

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

Löw F, Knöfel P, Conrad C (2015) Analysis of uncertainty in multi-temporal object-based classification. ISPRS J Photogramm Remote Sens 105:91–106. https://doi.org/10.1016/j.isprsjprs.2015.03.004

Mehdipoor H, Zurita-Milla R, Rosemartin A, Gerst K, Weltzin J (2015) Developing a workflow to identify inconsistencies in volunteered geographic information: a phenological case study. PLoS One 10(10):1–14

Meroni M, Rembold F, Verstraete MM, Gommes R, Schucknecht A, Beye G (2014) Investigating the relationship between the inter-annual variability of satellite-derived vegetation phenology and a proxy of biomass production in the Sahel. Remote Sens 6(6):5868–5884. https://doi.org/10.3390/rs6065868, http://www.mdpi.com/2072-4292/6/6/5868

Möller M, Gerstmann H, Feng G, Dahms T, Förster M (2017) Coupling of phenological information and $NDVI$ time series: limitations and potentials for the assessment and monitoring of soil erosion risk. CATENA 150:192–205

Murakami T, Ogawa S, Ishitsuka N, Kumagai K, Saito G (2001) Crop discrimination with multitemporal SPOT/HRV data in the Saga Plains, Japan. Int J Remote Sens 22(7):1335–1348. https://doi.org/10.1080/01431160151144378

Nagol JR, Sexton JO, Kim DH, Anand A, Morton D, Vermote E, Townshend JR (2015) Bidirectional effects in landsat reflectance estimates: is there a problem to solve? ISPRS J Photogramm Remote Sens 103:129–135

Nitze I, Barrett B, Cawkwell F (2015) Temporal optimisation of image acquisition for land cover classification with random forest and modis time-series. Int J Appl Earth Obs Geoinf 34:136–146

Peña-Barragàn JM, Ngugi MK, Plant RE, Six J (2011) Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens Environ 115(6):1301–1316. https://doi.org/10.1016/j.rse.2011.01.009

R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Rabus B, Eineder M, Roth A, Bamler R (2003) The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. ISPRS J Photogramm Remote Sens 57(4):241–262

Richter R, Schläpfer D (2015) Atmospheric/topographic correction for satellite imagery (ATCOR-2/3 User Guide, Version 9.0.0, June 2015). DLR report DLR-IB, pp 565–01

Rivera J, Verrelst J, Delegido J, Veroustraete F, Moreno J (2014) On the semi-automatic retrieval of biophysical parameters based on spectral index optimization. Remote Sens 6(6):4927–4951. https://doi.org/10.3390/rs6064927

Rouse J, Jr RH, Schell JA, Deering D (1974) Monitoring vegetation systems in the Great Plains with ERTS, NASA SP-351. In: Third ERTS-1 symposium, vol 1. NASA, Washington, DC, pp 309–317

Schmidt T, Schuster C, Kleinschmit B, Förster M (2014) Evaluating an intra-annual time series for grassland classification—how many acquisitions and what seasonal origin are optimal? IEEE J Sel Top Appl Earth Obs Remote Sens 7(8):3428–3439

Tyc G, Tulip J, Schulten D, Krischke M, Oxfort M (2005) The RapidEye mission design. Acta Astronaut 56(1–2):213–219

van Niel TG, McVicar TR (2004) Determining temporal windows for crop discrimination with remote sensing: a case study in south-eastern Australia. Comput Electron Agric 45(1–3):91–108. https://doi.org/10.1016/j.compag.2004.06.003

van Rijsbergen C (1979) Information retrival, 2nd edn. Springer, Berlin

Whitcraft A, Vermote E, Becker-Reshef I, Justice C (2015) Cloud cover throughout the agricultural growing season: impacts on passive optical earth observations. Remote Sens Environ 156:438–447

Xu X, Conrad C, Doktor D (2017) Optimising phenological metrics extraction for different crop types in Germany using the moderate resolution imaging spectrometer (MODIS). Remote Sens 9(3):254. https://doi.org/10.3390/rs9030254

## 6.2 CONCLUSIONS FROM PAPER 3

Paper 3 revealed typical phenologically defined temporal patterns of separability for the three tested crop types. These patterns are stable for differently favoured ecological regions.
It also provides evidence that the uncertainty of the underlying phenological model is the key driver for inaccuracies within the time frame detection procedure.

The application of the presented methodology was limited to the three main single-species crop types in the study sites, namely winter wheat, winter rapeseed and winter barley, but analogue applications to other crop types are possible.

# 7

SYNTHESIS

The two strategies to improve land use classifications that have been investigated in this thesis (see Section 1.2), have been proven to be successful within particular classification scenarios. The dynamic optimisation of spectral indices increased the classification accuracies (Chapter 4). The determined separability patterns could be interpreted successfully by utilisation of a phenological model (Chapter 5) to link the separability profiles to modelled phenophases (Chapter 6).

## 7.1 METHODOLOGICAL ENHANCEMENTS

The following sections present continuations and applications of the principle methodology. Particularly, the continuation of the method to detect phenological time frames for the remaining target crops (Section 7.1.1) is described. Second, an application-oriented example classification is presented. There, the data set selection process is demonstrated and the actual set-up of the example classification is described (Section 7.1.2).

### 7.1.1 *Derivation of phenological time frames for all target crops*

The principle methodology is further used to determine the phenological indicator time frames for those target crops that were not considered within the research papers (see Tab. 3) and to recalculate the time frames for winter barley, winter rapeseed and winter wheat using a different parametrisation of the approach.

In each of the Chapters 4 and 6 simplifications were applied to the methodology. For the index optimisation approach (Chapter 4), that was applied to a subset of the Harz study site, the classification was performed for a binary scenario. Each of the classes was represented by a single crop type, namely winter barley and winter wheat.

The process of coupling the separability peaks to phenological phases (Chapter 6) was limited to the three target species winter wheat, winter barley and winter rapeseed and to a small number of NDVI variants instead of the full index optimisation approach for the definition of the optimal VIs. Furthermore, band ratios were not considered.

These restrictions are finally resolved and all options of methodology were used here:

1. the analyses are performed for both study sites,

2. the optimal index is allowed to be selected from individual spectral bands, standard indices, simple band ratios and optimised indices, and

3. all crops are analysed.

The basic rationale for the index optimisation, that was permuted in Chapter 4 (Paper 1), is identical to a parametrised formulation of the Enhanced Vegetation Index (EVI) formula (Eq. 1, Huete et al. 1999):

$$I = G \times \frac{B_1 - B_2}{B_1 + c_1 \times B_2 + c_2 \times B_3 + L} \tag{1}$$

In Eq. 1, $I$ denotes a VI, $B_1$, $B_2$ and $B_3$ represent spectral bands and $L$, $G$, $c_1$ and $c_2$ are empirical weighting parameters. For computational issues and the finding in Paper 1 that $G$ is not distinctive for the index optimisation, the domains of the parameters $L$ and $G$ are reduced here to [0.0, 0.5] and [1], respectively.

The $F1$-measure, which describes class-wise classification accuracy and thus spectral separability, is used for selection of the optimised indices. This metric was defined as the harmonic mean of precision and recall by Van Rijsbergen (1979). It is calculated according to Eq. 2:

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \tag{2}$$

Precision for a specific class is defined as the number of correctly classified fields (true positives) divided by the sum of correctly classified fields and the fields that are falsely assigned to the class ("false positives"). Recall is calculated similarly, but the number of false positives in the denominator is replaced by the number of fields of the target class that have been assigned to another class ("false negatives").

Subsequently, the time frame reliability for each observed phenological phase and crop type is computed according to Eq. 3:

$$R^P = N^P_{Obs} \times (F^P_{max} + F^P_{mean}) \tag{3}$$

In Eq. 3, $N_{Obs}^P$ represents how often a phase was found to be optimal, and $F_{max}^P$ and $F_{mean}^P$ represent the maximum and mean *F1*-measure for these phases. Since the detected time frames for the Harz site are also valid for the North-Eastern German Lowlands site (Chapter 6), the optimisation results for both study sites are combined.

Ideally, a crop type was cultivated and reported on a sufficient number of fields in the North-Eastern German Lowlands site, as well as in the western and the eastern part of the Harz site. Sufficient in this context means that a crop was cultivated on at least 6 fields during a season.

This ideal case was given for winter wheat, maize, winter rapeseed and winter barley for the complete study period. The resulting data pool consequently consists of 15 observations per crop type from:

- 3 years for the North-Eastern German Lowlands, and

- 6 years and each RapidEye tile in the Harz site.

### 7.1.2  *Application-oriented example classification*

On the example of the eastern tile of the Harz study site (tile ID: 3262922), an exemplary classification is performed to demonstrate the application of the detected phenological time frames for all crops.

From the complete time series spanning the 2015 growing season, 3 data sets are selected. The selection ensures that the chosen data sets reflect at least one indicator phase, optimal or alternative, for each of the target crop types. All other crop types are aggregated to an additional "other" class.

Next, several VIs are calculated from the selected data sets. These VIs include the indices used in Chapter 6: NDVI, WDRVI (Wide-dynamic range vegetation index; Gitelson 2004), Green NDVI and Red edge NDVI. Further, optimised indices per crop type are also calculated for the individual data set that captures a respective indicator phase.

A multi-temporal classification is then conducted based on all of these calculated indices. The ground truth data is randomly split up in 1/3 (1,129 fields) and 2/3 (2,273 fields) fractions for training and validation, respectively. The training fraction shows approximately the class distribution of the full tile. A random forest classifier (Breiman 2001), implemented by Liaw & Wiener (2002) in the statistical computing environment *R* (R Core Team 2017) is trained and applied for classification. A general analysis of the classification performance is performed by the calculation of overall accuracy. Further, class-wise performance based on the *F1*-measures (see Eq. 2) per class

is assessed. The accuracy assessment and the related discussions are conducted in front of a practical rather than a detailed theoretical background.

To compare the optimised classification to established multi-temporal classifications, the same workflow is performed based on input data that consists of the NDVI stack of all 11 available data sets acquired in 2015.

## 7.2 RESULTS OF THE ENHANCED METHODOLOGY

This section presents the phenological indicator time frames of all target crop and discusses the findings with regard to botanical and management characteristics. Furthermore, it presents the results of the application-oriented example classification and results of the comparison with a traditional multi-temporal classification approach.

### 7.2.1 *Derived phenological time frames for all target crops*

The derived phenological time frames are displayed in Fig. 5 including their reliability score and their earliest and latest phase-specific optimal observation DOY.

A cluster is apparent between DOYs 120 to 180, during which the most of the species show at least an alternative maximum. Except for the late harvested crops sugar beet and maize as well as for the evergreen perennial grassland, all maxima detected on DOYs later than 230 might actually represent the reflectance of catch crops, crop residues or bare soil. Consequently, these maxima should be interpreted carefully.

For all time steps, sites and crops, the optimised indices outperformed the established indices, band ratios and single spectral bands. The frequency of occurrence of a spectral band in the best-performing index on the optimal acquisition day is summarised in Tab. 7. There, also the reliability measure $R^P$ (see Eq. 3 and Chapter 6) is given for all relevant phenological windows.

The spectral band frequencies can function as guidelines in which spectral region a further optimisation might be useful to achieve more accurate results when satellite data of higher spectral resolution is used, such as provided by the Sentinel-2 sensors. The blue and the red bands are included mostly during early stages such as "heading" and "first cut for silage". Green is most relevant for the separation of rapeseed. The red edge and NIR bands are included in the optimised index most frequently, especially for cereal species, rapeseed and perennial grassland.
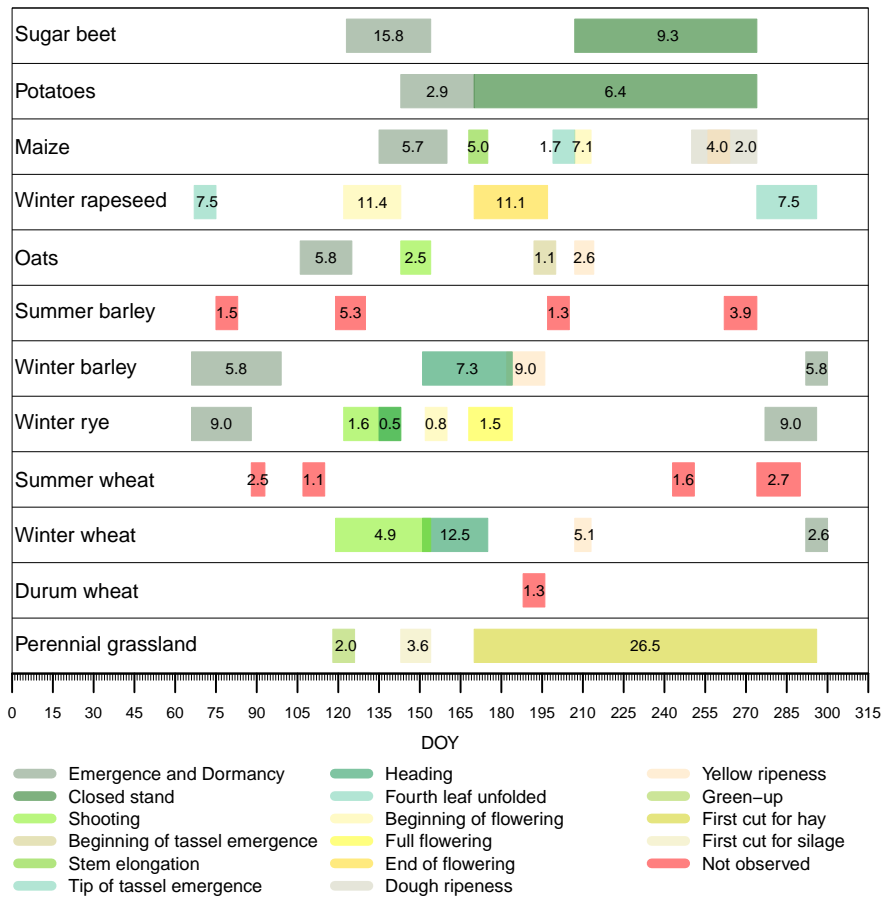
Figure 5: Detected phenological time frames for all target crops. The colour of the boxes corresponds to the phenological phases; red boxes represent species that are not part of the DWD phenological observation programme. Mixed colours indicate transitions between subsequent phases. The numbers printed in the phenological time frame boxes represent the time frame reliability $R^P$ (Eq. 3).

*Interpretation of the detected time frames*

Each of the target crops shows a distinct pattern of high separability time frames, during which the index composition and the separability can be interpreted according to the crop-specific cultivation practices.

PERENNIAL GRASSLAND    The separation of perennial grassland is best possible from June to the end of the growing season, since in June all other major crop types start to mature and reach ripening phases. This is related to a decrease of chlorophyll content and plant vitality. Due to its composition of different species and the multiple cutting and re-emerging, perennial grassland shows repeated vitality peaks during its full growing season. Furthermore, perennial grassland is harvested using scythes, so vital vegetation coverage remains on the fields during the entire season. Since all other crops are annual spe-

Table 7: Phenological indicator phases, time frame reliability ($R^P$, Eq. 3), frequency of a phase classified as optimal (N) and frequency of appearance of each spectral band in the optimised indices for all target crops that are monitored by the phenological observation programme

| CROP | PHASE | N | $R^P$ | B | G | R | RE | NIR |
|---|---|---|---|---|---|---|---|---|
| Perennial grassland | First cut for silage | 14 | 26.5 | 5 | 4 | 8 | 12 | 13 |
| Winter wheat | Heading | 7 | 12.5 | 3 | 2 | 4 | 5 | 7 |
| | Yellow ripening | 3 | 5.1 | 3 | 1 | 3 | 1 | 1 |
| | Shooting | 3 | 4.9 | 0 | 3 | 1 | 1 | 3 |
| Winter rye | Emergence and dormancy | 6 | 9.0 | 4 | 2 | 3 | 5 | 4 |
| Winter barley | Yellow ripening | 5 | 9.0 | 3 | 3 | 5 | 1 | 2 |
| | Heading | 5 | 7.3 | 4 | 4 | 1 | 5 | 1 |
| | Emergence and dormancy | 5 | 5.8 | 2 | 5 | 1 | 5 | 2 |
| Oats | Emergence | 5 | 5.8 | 5 | 2 | 4 | 2 | 2 |
| | Yellow ripening | 2 | 2.6 | 1 | 2 | 1 | 1 | 1 |
| Winter rapeseed | Beginning of flowering | 6 | 11.4 | 5 | 6 | 1 | 6 | 0 |
| | End of flowering | 6 | 11.1 | 1 | 4 | 5 | 2 | 5 |
| | Fourth leaf unfolded | 5 | 7.5 | 2 | 3 | 3 | 3 | 4 |
| Maize | Beginning of flowering | 4 | 7.1 | 1 | 0 | 3 | 4 | 4 |
| | Emergence | 3 | 5.7 | 0 | 1 | 3 | 3 | 2 |
| | Stem elongation | 3 | 5.0 | 1 | 3 | 2 | 1 | 2 |
| | Dough ripening | 3 | 4.0 | 1 | 1 | 2 | 3 | 2 |
| Potatoes | Closed stand | 8 | 6.4 | 2 | 3 | 4 | 1 | 6 |
| | Emergence | 3 | 2.9 | 3 | 1 | 2 | 2 | 1 |
| Sugar beet | Emergence | 9 | 15.8 | 6 | 5 | 8 | 4 | 4 |
| | Closed stand | 6 | 9.3 | 2 | 1 | 3 | 2 | 4 |

cies, they are removed completely from the fields or only dead crop residues are left on the field. The domination of red edge (12 appearances) and the NIR (13 appearances) channels in the optimised indices further confirms this high vitality during the whole growing season.

WINTER WHEAT    The separability maximum of winter wheat, characterised by an $R^P$ of 12.5, is related to "heading" (7 observations), since the heads and awns that are formed during that phase are of brighter reflectance than the green parts of the plant. Here, also the red and red edge bands are included most frequently in the optimised indices. The two alternative phases (3 observations, respectively), characterised by considerably less $R^P$ values of around 5.0, are "yellow ripening" and "shooting". During "yellow ripening", the blue and red channels are important for the optimised index, since the overall

reflectance increases during "heading". This leads to a brighter signal that is best detected using the visible bands. The green channel is consequently important for winter wheat separation during "shooting", when the plants turn green rapidly.

WINTER RYE    Winter rye is only of minor importance regarding its cultivation fraction in the two study sites. Consequently, when all agricultural fields are cultivated in summer, the rye-specific reflectance vanishes among the reflectance of the other crops. Hence, the best period to separate rye fields is after "emergence" in autumn or early spring with a fairly high $R^P$ of 9.0, until the summer crops start to emerge. Also no clear tendency is visible of which spectral regions are most important for rye detection, since bare soil, that dominates reflectance on the summer crop fields during the dormancy of rye, is characterised by overall low reflectance in the spectral regions covered by the RapidEye sensor and all visible bands are useful to separate vegetation from these soil-dominated fields.

WINTER BARLEY    A total of 3 phenological phases have been observed optimal on 5 times for winter barley separation. Of these phases, "yellow ripening" shows the highest maximum and mean *F1*-measure, leading to the highest $R^P$ measure. The very bright appearance is represented best by the visible channels. Consequently, blue, green and red are more frequently included in the optimised index than red edge and NIR. The red edge channel becomes more important during the alternative phases "heading" and "emergence / dormancy", which are the phases characterised by very high plant vitality.

OATS    The summer-cropped oats show their separability maximum during "emergence" (5 observations). This is typical for summer cereal species, since they are the only crop types that are vital in late March and April. Winter crops are still dormant during that time and the other relevant summer crops such as maize and sugar beet are planted and sowed later than summer cereals. The *F1*-measure for oats is comparatively low, leading to a fairly high $R^P$ value of 5.8.

WINTER RAPESEED    The typical yellow flowers determine the separability peak of winter rapeseed. Both flowering phases are characterised by a high $R^P$ value exceeding 11.0. It can be assumed that this separability maximum includes "beginning of flowering", "full flowering" and "end of flowering", unless only "beginning of flowering" and "end of flowering" are detected as optimal. The temporal gap between these two phases is a result of data gaps in the analysed time series and the lack of "full flowering" in the DWD observation programme for rapeseed. Interestingly, the NIR channel is never included in the

optimised index during "beginning of flowering", while red edge and green appear most frequently. This can be explained by decreasing photosynthetic activity of rapeseed during flowering that also determines a decreased green reflectance compared to the prior phases. However, during rapeseed flowering, all other crops are highly vital. Thus, the differences of green reflectance can be utilized to separate rapeseed from the other crops. As alternative phase, "Fourth leaf unfolded" has been detected 5 times as being also a period of high separability. This can be explained by the higher fractional vegetation coverage of the rapeseed rosettes compared to dormant winter cereal fields and the absence of vegetation on all fields that are not covered by winter crops or perennial grassland.

MAIZE    Since it is sown last during a growing season, the separability maxima of maize are located later than those of the other crops. In total, the phase "beginning of flowering" outperformed "emergence" and the other phenophases. Since three alternative phases have been detected at least three times, it can be stated, that the separation of maize from the other crops is possible with high accuracies during the full growing cycle of maize due to its delayed phenological cycle compared to the other crops.

POTATOES    The findings for the separation of potatoes are of fairly low reliability. This is, first, due to the high variety of the potato growing cycle. Harvesting of potatoes can be between June and September, depending on the respective cultivar. Further, potatoes are only cropped on small percentage of the study sites, which further reduces the potato separability. However, the optimal time frame is represented by the phase "closed stand", which can persist from May until September. Later-harvested cultivars can possibly be separated in late August and September, when the cereal crops and rapeseed are already harvested and only maize and sugar beet remained on the field. Contrarily, early harvested cultivars are hardly to separate from the other crops.

SUGAR BEET    Sugar beet can be separated with high accuracy during "emergence" and after all cereal species are harvested. "Emergence" of sugar beet, that occurs simultaneously with "emergence" of potato and maize, is the most suited time frame. This is due to the highly different seeding morphology. Sugar beet forms rosettes of leafs while potatoes and maize built a stem with several leafs. Consequently, fractional vegetation cover on sugar beet-planted fields increases faster than vegetation cover on fields cultivated with potatoes or maize. This difference is present in the reflectance signal, represented by a high frequency of visible spectral bands in the optimised indices.

DURUM WHEAT, SUMMER WHEAT AND SUMMER BARLEY    These three species are not covered by the DWD phenological observation programme, or were, in case of summer barley, only included into the programme in 2014 so that not enough observations are available for valid detection of phenological time frames. However, since they cover only a small percentage of the agriculturally used area, the detected separability maxima (independently of the plant-specific phenological behaviour), are randomly distributed along the growing season. Thus, no generally valid time frames could be detected for these summer cereal species.

### 7.2.2 *Application-oriented example classification*

The individual steps of the optimised classification are presented in detail in this section. Then, the selection of the optimal data sets is demonstrated practically. Second, the resulting classification is presented and discussed.

*Data set selection*

The example classification is based on the detected optimal temporal windows (Fig. 5) and the temporal distribution of the RapidEye data sets for the Harz lowlands sub-test site in 2015 (see Fig. 2 in Paper 3 / Section 6.1). The aggregated time frames that are understood as the temporal borders for data set selection should ensure that phenological indicator phases of as much crop types as possible are captured by the selected images.

A highly significant period for crop separation ranges from DOYs 120 to 160 when the distribution of the species-specific optimal time frames (Fig. 5) is considered. The optimal time frames for winter wheat and rye as well as for oats and winter rapeseed are enclosed in this period. For winter barley, the alternative phase "heading" is captured instead of the optimal phase "yellow ripening", but the difference of the $R^P$ value between these two phases is moderate. Furthermore, $R^P$ for "heading" is comparatively high. To classify the root crops, maize and perennial grassland, another aggregated time frame can be localized between DOYs 210 to 270. The highest inaccuracies during that time frame must be expected for the classification of potatoes, since their harvesting dates can vary significantly between cultivars.

The three data sets that have been selected are listed in Tab. 8. As most distinctive data set 2 August was selected, representing the optimal time frames for the spectral separation of maize, potatoes and perennial grassland. For sugar beet, oats and winter wheat, this data set was acquired during alternative time frames.

Table 8: RapidEye acquisitions from 2015 of tile 3262922 included in the classification experiment and the corresponding phenological time frames for the target crops

| TIME FRAME | ACQUISITION DATE | DOY | OPTIMAL TIME FRAME FOR | ALTERNATIVE TIME FRAME FOR |
|---|---|---|---|---|
| 1 | 2 August | 214 | Potatoes Maize Perennial grassland | Sugar beet Oats Winter wheat |
| 2 | 10 May | 130 | Sugar beet Winter rapeseed | Winter wheat Winter rye Summer barley |
| 3 | 1 July | 182 | Potato Perennial grassland | Winter barley Winter rye Winter rapeseed |

As the second-most distinctive data set, 10 May 2015 was selected, since it represents the optimal time frames for sugar beet and winter rapeseed and alternative time frames for winter wheat, winter rye and winter barley.

The inclusion of 1 July as third data set again displays optimal time frames for potatoes and perennial grassland. Alternative time frames for maize and summer barley are also captured by this data set.

Thus, after selecting three time frames, at least one data set captured the optimal indicator phase for seven out of nine phenologically observed target crops. Despite no data set was available that was acquired during the optimal time frames of winter wheat and winter barley, accurate separation of these two crop types is still possible since alternative time frames, two for winter wheat and one for winter barley, are captured by the three selected images.

*Classification results*

The results of the optimised and the traditional classifications are presented in Fig. 6. The false classified fields are highlighted for each of the classifications individually. Overall classification accuracies are 87.38 % for the optimised and 85.97 % for the traditional classification, respectively.

This overall gain in accuracy is not equally represented in every class (Fig. 7). For most of the classes, the *F1*-measures for the optimised classification are higher than for the traditional classification. The classifications of perennial grassland and summer barley represent two exceptions of this general tendency. For perennial grassland, this might be due to the long-lasting optimal time frame, which results in a high number of very significant data sets for this crop type. Regarding summer wheat, neither an optimal nor an alternative time

frame was captured by the selected images for the optimised classi-
fication, which results in the lowered classification accuracy for this
class.
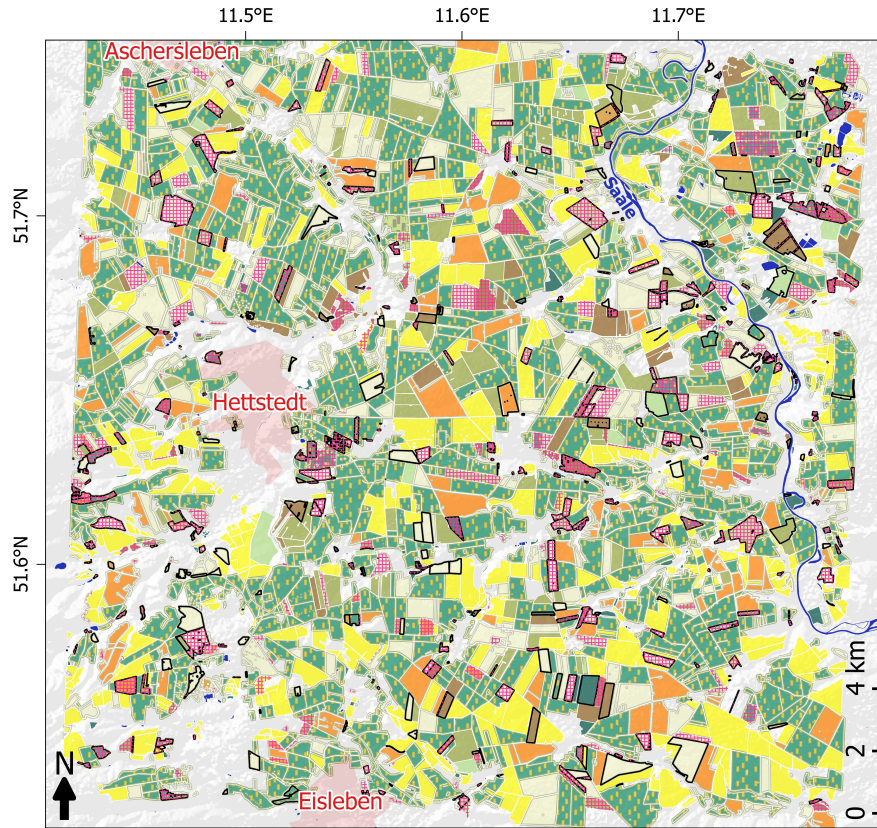


**False classified fields**

☐ Optimised classification (optimal time frames, optimised indices)

▨ Traditional classification (NDVI stack, all available data sets)

**Crops**

| | | | |
|---|---|---|---|
| ▮ Perennial grassland | ▮ Summer wheat | ▥ Oats | ▮ Sugar beet |
| ▮ Durum wheat | ▦ Winter rye | ▮ Winter rapeseed | ▮ Other |
| ▮ Winter wheat | ▮ Winter barley | ▮ Maize | |
| | ▮ Summer barley | ▮ Potato | |

Figure 6: Classification result of the eastern Harz study site in 2015. Col-
oured fields represent the actually cultivated crops. Non-hatched
fields without frame are classified correctly. Fields with black
frames are false classified by the optimised classification, hatched
fields are classified false by the NDVI classification. A hillshade of
$30 \times 30$ m resolution derived from the DEM is used as background.
Topographic information: © GeoBasis-DE / BKG 2018 and Open-
StreetMap contributors. Projection: WGS84, Pseudo Mercator

No significant difference could be observed for maize and the "other"
class. The highest improvements could be observed for winter rye,
winter barley and summer barley. The crop types with the highest

*F1*-measures are winter wheat, winter barley and winter rapeseed, which are the most frequently cultivated single-species crop types. Perennial grassland and maize, which also cover more than 5 % of the agricultural fields, are less accurately classified. For maize, the classification might be more accurate if a fourth data set is included that captures the periods when maize is almost the only vital plant on the fields. However, the effect of catch crops during that later periods of the season must be considered here.
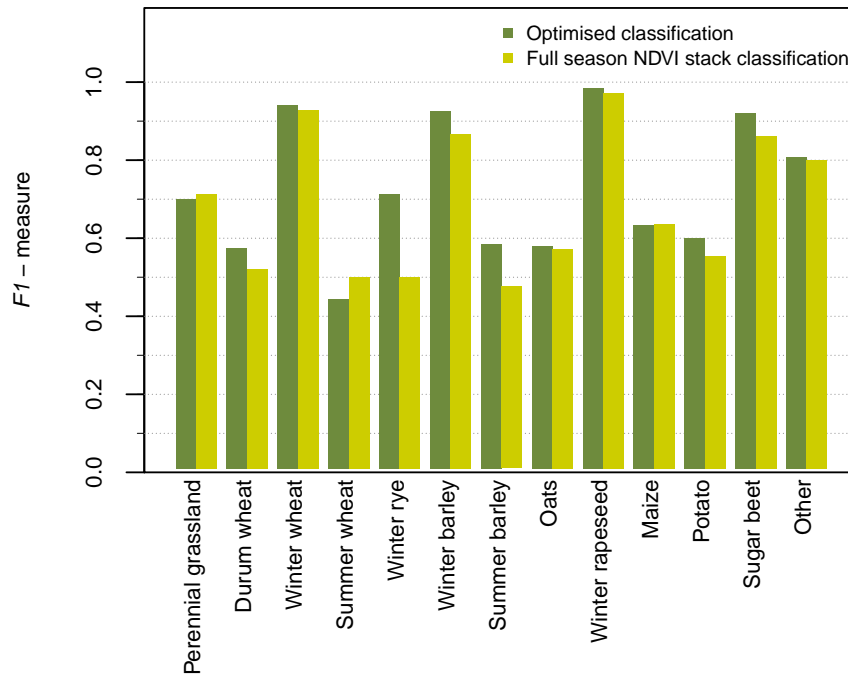


Figure 7: Class-wise *F1*-measure for traditional and optimised classification of the eastern sub-site of the Harz study site.

The overall number of falsely classified fields could be reduced. This reduction indicates the positive effect of the classification optimisation on overall accuracy, which can be traced back to most of the crop classes. With the traditional approach, 479 fields could not be correctly classified. For the optimised classification, their number decreased by 10% to 430 fields. A fraction of 279 fields was falsely classified by both classifications.

*Discussion of the classification experiment*

The classification and its inputs are not completely independent from the phenological time frame detection. Both the detection of the time frames as well as the classification were performed partly in the same area, which might lead to tendencies of overfitting. However, it can be assumed that spectral variations of the available reference objects and objects to classify are sufficiently high for the training of the clas-

sifier due to the large number of samples. This especially applies to frequently represented crops like winter wheat, winter barley, perennial grassland, maize and sugar beet. The classification uncertainties for the other crops (potatoes, winter rye, summer cereal species) are somewhat higher due to their relatively rare occurrence in the study site.

The findings of the classification experiment could be strengthened by further optimisation of sample sizes and the sampling design. For the tested splitting fractions of the fields, the improvement of the optimised classification compared to the traditional classification design could be demonstrated. Future research should address the verification of the results for smaller sample sizes. In doing so, the universal applicability of the approach could be substantiated. The sampling design is completely random, leading to congruent representation percentages of each crop type both in the training and the validation fractions. Consequently, the data base for highly represented crops like winter wheat is significantly larger than for other classes, leading to much less classification uncertainties for those classes. To prove the stability of the approach for other study regions, the crop types should be equally distributed in the training data set. However, this was not possible for the study site, since some crops, that are highly relevant for entire Germany, are significantly under-represented in the Harz region.

From the spatial perspective, the false classified fields are irregularly distributed without any obvious clusters. However, hidden dependencies on other factors like terrain, irrigation, fertilisation, soil types or differences between conventional and organic farming might be undetected.

Despite for the spectrally highly diverse 62 "other" fields which are false classified by both classifications, the class membership of the falsely classified fields can be used for further method refinement. For the remaining fields, various analysis can be applied to refine the optimised classification and to address potentially undetected dependencies to other variables. These analyses could, among others, investigate:

- the class distribution of the fields that are falsely classified once to detect class-specific method weaknesses of the optimised classification;

- the class distribution of fields that are falsely classified by both classifications to detect general difficulties for the classification of certain crop classes;

- the enrichment of the spectral information by additional variables such as terrain attributes or soil types to define variables that can be used as criteria for stratified sampling.

# COMPREHENSIVE DISCUSSION AND OUTLOOK

In Chapters 4 to 6 and Section 7.2.1 it could be demonstrated, that land use and land cover classification still is hampered by several shortcomings that can be addressed by consideration of additional data sources and technical effort. When these increased processing complexity is accepted by the user, the integration of the pre-processing results can result in land use classification of improved accuracy. In addition to the individual discussions about the components of the methodology addressed in each of the research papers and in Section 7.2.2, overall considerations regarding the interconnection of the components and possible subjects of future research are given here.

## 8.1 OPTIMISATION OF VEGETATION INDICES

The automated selection procedure of spectral indices can contribute to increased LULC classification accuracy, when the optimised indices are passed to the classifier as independent variables. The optimisation leads to significant improvements in classification accuracy compared to traditional multi-temporal classification approaches based on VIs such as NDVI, which could be proven both in Chapter 4 and Section 7.2.2. However, it was not assessed, at what extent the two optimisations, data set selection and index optimisation, individually contribute to the overall gain in classification performance.

### 8.1.1 *Computational effort*

The high computational effort of the index optimisation must still be evaluated to assess the efficiency of the approach compared to traditional techniques. Hence it must be evaluated whether the gain in accuracy justifies index optimisation as an additional pre-processing chain for operational classification frameworks or whether if other strategies are more effective. For a region of limited spatial size the effort might be inefficient, especially when the ground truth data base can be sufficiently large by applying traditional ground truth sampling methods.

### 8.1.2 *Natural conditions*

Additional uncertainty remains due to different natural conditions between training and application areas. The more the region differs regarding its internal ecological characteristics, the more uncertain is the reliability of the optimisation results. Influences of soil types, highly differing annual precipitation sums as well as irrigation practises reduce the universal applicability of the optimised indices even if the same crop types are investigated. So, if applied on large scale areas such as countries or on supra-national scale, separate optimisations should be performed to optimise indices suitable for specific natural regions, soil types, climatic regions or other stratification strategies based on additional environmental or management variables (Franklin & Wulder 2002; Metzger et al. 2005).

### 8.1.3 *Spectral separability indicators*

In Paper 1 (Chapter 4) the effect size measure $\eta^2$ was proposed as spectral similarity indicator since, as a non-parametric statistic, it is insensitive against non-normally distributed data sets. The metric is best-performing on very large sample sizes, which is given for raster data of moderate or high geometric resolution such as provided by RapidEye, Landsat or Sentinel-2 satellite sensors. For the time frame derivation, the raster-based classification approach was replaced by an object-based classification, which was found or referred to be superior for crop classifications in several studies (Blaschke 2010; Löw et al. 2013, 2015). For this purpose and to allow cross-comparison with these studies and the results presented here, the $\eta^2$ metric was replaced by the *F1*-measure as separability indicator in Paper 3 (Chapter 6) and the derivation of the time frames for the remaining target crops as well as for the classification experiment. However, the comparison of the actual effects of the two metrics can be subject of future research.

### 8.1.4 *Atmospheric effects optimised indices*

Atmospheric effects such as aerosol content are also effective on most VIs (Huete & Liu 1994). Correction algorithms can remove the most influencing atmospheric disturbances on reflectance. However, every model is based on assumptions and generalisations that prevent it to reach 100% accuracy. Hence atmospheric correction algorithms can only aspire to transform the recorded top-of-atmosphere reflectance as close as possible to actual bottom-of-atmosphere reflectance. So, when optimised indices are computed for a certain year or location,

it might be that these indices are not necessarily the best-performing indices in another year or another site, even when the phenological situation is exactly equal.

### 8.1.5 *Complexity of the optimisation implementation*

The optimisation can be performed in different complexity settings and still has potentials to be further improved. The domain of the weighting parameters, through which the algorithm iterates, can be theoretically extended without restrictions, but with the consequence of rapidly rising computational effort. In Paper 1 (Chapter 4), this complexity was set relatively high (see Table 2 of Paper 1 / Chapter 4). For computational reasons, the weighting parameter domains are diminished for the derivation of the phenological time frames, which decreases the possible number of indices that are compared. However, according to Paper 1 (Chapter 4), this effect is expected to be negligible for the absolute factor $G$ (see Eq. 1), since $G$ simply stretches the range of the resulting index. The influence of the domain reduction of $L$ was not assessed in this study.

Further, only the Enhanced Vegetation Index (EVI, Huete et al. 1999) rationale and band ratios function as mathematical framework for the iteration procedure. Implementation of alternative types of indices according to Rivera et al. (2014) can further improve the algorithm performance, although the EVI rationale is the base for the vast majority of established VIs.

RapidEye data, which were used as data base for method development, are of relatively low spectral resolution. Other, more recent multi- and super-spectral sensors such as Landsat-8 or Sentinel-2, collect reflectance in additional spectral regions (Drusch et al. 2012) or with finer spectral resolution. Especially the three red edge bands of the Sentinel-2 sensor and the SWIR of Landsat-8 and Sentinel-2 proved to provide highly significant information for crop separation (Immitzer et al. 2016). If the optimisation is applied on these sensors, the computational effort increases exponentially with the number of bands that are used to find the optimised indices.

As presented in Chapter 4, the optimisation is implemented in a binary design, to optimise an index for the separation of one land cover class from another. Other optimisation strategies could employ multi-class optimisation, to detect an index that is optimised for the classification of larger set of land cover classes. For instance, an index might be found which is optimised for the classification of winter wheat, winter barley and winter rye during the same classification run.

## 8.2 INTER-RELATION OF PHENOLOGY AND SPECTRAL SEPARABILITY

Phenology derived from ground observations was proven to be a stable indicator for crop separability in various studies (e. g. Haralick et al. 1980; Maus et al. 2016; Wardlow et al. 2007, see Section 2.1.4). This was verified in Chapter 4, where spectral-temporal separability patterns could be clearly explained by phenological information based on ground observation data. The phenological model PHASE could be implemented to reduce the uncertainty-evoking factors of volunteered geographic information collections (Flanagin & Metzger 2008; Mehdipoor et al. 2015) by substantiating the visually collected point observations with a plant physiological background.

Förster et al. (2012) and Schmidt et al. (2014) already proved that the data derived by the DWD observations programme is of suitable quality to can be employed as variables for crop or grassland classifications. Their implications, that the phenological observations are linked to spectral-temporal profiles and separability patterns could be verified by the successful transferability assessment, which is presented in Chapter 6. The vegetation period starts later and ends earlier in regions of cooler climate and less fertile soils compared to highly favoured, warmer regions. The spectral separability profiles correspond to these differences, showing their maxima rather during the same phenological phase than during the same calendar period.

### 8.2.0.1 *Uncertainties in phenological modelling*

The spatially explicit phenological data produced by the model PHASE still is affected by uncertainty. Geostatistical interpolation and a strong generalisation of the growing-degree-days approach to one single indicator temperature sum account for these uncertainties. Similar to the index optimisation, also the calculation of the indicator temperature sum for stratified regions (Siebert & Ewert 2012) can have positive effects on the model performance and accuracy.

### 8.2.1 *Input variable selection*

Day length was included as proxy for latitude, but other factors that might be affecting plant phenology are not considered. This regards factors such as cloud coverage, global radiation, sea proximity and population density (Hense & Müller 2007).

Although precipitation was proved to be not significantly efficient for most of the phenological growing stages, it possibly has stronger influence in semi-arid climates of Southern Europe. Under semi-arid conditions it was reported that precipitation and water availability are often the limiting factor for phenological development and not only

for crop yield (Fu et al. 2014). Also, climate projections performed by the Intergovernmental Panel on Climate Change predict that in parts of Central Europe water could also become less available. This applies mainly for the regions in Eastern Germany (Kovats et al. 2014), including the eastern parts of the Harz study site. Thus, the PHASE model must be adjusted in the future in order to respect this increasing probability of water becoming the limiting resource for agricultural land use.

### 8.2.2 *Requirements for a supra-national observation programme*

Although it provides phenological data of unique density and thematic depth, deriving phenological information from remote sensing data instead of ground surveys is currently more promising than the establishment of a comparable observation programme on supranational scale. Recent studies report a gain in accuracy and reliability of phenological information from satellite data (White et al. 2014; Xu et al. 2017), but criticise the lack of links to ground observations. Although volunteered geographic information become more and more a source of scientific data and have the potential to close the gap between space- and ground-borne data, a pan-European monitoring network of spatially stable quality would be complicated to establish due to high demands in perspective of administration and harmonisation (Senaratne et al. 2017).

However, if such a network is approached, the spectrum of the observed species should be extended when the observation scale is planned to be enlarged. The proportion of cultivated crop types is not only determined by natural conditions, but also by socio-economic factors. For instance, triticale *(x Triticosecale)* plays a much larger role in southern-European and Polish agricultural production than in Germany but the natural conditions of Northern Poland are relatively similar to those of North-Eastern Germany. However, Poland is the greatest triticale producer worldwide and triticale is cultivated on larger percentages of the arable land than barley and rye (Rozkrut 2017). This would justify the integration of triticale to the monitored crops in a pan-national programme. Other examples are durum wheat in southern Europe and soy in France, Romania and Italy, where such considerations should be made accordingly.

## 8.3 PHENOLOGICAL INDICATOR PHASES

For all target crops, multi-annually stable phenological time frames of high separability could be defined. The selection of the data set to include into the sample classification experiment verified findings of Conrad et al. (2014), Murakami et al. (2001), Schmidt et al. (2014)

and Wardlow et al. (2007), that a small number of highly significant acquisition dates provide sufficient information for accurate LULC classifications. This is of high relevance in regions that are affected by cloud coverage during significant parts of the year. The findings presented in this thesis demonstrate that sophisticated data set selections reduce the required number of images and thus make operational and non-operational applications more robust against unfavoured weather conditions.

### 8.3.1  *Sample size of the ground truth data*

Phenology-based classifications have, as demonstrated by Zhong et al. (2011), the potential to decrease ground data sampling requirements significantly, when the methodology is applied on larger scales. However, the still necessary ground surveys must be designed well, so that they include reference fields for each target crop. The number of reference fields must be oriented on the estimated class fraction to capture the intra-class reflectance differences by management, terrain, soil, and other factors. But, to ensure that enough fields for rarer classes are available to train the classifier, these classes must be represented disproportionally high. For the demonstrated example classification, the used sample size for training is relatively high. So, to verify the prescribed assumptions, comparable classifications must be performed and assess in detail for different size of the training set of fields.

The required ground sampling effort is reduced further, as the samples, optimised indices and phenological time frames are also applicable in subsequent growing seasons. Such cross-annual applications of PBC have been investigated by Zhong et al. (2014). This study could verify these findings, since the time frames are computed based on a multi-annual pool of separability analyses and phenological coupling. Specifically, this implicates that a sampling strategy for a large scale monitoring initiative would be most effective when it collects ground truth information for the relevant ecological regions and the regionally dependent dominating crop types for a few number of subsequent years. Once this ground truth data pool is sufficiently large, it can be used for classification of any growing season, both retrospective or of subsequent seasons. Analogue to Maus et al. (2016) and Belgiu & Csillik (2018), the data set selection based on phenological time frames can be also transferred to regions other than the training site.

8.3.2   *Number of optimal time frames*

The incorporation of the fourth and fifth time frame does not provide information to further increase classification accuracies. This supports findings that additional variables can obliterate distinctive spectral features capable for accurate crop discrimination, which results in (almost) constant accuracy (Conrad et al. 2014) or declining classification accuracies with increasing number of scenes (Murakami et al. 2001).

The number of 3 images, that has been assumed to be optimal in this study, probably applies only in landscapes dominated by single-cropping agriculture such as Germany. The investigations of Conrad et al. (2014), for example, were performed in landscapes in Uzbekistan in which double-cropping of major crop types, e.g. wheat followed by rice or mixed-crop cultivation, is much more common. To ensure that both crop types of the double cropping systems can be classified, the number of optimal time frames must probably higher since two phenological seasons expire parallel to one phenological season of single-cropping farming systems.

8.3.3   *Impact of data gaps on classification accuracy*

Data gaps caused by clouds are an issue for remote sensing worldwide (Whitcraft et al. 2015). However, the higher the temporal distribution of optical satellite data is, the better classification algorithms and data set selection approaches can perform. Assuming a combination of the Sentinel-2, RapidEye and Landsat-8 imagery and the theoretical temporal resolution lower than 5 days, persisting cloudy conditions can still have the effect that for parts of the areas under investigation data sets are lacking. Furthermore, multi-sensor monitoring approaches are affected by other issues such as different spatial, spectral or radiometric resolution.

To overcome the limitations by cloudy conditions that might affect optical remote sensing data, other strategies can get valuable input from accurate phenological time frames are possible helpful that aim on increasing the temporal resolution of satellite image time series. For instance, image compositing approaches can mosaic cloud-free parts of different satellite images which are acquired during a single time frame to one cloud-free data set per relevant phenological time frame (Frantz et al. 2017). Other recently tested methods include multi-sensor merging such as optical and RADAR sensors, which are insensitive to clouds but also can be utilised to map crop types and growth stages (Joshi et al. 2016). Finally, simulation of dense time series from vegetation indices using daily acquiring, coarse resolution sensors (MODIS) and less frequently acquiring, high resolution sensors (Landsat, RapidEye, Sentinel-2) is another strategy. Möller et

al. (2017) published promising results from a simulation experiment of Landsat-8 and MODIS with the STAR-FM algorithm (Gao et al. 2006). They could show that the simulation accuracy is controlled by the ability of the actual Landsat image to properly represent the phenological phase of the Landsat image simulated by MODIS.

### 8.3.4  *Socioeconomic sources of uncertainties*

In Germany, typical catch crops such as grass species, trefoil or rapeseed are cultivated during the autumn and winter months and are harvested before ripening to ensure an ideal sowing date for subsequently grown main crop. The main purpose of growing catch crops is to improve the nitrogen content of the soil or to form a protective layer against soil erosion. Mainly the harvested catch crops are used for animal feed. Thus they are often harvested before reaching the ripening growing stages.

The unclear composition and land use of the parcels covered by catch crops is somehow random. This determines that no data set acquired later than 2 August and earlier than 10 May could be selected as optimal. Furthermore, a catch crop like rapeseed is cultivated after harvest of the main crop during that season. Afterwards it emerges and reaches the phenological stages of highest vitality. This phase possibly overlaps to the vitality maxima of the latest sown main crop types, maize and sugar beet, and hence results in a decrease of separability for these two main crops.

The regionally predominating intended use of the grown crops also influences the applicability of the phenological time frames. Barley grown for forage is usually harvested before yellow ripening. However, yellow ripening was detected as optimal phase for the separation of winter barley. So, if one selects an acquisition date for which the PHASE model predicted yellow ripening to be present, winter barley might be already harvested and hence cannot be classified correctly. Consequently, regional predominating usage of the crops must be considered when the phenological time frames are used for data set selection. Also, alternative phenological time frames are potentially to be preferred over the optimal time frame.

The phenological time frames are defined and tested for regions of relatively large field sizes, which thus provide a high number of pixels for averaging the reflectance value per field parcel. In agricultural regions with very fragmented landscapes such as large parts of Southern Germany, Austria, Romania, Poland, Greece and Italy, this number of pixels is much less and thus the fraction of mixed pixels increases. Consequently, the spectral similarity and thus classi-

fication results decrease in accuracy. Contrarily, small and medium-sized fields have less intra-field reflectance variations due to a more homogeneous terrain. This could possibly lead to the contrary effect, an increase of classification accuracy in such areas. The dimensions of these effects have to be assessed in the future.

## 8.4 APPLICATION SCENARIOS

The determined phenological time frames for crop separation can be utilised in multiple ways, e. g., to improve hydrological models, to estimate effects of changes in cropping patterns and structures on carbon capture and storage or other usage scenarios in regions of low density regarding crop phenological observations. Four exemplary scenarios are briefly described in the following subsections 8.4.1, 8.4.2, 8.4.3 and 8.4.4.

### 8.4.1  *Scenario 1: Changes in cropping area of oats*

From a historical perspective, oats, barley and rye have been the most important cereal species in Northern Europe until the twentieth century. The production of these species declined rapidly during the last decades because of (bio-) technical proceedings that allow more frequent cultivation of more demanding crops with higher yields under the same climatic conditions. However, recent diet trends and increasing awareness of health-related benefits of oats can possibly stimulate a renaissance leading to a future increase in oats production (Tiwari 2010). To estimate the trend of acreage of oats, an accurate classification scheme should be based on 2-3 satellite images per season. The first acquisition date can be used for separation of winter and summer crops, of which the latter are characterised by bare soil approximately until April. The second and most important data set should be acquired during the first two thirds of April, when oats usually emerge. During that time, no other crop shows a comparably high growing activity and parcels of high activity that have been covered by bare soil in the first time step are probably covered by emerging oats. To further increase the oats classification accuracy, one image acquired during "yellow ripening" between late July and early August can be additionally used, on which oats should appear very dry with a light brown colour.

### 8.4.2  *Scenario 2: Modelling water balance of transnational river catchments*

Hydrological modelling of large river catchments, that span multiple countries, is often hampered by country-specific limitations regarding the possibility to access cropping information on parcel scale due to

political and administrative differences (Conrad et al. 2013; Courault et al. 2016). For trans-national catchments that are completely contained by member states of the EU, such as the rivers Elbe, Oder or Tejo, these information are collected by the member states according to standardised criteria that have been defined as consequence of the EU's common agricultural policy (Pérez 2005). However, accessibility of these data is dependant on country-specific legal situations.

Other catchments that are of high interest for ecosystem modelling span several countries for which land use data on field scale is not collected or the access is restricted due to political conflicts between these countries. An example for this situation are the Amu Darya and Syr Darya catchments in central Asia (Chemin et al. 2004). In such areas, the presented methodology, applied on the regionally dominating crop types, can potentially reduce the associated hydrological model uncertainties.

### 8.4.3  *Scenario 3: Crop classification in inaccessible study sites*

As demonstrated by Zhong et al. (2011), especially phenology-based mapping approaches are suited for crop mapping for food supply assessment and water resources management in regions of limited ground truth availability. Foreign aid or famine relief require accurate crop maps of the affected countries that can assist these actions. However, some required data of these countries are hardly to access or to observe due to political isolation, such as North Korea, or due to security reasons e. g. in Somalia. To achieve accurate crop map products anyhow, one strategy is to train the methodology in adjacent countries of similar environmental conditions (e. g. South Korea or the Chinese province Jilin as well as Ethiopia or Kenya, respectively) but better data or field work situations. In doing so, phenological observations can be recorded in these alternative study sites and the time frames can be subsequently transferred to the isolated target countries.

### 8.4.4  *Scenario 4: Spatial and temporal identification of extreme weather events and conditions*

The monitoring of extreme weather conditions and events is crucial to adapt measures for farmers, support decision making and refining soil policies. This is especially relevant in the context of climate change. A precondition for an effective monitoring is the availability of indices representing the spatio-temporal dynamics of influencing factors like precipitation, temperature or fractional soil coverage. First, the monitoring of such erosion pattern requires up-to-date phase-specific soil coverage information on field scale. As shown by Möller et al. (2017), such dynamic data sets can be obtained by data

fusion of satellite imagery of different temporal and geometric resolution in combination with the phenological spatial data introduced in this thesis. Furthermore, designing agricultural weather indices for farmers' insurances is challenging, since complex and dynamic crop-climate relationships have to be considered. The modelling results of the PHASE model have been proven as suitable for the dynamic determination of phenological windows, which are sensitive to weather conditions of winter wheat (Möller et al. 2018).

## 8.5 CONCLUSIONS

This thesis addressed two strategies to improve land use classifications in heterogeneous agricultural regions. It presents evidence that vegetation indices can be optimised for a specific problem in an automated manner to support the selection of spectral features for land use classifications. Furthermore, phenological volunteer-based ground observation data were successfully modelled to spatially-explicit phenological raster data sets and could be employed to couple spectral separability to plant growth stages.

This coupling procedure leads to the derivation of phenological indicator phases, which can function as proxy for the detection of temporal windows of maximal spectral separability.
Finally, it could be demonstrated that both index optimisation as well as phenological indicator phases increase land use classification accuracy and can decrease the effort that is required for appropriate ground truth data collection campaigns.

The presented methodology is designed flexible regarding the crop types and application regions, as long as appropriate input data sets are available. Since large-scale remote sensing-based monitoring of agricultural landscapes increasingly gains importance in various contexts, the methods presented here can contribute to such initiatives facing problems relating to climate change and food security.

Asrar, G.; Fuchs, M; Kanemasu, E. & Hatfield, J. (1984). Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. *Agronomy journal* 76 (2), 300–306. DOI: `10.2134/agronj1984.00021962007600020029x`.

BGR (2015a). *Bodenübersichtskarte der Bundesrepublik Deutschland 1 : 200,000*. Bundesanstalt für Geowissenschaften und Rohstoffe.

BGR (2015b). *Geologische Übersichtskarte der Bundesrepublik Deutschland 1 : 200,000*. Bundesaanstalt für Geowissenschaften und Rohstoffe.

Bannari, A; Morin, D; Bonn, F & Huete, A. (1995). A review of vegetation indices. *Remote sensing reviews* 13 (1-2), 95–120. DOI: `10.1080/02757259509532298`.

Beggs, P. J. (2004). Impacts of climate change on aeroallergens: past and future. *Clinical & Experimental Allergy* 34 (10), 1507–1513. DOI: `10.1111/j.1365-2222.2004.02061.x`.

Belgiu, M. & Csillik, O. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment* 204, 509 –523. ISSN: 0034-4257. DOI: `10.1016/j.rse.2017.10.005`.

Belward, A. S. & Skøien, J. O. (2015). Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing* 103. Global Land Cover Mapping and Monitoring, 115 –128. ISSN: 0924-2716. DOI: `10.1016/j.isprsjprs.2014.03.009`.

Blaes, X.; Vanhalle, L. & Defourny, P. (2005). Efficiency of crop identification based on optical and {SAR} image time series. *Remote Sensing of Environment* 96 (3–4), 352 –365. ISSN: 0034-4257. DOI: `10.1016/j.rse.2005.03.010`.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (1), 2–16. DOI: `10.1016/j.isprsjprs.2009.06.004`.

Bleiholder, H.; Weber, E.; Feller, C.; Hess, M.; Wicke, H.; Meier, U.; Boom, T. van den; Lancashire, P.; Buhr, L.; Hack, H.; Hack, H. & Stauss, R. (2001). *Growth stages of mono- and dicotyledonous plants*. Tech. rep. Brunswick, Germany: Federal Biological Research Centre for Agriculture and Forestry. DOI: `10.5073/bbch0515`.

Bogena, H. R. (2016). TERENO: German network of terrestrial environmental observatories. *Journal of large-scale research facilities JLSRF* 2, 52. DOI: `10.17815/jlsrf-2-98`.

Bolton, D. K. & Friedl, M. A. (2013). Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agricultural and Forest Meteorology* 173, 74–84. DOI: `10.1016/j.agrformet.2013.01.007`.

Breiman, L. (2001). Random forests. *Machine learning* 45 (1), 5–32.

Chemin, Y.; Platonov, A.; Ul-Hassan, M. & Abdullaev, I. (2004). Using remote sensing data for water depletion assessment at administrative and irrigation-system levels: case study of the Ferghana Province of Uzbekistan. *Agricultural Water Management* 64 (3), 183–196. ISSN: 0378-3774. DOI: `10.1016/S0378-3774(03)00209-9`.

Chen, B.; Huang, B. & Xu, B. (2017). Multi-source remotely sensed data fusion for improving land cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 124, 27–39. DOI: `10.1016/j.isprsjprs.2016.12.008`.

Chuine, I.; Garcia de Cortazar-Atauri, I.; Kramer, K. & Hänninen, H. (2013). Plant Development Models. In: *Phenology: An Integrative Environmental Science*. Ed. by Schwartz, M. D. Dordrecht: Springer Netherlands, 275–293. ISBN: 978-94-007-6925-0. DOI: `10.1007/978-94-007-6925-0_15`.

Congalton, R. G.; Gu, J.; Yadav, K.; Thenkabail, P. & Ozdogan, M. (2014). Global Land Cover Mapping: A Review and Uncertainty Analysis. *Remote Sensing* 6 (12), 12070–12093. ISSN: 2072-4292. DOI: `10.3390/rs61212070`.

Conrad, C.; Rahmann, M.; Machwitz, M.; Stulina, G.; Paeth, H. & Dech, S. (2013). Satellite based calculation of spatially distributed crop water requirements for cotton and wheat cultivation in Fergana Valley, Uzbekistan. *Global and Planetary Change* 110, 88–98. DOI: `10.1016/j.gloplacha.2013.08.002`.

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. *Computers and Electronics in Agriculture* 103, 63–74. DOI: `10.1016/j.compag.2014.02.003`.

Courault, D.; Demarez, V.; Guérif, M.; Page, M. L.; Simonneaux, V.; Ferrant, S. & Veloso, A. (2016). 4 - Contribution of Remote Sensing for Crop and Water Monitoring. In: *Land Surface Remote Sensing in Agriculture and Forest*. Ed. by Baghdadi, N. & Zribi, M. Elsevier, 113–177. ISBN: 978-1-78548-103-1. DOI: `10.1016/B978-1-78548-103-1.50004-2`.

DESTATIS (2015). *Land- und Forstwirtschaft, Fischerei. Bodennutzung der Betriebe (Landwirtschaftlich genutzte Flächen)*. German. Ed. by Bundesamt, Statistisches. Statistisches Bundesamt, Wiesbaden.

DWD (2015). *Vorschriften und Betriebsunterlagen für die phänologischen Beobachter des Deutschen Wetterdienstes*. Offenbach, Germany: Deutscher Wetterdienst.

DWD (2018a). *Multi-annual grids of precipitation height over Germany 1981-2010*. DWD Climate Data Center (CDC) 1981-2010, version v1.0., Deutscher Wetterdienst.

DWD (2018b). *Multi-annual means of grids of air temperature (2m) over Germany*. DWD Climate Data Center (CDC) 1981-2010, version v1.0., Deutscher Wetterdienst.

Drusch, M.; Bello, U. D.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; Meygret, A.; Spoto, F.; Sy, O.; Marchese, F. & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment* 120. The Sentinel Missions - New Opportunities for Science, 25 –36. ISSN: 0034-4257. DOI: `10.1016/j.rse.2011.11.026`.

Epting, J.; Verbyla, D. & Sorbel, B. (2005). Evaluation of remotely sensed indices for assessing burn severity in interior Alaska using Landsat TM and ETM+. *Remote Sensing of Environment* 96 (3), 328–339. DOI: `10.1016/j.rse.2005.03.002`.

Feddema, J. J.; Oleson, K. W.; Bonan, G. B.; Mearns, L. O.; Buja, L. E.; Meehl, G. A. & Washington, W. M. (2005). The importance of land-cover change in simulating future climates. *Science* 310 (5754), 1674–1678. DOI: `10.1126/science.1118160`.

Filella, I; Serrano, L; Serra, J. & Penuelas, J (1995). Evaluating wheat nitrogen status with canopy reflectance indices and discriminant analysis. *Crop Science* 35 (5), 1400–1405. DOI: `10.2135/cropsci1995.0011183X003500050023x`.

Flanagin, A. J. & Metzger, M. J. (2008). The credibility of volunteered geographic information. *GeoJournal* 72 (3), 137–148. ISSN: 1572-9893. DOI: `10.1007/s10708-008-9188-y`.

Förster, S.; Kaden, K.; Förster, M. & Itzerott, S. (2012). Crop type mapping using spectral–temporal profiles and phenological information. *Computers and Electronics in Agriculture* 89, 30–40. DOI: `10.1016/j.compag.2012.07.015`.

Franklin, S. & Wulder, M. (2002). Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. *Progress in Physical Geography* 26 (2), 173–205. DOI: `10.1191/0309133302pp332ra`.

Frantz, D.; Röder, A.; Stellmes, M. & Hill, J. (2017). Phenology-adaptive pixel-based compositing using optical earth observation imagery. *Remote Sensing of Environment* 190, 331 –347. ISSN: 0034-4257. DOI: `10.1016/j.rse.2017.01.002`.

Fu, Y.; Zhang, H.; Dong, W. & Yuan, W. (2014). Comparison of phenology models for predicting the onset of growing season over the Northern Hemisphere. *PloS one* 9 (10), e109544. DOI: `10.1371/journal.pone.0109544`.

Gao, F.; Masek, J.; Schwaller, M. & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote sensing* 44 (8), 2207–2218. DOI: `10.1109/TGRS.2006.872081`.

Gitelson, A. A. (2004). Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. *J. Plant Physiol.* 161 (2), 165 –173. ISSN: 0176-1617. DOI: `10.1078/0176-1617-01176`.

Goodin, D. G.; Gao, J. & Henebry, G. M. (2004). The effect of solar illumination angle and sensor view angle on observed patterns of spatial structure in tallgrass prairie. *IEEE Transactions on Geoscience and Remote Sensing* 42 (1), 154–165. ISSN: 0196-2892. DOI: `10.1109/TGRS.2003.815674`.

Griffiths, P.; Linden, S. van der; Kümmerle, T. & Hostert, P. (2013). A Pixel-Based Landsat Compositing Algorithm for Large Area Land Cover Mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6 (5), 2088–2101. ISSN: 1939-1404. DOI: `10.1109/JSTARS.2012.2228167`.

Guerschman, J.; Paruelo, J.; Bella, C. D.; Giallorenzi, M. & Pacin, F (2003). Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. *International Journal of Remote Sensing* 24 (17), 3381–3402. DOI: `10.1080/0143116021000021288`.

Gómez, C.; White, J. C. & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 116, 55 –72. ISSN: 0924-2716. DOI: `10.1016/j.isprsjprs.2016.03.008`.

Haralick, R. M.; Hlavka, C. A.; Yokoyama, R. & Carlyle, S. (1980). Spectral-temporal classification using vegetation phenology. *IEEE Transactions on Geoscience and Remote Sensing* (2), 167–174. DOI: `10.1109/TGRS.1980.350269`.

Harris, S.; Veraverbeke, S. & Hook, S. (2011). Evaluating spectral indices for assessing fire severity in chaparral ecosystems (Southern California) using MODIS/ASTER (MASTER) airborne simulator data. *Remote sensing* 3 (11), 2403–2419. DOI: `10.3390/rs3112403`.

Hede, A. N. H.; Kashiwaya, K.; Koike, K. & Sakurai, S. (2015). A new vegetation index for detecting vegetation anomalies due to mineral deposits with application to a tropical forest area. *Remote Sensing of Environment* 171, 83–97. DOI: `10.1016/j.rse.2015.10.006`.

Hense, A & Müller, M (2007). Geostatistische Modellierung und Qualitätskontrolle von phänologischen Beobachtungen. *Promet* 33 (1), 7–13.

Huete, A; Justice, C & Leeuwen, W. (1999). *Modis Vegetation Index (MOD13) Version 3. Algorithm Theoretical Basis Document [M/OL]*.

Huete, A. R. & Liu, H. Q. (1994). An error and sensitivity analysis of the atmospheric-and soil-correcting variants of the NDVI for the MODIS-EOS. *IEEE Transactions on Geoscience and Remote sensing* 32 (4), 897–905. DOI: 10.1109/36.298018.

Hunt Jr., E. R.; Doraiswamy, P. C.; McMurtrey, J. E.; Daughtry, C. S.; Perry, E. M. & Akhmedov, B. (2013). A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *International Journal of Applied Earth Observation and Geoinformation* 21, 103 –112. ISSN: 0303-2434. DOI: 10.1016/j.jag.2012.07.020.

Immitzer, M.; Vuolo, F. & Atzberger, C. (2016). First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sensing* 8 (3/166). ISSN: 2072-4292. DOI: 10.3390/rs8030166.

Inan, H.; Sagris, V.; Devos, W.; Milenov, P.; Oosterom, P. van & Zevenbergen, J. (2010). Data model for the collaboration between land administration systems and agricultural land parcel identification systems. *Journal of Environmental Management* 91 (12), 2440–2454. DOI: 10.1016/j.jenvman.2010.06.030.

Jarvis, A.; Reuter, H. I.; Nelson, A. & Guevara, E. (2008). *Hole-filled seamless SRTM data V4, International Centre for tropical Agriculture (CIAT)*.

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 Sensing* 8 (1). ISSN: 2072-4292. DOI: 10.3390/rs8010070.

Khatami, R.; Mountrakis, G. & Stehman, S. V. (2016). A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sensing of Environment* 177, 89 –100. ISSN: 0034-4257. DOI: 10.1016/j.rse.2016.02.028.

Knight, J. F.; Lunetta, R. S.; Ediriwickrema, J. & Khorram, S. (2006). Regional scale land cover characterization using MODIS-NDVI 250 m multi-temporal imagery: A phenology-based approach. *GIScience & Remote Sensing* 43 (1), 1–23. DOI: 10.2747/1548-1603.43.1.1.

Kovats, R.; Valentini, R.; Bouwer, L.; Georgopoulou, E.; Jacob, D.; Martin, E.; Rounsevell, M. & Soussana, J.-F. (2014). Europe. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by Barros, V.; Field, C.; Dokken, D.; Mastrandrea, M.; Mach, K.; Bilir, T.; Chatterjee, M.; Ebi, K.; Estrada, Y.; Genova, R.; Girma, B.; Kissel,

E.; Levy, A.; MacCracken, S.; Mastrandrea, P. & White, L. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. Chap. 23, 1267–1326.

Liaw, A. & Wiener, M. (2002). Classification and Regression by randomForest. *R News* 2 (3), 18–22.

Lieth, H. (2013). *Phenology and seasonality modeling*. Vol. 8. Springer Science & Business Media.

Lloyd, D. (1990). A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing* 11 (12), 2269–2279. DOI: `10.1080/01431169008955174`.

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 Journal of Photogrammetry and Remote Sensing* 85, 102–119. DOI: `10.1016/j.isprsjprs.2013.08.007`.

Löw, F.; Conrad, C. & Michel, U. (2015). Decision fusion and non-parametric classifiers for land use mapping using multi-temporal RapidEye data. *ISPRS Journal of Photogrammetry and Remote Sensing* 108, 191 –204. ISSN: 0924-2716. DOI: `10.1016/j.isprsjprs.2015.07.001`.

Lunetta, R. S.; Knight, J. F.; Ediriwickrema, J.; Lyon, J. G. & Worthy, L. D. (2006). Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment* 105 (2), 142–154. DOI: `10.1016/j.rse.2006.06.018`.

Mather, P. & Tso, B. (2016). *Classification Methods for Remotely Sensed Data*. 2nd ed. Environmental engineering. CRC Press. ISBN: 9781420090741.

Maus, V.; Câmara, G.; Cartaxo, R.; Sanchez, A.; Ramos, F. M. & de Queiroz, G. R. (2016). A time-weighted dynamic time warping method for land-use and land-cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (8), 3729–3739. DOI: `10.1109/JSTARS.2016.2517118`.

Mehdipoor, H.; Zurita-Milla, R.; Rosemartin, A.; Gerst, K. & Weltzin, J. (2015). Developing a Workflow to Identify Inconsistencies in Volunteered Geographic Information: A Phenological Case Study. *PLoS ONE* 10 (10), 1–14. DOI: `10.1371/journal.pone.0140811`.

Meier, U. (2003). Phenological Growth Stages. In: *Phenology: An Integrative Environmental Science*. Ed. by Schwartz, M. D. Dordrecht: Springer Netherlands, 269–283. ISBN: 978-94-007-0632-3. DOI: `10.1007/978-94-007-0632-3_17`.

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

Meynen, E. (1959). *Handbuch der naturräumlichen Gliederung Deutsch-lands*. Vol. 2. 6-9. sl.

Möller, M.; Gerstmann, H.; Gao, F.; Dahms, T. C. & Förster, M. (2017). Coupling of phenological information and simulated vegetation index time series: Limitations and potentials for the assessment and monitoring of soil erosion risk. *Catena* 150, 192–205.

Möller, M.; Doms, J.; Gerstmann, H. & Feike, T. (2018). A framework for standardized calculation of weather indices in Germany. *Theoretical and Applied Climatology*. ISSN: 1434-4483. DOI: `10.1007/s00704-018-2473-x`.

Morain, S. A. (1974). Phenology and remote sensing. In: *Phenology and seasonality modeling*. Springer, 55–75.

Murakami, T.; Ogawa, S.; Ishitsuka, N.; Kumagai, K. & Saito, G. (2001). Crop discrimination with multitemporal SPOT/HRV data in the Saga Plains, Japan. *International Journal of Remote Sensing* 22 (7), 1335–1348. DOI: `10.1080/01431160151144378`.

Pérez, J. M. (2005). Cadastre and the Reform of European Union's Common Agricultural Policy. Implementation of the SIGPAC. *CT / CATASTRO Julio*.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Réaumur, M. (1735). Observations du thermométre faites á Paris pendant l'année 1735 comparées avec celles qui ont été faites sous la ligne á l' Ile de France, á Alger et en quelques-unes de nos íles de l'Amérique. *Académie Royale des Sciences*, 545–576.

Rivera, J.; Verrelst, J.; Delegido, J.; Veroustraete, F. & Moreno, J. (2014). On the Semi-Automatic Retrieval of Biophysical Parameters Based on Spectral Index Optimization. *Remote Sensing* 6 (6), 4927–4951. ISSN: 2072-4292. DOI: `10.3390/rs6064927`.

Rodriguez-Galiano, V.; Ghimire, B.; Rogan, J.; Chica-Olmo, M. & Rigol-Sanchez, J. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 67, 93 –104. ISSN: 0924-2716. DOI: `10.1016/j.isprsjprs.2011.11.002`.

Rozkrut, D., ed. (2017). *Statistical yearbook of agriculture*. Warsaw: Central Statistical Office of Poland.

Schmidt, G.; Schönrock, S. & Schröder, W. (2014). *Plant Phenology as a Biomonitor for Climate Change in Germany: A Modelling and Mapping Approach*. Springer.

Schowengerdt, R. (2006). *Remote Sensing: Models and Methods for Image Processing*. Elsevier Science. ISBN: 9780080480589.

Schuster, C.; Förster, M. & Kleinschmit, B. (2012). Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *International Journal of Remote Sensing* 33 (17), 5583–5599. DOI: `10.1080/01431161.2012.666812`.

Senaratne, H.; Mobasheri, A.; Ali, A. L.; Capineri, C. & Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science* 31 (1), 139–167. DOI: `10.1080/13658816.2016.1189556`.

Siachalou, S.; Mallinis, G. & Tsakiri-Strati, M. (2015). A Hidden Markov Models Approach for Crop Classification: Linking Crop Phenology to Time Series of Multi-Sensor Remote Sensing Data. *Remote Sensing* 7 (4), 3633–3650. ISSN: 2072-4292. DOI: `10.3390/rs70403633`.

Siebert, S. & Ewert, F. (2012). Spatio-temporal patterns of phenological development in Germany in relation to temperature and day length. *Agricultural and Forest Meteorology* 152, 44–57. DOI: `10.1016/j.agrformet.2011.08.007`.

Son, N.-T.; Chen, C.-F.; Chen, C.-R.; Duc, H.-N. & Chang, L.-Y. (2013). A phenology-based classification of time-series MODIS data for rice crop monitoring in Mekong Delta, Vietnam. *Remote Sensing* 6 (1), 135–156. DOI: `10.3390/rs6010135`.

Tehrany, M. S.; Pradhan, B. & Jebuv, M. N. (2014). A comparative assessment between object and pixel-based classification approaches for land use/land cover mapping using SPOT 5 imagery. *Geocarto International* 29 (4), 351–369. DOI: `10.1080/10106049.2013.768300`.

Tiwari, V. (2010). Growth and production of oat and rye. In: *Soils, Plant Growth and Crop Production*. Ed. by Verheye, W. Encyclopedia of life support systems. Food and agricultural sciences, engineering and technology resources Vol. 2. Eolss Publishers Company Limited.

Tyc, G.; Tulip, J.; Schulten, D.; Krischke, M. & Oxfort, M. (2005). The RapidEye mission design. *Acta Astronautica* 56 (1), 213–219. DOI: `10.1016/j.actaastro.2004.09.029`.

USGS (2004). *Shuttle Radar Topography Mission*. Global Land Cover Facility.

Van Niel, T. G. & McVicar, T. R. (2004). Determining temporal windows for crop discrimination with remote sensing: a case study in south-eastern Australia. *Computers and Electronics in Agriculture* 45 (1), 91 –108. ISSN: 0168-1699. DOI: `10.1016/j.compag.2004.06.003`.

Van Rijsbergen, C. (1979). *Information Retrival*. 2nd ed. Springer.

Van der Meer, F. D.; Werff, H. M. Van der; Ruitenbeek, F. J. van; Hecker, C. A.; Bakker, W. H.; Noomen, M. F.; Meijde, M. van der; Carranza, E. J. M.; Smeth, J. B. de & Woldai, T. (2012). Multi- and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation* 14 (1), 112–128. DOI: `10.1016/j.jag.2011.08.002`.

Waldner, F.; Lambert, M.-J.; Li, W.; Weiss, M.; Demarez, V.; Morin, D.; Marais-Sicre, C.; Hagolle, O.; Baret, F. & Defourny, P. (2015). Land Cover and Crop Type Classification along the Season Based on Biophysical Variables Retrieved from Multi-Sensor High-Resolution Time Series. *Remote Sensing* 7 (8), 10400–10424. ISSN: 2072-4292. DOI: `10.3390/rs70810400`.

Wardlow, B. D.; Egbert, S. L. & Kastens, J. H. (2007). Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. *Remote Sensing of Environment* 108 (3), 290–310. DOI: `10.1016/j.rse.2006.11.021`.

Whitcraft, A.; Vermote, E.; Becker-Reshef, I. & Justice, C. (2015). Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. *Remote Sensing of Environment* 156, 438–447. DOI: `10.1016/j.rse.2014.10.009`.

White, K.; Pontius, J. & Schaberg, P. (2014). Remote sensing of spring phenology in northeastern forests: A comparison of methods, field metrics and sources of uncertainty. *Remote Sensing of Environment* 148, 97–107. DOI: `10.1016/j.rse.2014.03.017`.

Wu, H. & Li, Z.-L. (2009). Scale issues in remote sensing: A review on analysis, processing and modeling. *Sensors* 9 (3), 1768–1793. DOI: `10.3390/s90301768`.

Xu, X.; Conrad, C. & Doktor, D. (2017). Optimising Phenological Metrics Extraction for Different Crop Types in Germany Using the Moderate Resolution Imaging Spectrometer (MODIS). *Remote Sensing* 9 (3/254). ISSN: 2072-4292. DOI: `10.3390/rs9030254`.

Zhang, S.; Liu, L.; Liu, X. & Liu, Z. (2016). Development of a New BRDF-Resistant Vegetation Index for Improving the Estimation of Leaf Area Index. *Remote Sensing* 8 (11), 947. DOI: `10.3390/rs8110947`.

Zhao, M.; Peng, C.; Xiang, W.; Deng, X.; Tian, D.; Zhou, X.; Yu, G.; He, H. & Zhao, Z. (2013). Plant phenological modeling and its application in global climate change research: overview and future challenges. *Environmental Reviews* 21 (1), 1–14. ISSN: 1208-6053. DOI: `10.1139/er-2012-0036`.

Zhong, L.; Hawkins, T.; Biging, G. & Gong, P. (2011). A phenology-based approach to map crop types in the San Joaquin Valley, California. *International Journal of Remote Sensing* 32 (22), 7777–7804. DOI: `10.1080/01431161.2010.527397`.

Zhong, L.; Gong, P. & Biging, G. S. (2014). Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sensing of Environment* 140, 1–13. DOI: `10.1016/j.rse.2013.08.023`.

# A

## APPENDIX

### A.1  LIST OF PUBLICATIONS

*Peer-reviewed journals*

Gerstmann, H.; Gläßer, C.; Thürkow, D. & Möller, M. (2018). Detection of Phenology-Defined Data Acquisition Time Frames For Crop Type Mapping. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 86 (1), 15–27. ISSN: 2512-2819. DOI: 10.1007/s41064-018-0043-6

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. *Computers and Electronics in Agriculture* 127, 726 –738. ISSN: 0168-1699. DOI: 10.1016/j.compag.2016.07.032

Gerstmann, H.; Möller, M. & Gläßer, C. (2016). Optimization of spectral indices and long-term separability analysis for classification of cereal crops using multi-spectral RapidEye imagery. *International Journal of Applied Earth Observation and Geoinformation* 52, 115 –125. ISSN: 0303-2434. DOI: 10.1016/j.jag.2016.06.001

Möller, M.; Doms, J.; Gerstmann, H. & Feike, T. (2018). A framework for standardized calculation of weather indices in Germany. *Theoretical and Applied Climatology*. ISSN: 1434-4483. DOI: 10.1007/s00704-018-2473-x

Möller, M.; Gerstmann, H.; Gao, F.; Dahms, T. C. & Förster, M. (2017). Coupling of phenological information and simulated vegetation index time series: Limitations and potentials for the assessment and monitoring of soil erosion risk. *Catena* 150, 192–205

Götze, C.; Gerstmann, H.; Gläßer, C. & Jung, A. (2017). An approach for the classification of pioneer vegetation based on species-specific phenological patterns using laboratory spectrometric measurements. *Physical Geography* 38 (6), 524–540

Kinkeldey, C.; Schiewe, J.; Gerstmann, H.; Götze, C.; Kit, O.; Lüdeke, M.; Taubenböck, H. & Wurm, M. (2015). Evaluating the use of uncertainty visualization for exploratory analysis of land cover change: A qualitative expert user study. *Computers & Geosciences* 84, 46–53

*Conference contribution as first author*

Gerstmann, H. & Gläßer, C. (2016). *Optimierung spektraler Indizes aus multi-sensoralen Satellitenbildzeitreihen für Landnutzungsklassifi-kationen*. German. Oral presentation. 5. gemeinsame Jahrestagung der Arbeitskreise Fernerkundung (DGfG) und Auswertung von Fernerkundungsdaten (DGPF), 28. - 30. September, Halle (Saale)

Gerstmann, H.; Möller, M.; Thürkow, D.; Knöfel, P.; Xu, X. & Gläßer, C. (2015). *Nutzung phänologischer Zeitfenster für satellitengestützte Klassifikation landwirtschaftlicher Nutzflächen*. German. Oral presentation. Deutscher Kongress für Geographie, 1 - 6 October, Berlin

Gerstmann, H.; Möller, M. & Gläßer, C. (2015). *Detetermination of phenological indicator phases for optimal satellite data set selection for land cover classification*. Conference paper, Poster presentation. 8th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multitemp), 22 - 24 July 2015, Annecy, France

Gerstmann, H.; Gläßer, C. & Götze, C. (2015). *Comparison of methods for hyperspectral change detection of pioneer vegetation in a post mining landscape in Eastern Germany*. Oral presentation. 9th edition of the European Association of Remote Sensing Laboratories (EARSeL) Special Interest Group on Imaging Spectroscopy workshop. 14 - 16 April 2015, Luxembourg

Gerstmann, H.; Möller, M.; Thürkow, F. & Gläßer, C. (2015). *Entwicklung eines OpenSource-WebGIS zur Modellierung von Pflanzenentwicklungsphasen*. German. Oral presentation. 7th Geofachtag, 18 February 2015, Halle

Gerstmann, H.; Gläßer, C. & Götze, C. (2014). *Veränderungsanalysen von Pioniervegetation in einer Bergbaufolgelandschaft in Mitteldeutschland mittels hyperspektraler HyMap-Flugzeugscannerdaten*. German. Conference paper, Poster presentation. 34th Wissenschaftlich-Technische Jahrestagung der DGPF, 62. Deutscher Kartographentag der DGfK, Geoinformatik 2014 der GfGI und des GIN, 26 – 28 March 2014, Hamburg

Gerstmann, H.; Möller, M.; Thürkow, D. & Gläßer, C. (2014). *PhenoS - A joint research project to improve land cover classification using Sentinel-2 satellite data*. Poster presentation. 5th Workshop of the EARSeL Special Interest Group on Land Use and Land Cover, 17 - 18 March 2014, Berlin

## A.2    CURRICULUM VITAE

*Personal Information*

- **Name:** Henning Gerstmann
- **Date of birth:** 27 May 1989
- **Place of birth:** Berlin
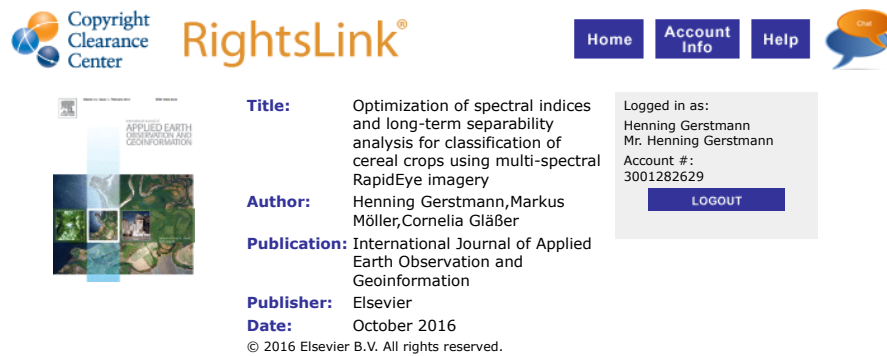- **Citizenship:** Germany

*Academic merits*

- 2017 – **Research associate** Martin Luther University Halle-Wittenberg, Institute for Geosciences and Geography

- 2013 – 2017 **Research associate** Martin Luther University Halle-Wittenberg, Institute for Geosciences and Geography, Department Remote Sensing and Thematic Cartography

- 2013 – 2018 **Doctoral student** Martin Luther University Halle-Wittenberg, Institute for Geosciences and Geography, Department Remote Sensing and Thematic Cartography

- 2012 **Exchange semester** Høgskulen is Sogn og Fjordane (now: Høgskulen på Vestlandet), Sogndal, Norway, Study programme: From Mountain to Fjord

- 2010 – 2013 **Academic degree: Master of Science in Geography** Martin Luther University Halle-Wittenberg (Major subject: Geospatial Data Handling, minor subject: Natural sciences).

- 2007 – 2010: **Academic degree: Bachelor of Science in Geography** Dresden Technical University (Major subject: Physical Geography, minor subjects: Meteorology, Water management).

- 2010 **Internship:** Alfred-Wegener-Institute for Polar and Marine Research, Potsdam. Research group: Sensitivity of Permafrost in the Arctic

- 2009 **Internship:** Municipal Environmental Agency, Dresden

PRE-ACADEMIC MERITS

- 1995 – 1999: Elementary school (14. Grundschule Dresden)
- 1999 – 2007: High School (Fritz-Löffler-Gymnasium Dresden) (Majors: Mathematics, English; Further exams in: Geography, German)

## A.3 REPRINTING PERMISSIONS OF THE PUBLISHERS



**Title:** Optimization of spectral indices and long-term separability analysis for classification of cereal crops using multi-spectral RapidEye imagery
**Author:** Henning Gerstmann, Markus Möller, Cornelia Gläßer
**Publication:** International Journal of Applied Earth Observation and Geoinformation
**Publisher:** Elsevier
**Date:** October 2016
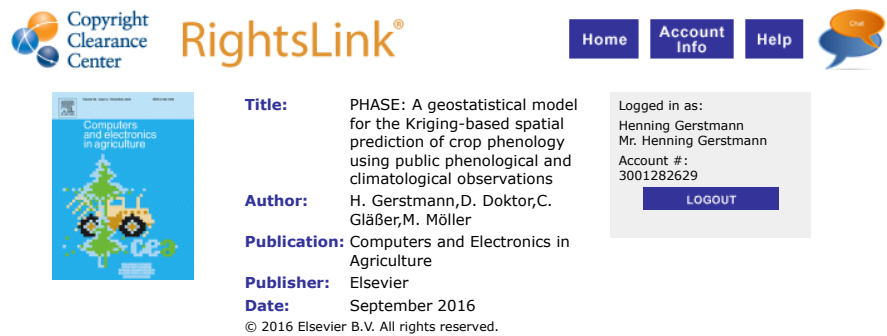© 2016 Elsevier B.V. All rights reserved.

Logged in as:
Henning Gerstmann
Mr. Henning Gerstmann
Account #: 3001282629

LOGOUT

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK   CLOSE WINDOW

**Title:** PHASE: A geostatistical model for the Kriging-based spatial prediction of crop phenology using public phenological and climatological observations
**Author:** H. Gerstmann, D. Doktor, C. Gläßer, M. Möller
**Publication:** Computers and Electronics in Agriculture
**Publisher:** Elsevier
**Date:** September 2016
© 2016 Elsevier B.V. All rights reserved.

Logged in as:
Henning Gerstmann
Mr. Henning Gerstmann
Account #: 3001282629

LOGOUT

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK   CLOSE WINDOW

Elsevier B.V.

**DEUTSCHE GESELLSCHAFT FÜR PHOTOGRAMMETRIE, FERNERKUNDUNG UND GEOINFORMATION (DGPF) E. V.**

DGPF e.V., Hochschule Neubrandenburg,
Brodaer Straße 2, 17033 Neubrandenburg

Herrn
Henning Gerstmann
Institute for Geosciences and Geography
Martin Luther University Halle-Wittenberg
Von-Seckendorff-Platz 4
06120 Halle (Saale)
Germany

**Hauptschriftleiter**

Prof. Dr. Wolfgang Kresse
Hochschule Neubrandenburg
Brodaer Straße 2
17033 Neubrandenburg

Tel.: +49 (0) 395 5693 4106
Fax: +49 (0) 395 5693 4199
E-Mail: kresse@hs-nb.de
http://www.dgpf.de

Neubrandenburg,
den 29.5.2018

Sehr geehrter Herr Gerstmann,

the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) hereby grants Henning Gerstmann, the author of the publication "Detection of Phenology-Defined Data Acquisition Time Frames For Crop Type Mapping" (DOI: https://doi.org/10.1007/s41064-018-0043-6), published in "PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science" (ISSN: 2512-1819) via Springer Nature, the right to reuse the article for a thesis/dissertation. This permission includes the reuse and publication of the non-translated article in its finally published layout both in printed and electronic form.

Hiermit erteilt die Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V. (DGPF), Henning Gerstmann, dem Erstautor des Artikels „Detection of Phenology-Defined Data Acquisition Time Frames For Crop Type Mapping" (DOI: https://doi.org/10.1007/s41064-018-0043-6), veröffentlicht in "PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science" (ISSN: 2512-1819) durch Springer Nature das Recht, den Artikel im Rahmen seiner Dissertation wiederzuverwenden. Diese Genehmigung umfasst die Wiederverwendung und -veröffentlichung des nicht-übersetzten Artikels sowohl in gedruckter als auch elektronischer Form.

Mit freundlichen Grüßen

*Wolfgang Kresse*

DGPF-Vorstand:
Prof. Dr. Uwe Stilla (Präsident)
Prof. Dr. Uwe Sörgel (Vizepräsident)
Prof. Dr. Wolfgang Kresse
Prof. Dr. Eberhard Gülch
Dr. Herbert Krauß

Geschäftsstelle der Gesellschaft: München
Eingetragen beim Amtsgericht München
Vereinsregister-Nr. VR 4564

Steuer-Nr. 143/212/60463

Deutsche Gesellschaft für Photogrammetrie und Fernerkundung

# SPRINGER NATURE LICENSE
## TERMS AND CONDITIONS

May 04, 2018

This Agreement between Mr. Henning Gerstmann -- Henning Gerstmann ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

| | |
|---|---|
| License Number | 4341851078270 |
| License date | May 04, 2018 |
| Licensed Content Publisher | Springer Nature |
| Licensed Content Publication | PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science |
| Licensed Content Title | Detection of Phenology-Defined Data Acquisition Time Frames For Crop Type Mapping |
| Licensed Content Author | Henning Gerstmann, Cornelia Gläßer, Detlef Thürkow et al |
| Licensed Content Date | Jan 1, 2018 |
| Type of Use | Thesis/Dissertation |
| Requestor type | academic/university or research institute |
| Format | print and electronic |
| Portion | full article/chapter |
| Will you be translating? | no |
| Circulation/distribution | <501 |
| Author of this Springer Nature content | yes |
| Title | Phenological and spectral optimisation of multi-temporal land use classification |
| Instructor name | Prof. Dr. Cornelia Gläßer |
| Institution name | Martin-Luther-University Halle-Wittenberg, Institute for Geosciences and Geography |
| Expected presentation date | Nov 2018 |
| Requestor Location | Mr. Henning Gerstmann<br>Von-Seckendorff-Platz 4<br><br>Halle, 06120<br>Germany<br>Attn: Mr. Henning Gerstmann |
| Billing Type | Invoice |
| Billing Address | Mr. Henning Gerstmann<br>Von-Seckendorff-Platz 4<br><br>Halle, Germany 06120<br>Attn: Mr. Henning Gerstmann |
| Total | **0.00 EUR** |
| Terms and Conditions | |

**Springer Nature Terms and Conditions for RightsLink Permissions**

**Springer Customer Service Centre GmbH (the Licensor)** hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

   If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where **print only** permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.

Springer Nature

3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.

5. Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.

7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.

10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval. However, the adaptation should be credited as shown in Appendix below.

### Appendix — Acknowledgements:

**For Journal Content:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

**For Adaptations/Translations:**
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

**Note: For any republication from the British Journal of Cancer, the following credit line style applies:**

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

**For Book content:**
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

**Other Conditions**:

Version  1.0

**Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.**

## 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*