



Suitability of satellite remote sensing data for yield estimation in northeast Germany

Claudia Vallentin¹ · Katharina Harfenmeister¹ · Sibylle Itzerott¹ · Birgit Kleinschmit² · Christopher Conrad³ · Daniel Spengler¹

Accepted: 7 June 2021 / Published online: 17 June 2021
© The Author(s) 2021

Abstract

Information provided by satellite data is becoming increasingly important in the field of agriculture. Estimating biomass, nitrogen content or crop yield can improve farm management and optimize precision agriculture applications. A vast amount of data is made available both as map material and from space. However, it is up to the user to select the appropriate data for a particular problem. Without the appropriate knowledge, this may even entail an economic risk. This study therefore investigates the direct relationship between satellite data from six different optical sensors as well as different soil and relief parameters and yield data from cereal and canola recorded by the thresher in the field. A time series of 13 years is considered, with 947 yield data sets consisting of dense point data sets and 755 satellite images. To answer the question of how well the relationship between remote sensing data and yield is, the correlation coefficient r per field is calculated and interpreted in terms of crop type, phenology, and sensor characteristics. The correlation value r is particularly high when a field and its crop are spatially heterogeneous and when the correct phenological time of the crop is reached at the time of satellite imaging. Satellite images with higher resolution, such as RapidEye and Sentinel-2 performed better in comparison with lower resolution sensors of the Landsat series. The additional Red Edge spectral band also has advantage, especially for cereal yield estimation. The study concludes that there are high correlation values between yield data and satellite data, but several conditions must be met which are presented and discussed here.

Keywords Vegetation Index · Yield estimation · Multispectral remote sensing · Precision farming

✉ Katharina Harfenmeister
katharina.harfenmeister@gfz-potsdam.de

¹ Helmholtz Centre Potsdam – GFZ German Research Centre for Geosciences, Remote Sensing and Geoinformatics, Potsdam, Germany

² Technische Universität Berlin, Geoinformation in Environmental Planning Lab, Berlin, Germany

³ Martin-Luther-Universität Halle-Wittenberg, Institut Für Geowissenschaften und Geographie, Halle, Germany

Introduction

Modern agriculture is increasingly primarily a digital agriculture and solutions to major challenges can hardly do without data. In precision farming, geographic information systems (GIS) and increasingly satellite data serve as a basis for the management. The flood of data from space is both a curse and a blessing. With the launch of the Sentinel satellite series of the European Space Agency (ESA), highly relevant data for agriculture due to their spectral, spatial, and temporal resolution are now freely available. In cloud-free conditions, a data set can provide information on the growth and condition of the crop up to once a week, allowing conclusions to be drawn on parameters such as biomass (Campos et al., 2019; Punalekar et al., 2018), plant density (Clevers et al., 2017; Pasqualotto et al., 2019) and yield (Battude et al., 2016; Hunt et al., 2019). Yield as a parameter is important for the farmer and the authorities, historically and during the season.

The derivation of yield from satellite, soil and relief data is not trivial because the yield formation of each plant species depends on complex factors and is different for every crop type (Geisler, 1988). The final yield of a field is mainly dependent on the number of seeds, soil type and therefore soil fertility, water and nutrient supply, and duration of sunshine throughout the season (Evans & Fischer, 1999; Geisler, 1988). The grain yield of cereals, for example, cannot be measured directly from satellite data, which is why those methods are based on proxies such as biomass (Babar et al., 2006; Ren et al., 2008), leaf area index (LAI) (Gasó et al., 2019; Peng et al., 2019) or chlorophyll content (Guo et al., 2018; Serrano et al., 2000). These proxies are often modelled using vegetation indices such as the Normalized Difference Vegetation Index (NDVI) (Bognár et al., 2017; Marti et al., 2007), which is considered to be reliable. Besides the NDVI, there are a number of further vegetation indices that correlate with plant parameters such as biomass, LAI or chlorophyll content, which can thus be related to yield (Barnes et al., 2000; Viña et al., 2011).

A meaningful correlation is not only a question of the right vegetation index, but also of the time of acquisition. The correlation between a vegetation index and the crop yield is not congruent in every phenological stage. For wheat, the phenological growth stages stem elongation (BBCH-Code 31), heading (BBCH-Code 51) and development of fruit until early ripening (BBCH-Code 75–83) of the BBCH scale (BBCH Working Group, 2001; Hack et al., 1992) are stated to be suitable to derive spatial yield patterns from satellite data (Knoblauch et al., 2017; Marti et al., 2007). One limiting factor, however, is the spatial resolution of the products of many publications. This is because yield models on a (sub) national or regional basis (Baruth et al., 2008; Ren et al., 2007, 2008) with spatial resolutions of 250 m (MODIS) to 1000 m (SPOT VEGETATION) can only be used to a limited extent in agricultural practice. In European countries, the average farm size of around 56 hectares (Statistische Ämter des Bundes und der Länder, 2011) is many times smaller than, for example, in the USA with around 447 hectares (Macdonald et al., 2013). Since precision farming commonly addresses variable conditions within the field (Finch et al., 2014), information at field level and particularly at within-field level is needed.

So far, correlation analyses between vegetation indices, soil data, relief data and yield data were mainly conducted for single crop types and years, or for specific sensors and a limited number of vegetation indices (Ali et al., 2019; Gómez et al., 2017; Panek et al., 2020; Zhao et al., 2020). A common application of correlation analysis is, for example, to find optimal hyperspectral bands correlating with yield data (Prey et al., 2020; Thenkabail et al., 2000; Zhang et al., 2018b). An extensive correlation analysis using multiple multi-spectral sensors and data of a large time period of over 10 years does not exist yet.

The hypothesis of this study is that there are optimal combinations of sensors, vegetation indices and time periods to achieve best possible yield estimations. These combinations are valid for various years and various agricultural fields of a certain crop type in a certain region. To test this hypothesis, this study analyzes the statistical correlation of high-resolution yield data and satellite data of various spatial and spectral resolutions and aims to answer the following questions: Which satellite sensor with which spatial, spectral, and temporal resolution is best suited for yield estimation of the individual crop? At what point in time do satellite data have to be acquired to obtain a robust estimate and can soil data like soil type, soil quality, and wetness as well as relief data do the same?

This study aims to provide a guideline for future approaches of yield prediction. Its aim is not to reach best possible correlations, but to give an objective overview about suitable sensors, vegetation indices, soil and relief data, time periods and further factors that should be considered or could be neglected for a successful yield prediction.

For this purpose, 947 yield data sets from wheat, barley, rye, and canola fields collected between 2006 and 2018 are analyzed. In total, 755 satellite images are available in this period, from which 15 different indices were calculated. Furthermore, soil data such as soil type, soil quality, and wetness index as well as relief data were also included in the analysis. The study evaluates for which crop type and which phenological stage yield estimation is mostly suited and which sensors with which spatial and spectral resolution are necessary.

Materials and methods

Study area

The study area is located in the north-eastern lowlands of Germany (Fig. 1). The analyzed fields are located within the DEMMIN test site, which is part of the TERENO project (Terrestrial Environmental Observatories) (Heinrich et al., 2018) as well as of the Joint

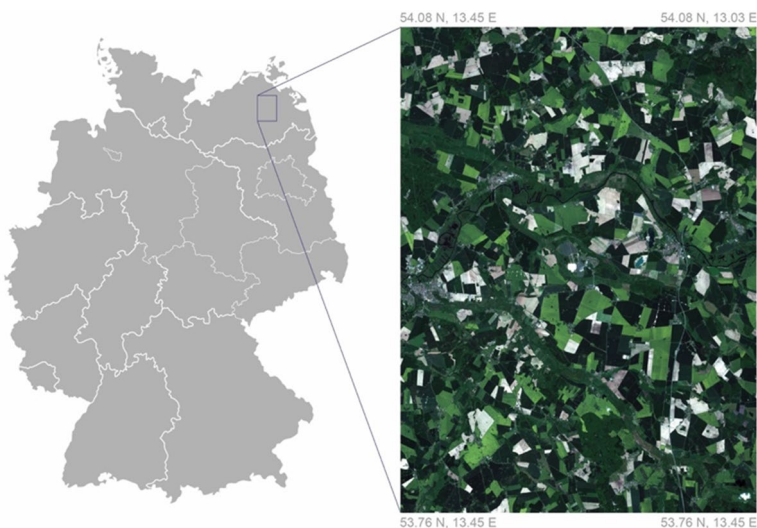


Fig. 1 Study area and location in Germany. Satellite image of Sentinel-2 satellite (ESA)

Experiment of Crop Assessment and Monitoring (JECAM) (Spengler et al., 2018). The test site is operated by Helmholtz Centre Potsdam – GFZ German Research Centre for Geosciences (GFZ). The central point of the study is around 53.948 N, 13.186 E. Geologically, the region was characterized by repeated glacial processes during the last glaciation and transformed into a young moraine landscape with representative glacier characteristics such as extensive, flat sand regions, hills and sinks as well as numerous lakes and bogs. The region is extensively used for agriculture today and sparsely populated. Agricultural fields in the study area are characterized by young moraine soil types, which are basically sandy and loamy, and mainly cultivated with cereals such as wheat, rye, and barley, but also with corn and canola. The average field size of the analyzed fields in this study is 43 ha.

The mean annual temperature is 8.8 °C according to the reference period 1981–2010 at the weather station Demmin operated by the German National Meteorological Service (Deutscher Wetterdienst (DWD) 2020). The yearly precipitation amount is around 600 mm.

Satellite remote sensing data

All available optical satellite images recorded by Landsat 5, Landsat 7, Landsat 8, RapidEye, Sentinel-2 and PlanetScope covering the study area between 2006 and 2018 were obtained (Fig. 2; Table 1). If possible, all data were obtained with radiometric and geometric corrections already applied to the data, namely as level L1TP (Landsat), 3A (Rapid Eye and PlanetScope) or 1C (Sentinel-2). RapidEye data at level 1B has no geometric correction, whereas Sentinel-2 data at level 2A is already atmospherically corrected and provides Bottom of Atmosphere (BOA) reflectance data.

Landsat images were downloaded at Level L1TP from the Earth Explorer provided by NASA. RapidEye images were downloaded from the RESA Archive (Planet Team 2020) at level 1B or 3A depending on availability. They are acquired with 6.5 m spatial resolution and resampled to 5 m by default. PlanetScope data at Level 3A were downloaded from the Planet API (Planet Team 2020). Sentinel-2 data were provided by the Copernicus Open Access Hub and were downloaded at level 1C and 2A depending on availability.

Atmospheric correction was performed on all data sets except for Sentinel-2 data using the ATCOR (Atmospheric and Topographic CORrection) software (Richter & Schläpfer, 2005). Atmospheric correction for Sentinel-2 data at level 1C was performed by the SICOR algorithm (Hollstein et al., 2016). The georeferencing of RapidEye and Sentinel-2

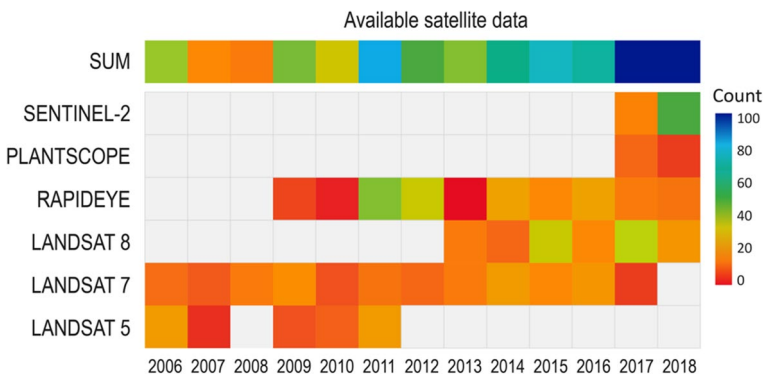


Fig. 2 Satellite data availability for each sensor in each year

Table 1 Overview of the availability and characteristics of the individual remote sensing sensors, as well as the performed processing steps

Sensor	Years	Data sets	Spatial resolution	Spectral bands	Processing steps
Landsat 5 Level L1TP	2006, 2007, 2009–2011	105	30 m	Blue, Green, Red, NIR	Atmospheric Correction (ATCOR)
Landsat 7 Level L1TP	2006–2017	280	30 m	Blue, Green, Red, NIR	Atmospheric Correction (ATCOR)
Landsat 8 Level L1TP	2013–2018	127	30 m	Blue, Green, Red, NIR	Atmospheric Correction (ATCOR)
PlanetScope Level 3A	2017–2018	21	3 m	Blue, Green, Red, NIR	–
RapidEye Level 1B, 3A	2009–2018	179	5 m	Blue, Green, Red, Red Edge, NIR	Atmospheric Correction (ATCOR) Coregistration (AROSICS)
Sentinel-2, Level 1C, 2A	2017–2018	43	10 m & 20 m (only Red Edge Bands)	Blue, Green, Red, 3 Red Edge Bands, NIR	Atmospheric Correction (SICOR, only Level 2A) Coregistration (AROSICS)

Please note, that not every data set covers the whole study area. Spatial and spectral resolution refer to those used in this study

data were partly not exact and additionally corrected with the software AROSICS (Automated and Robust Open-Source Image Co-Registration Software) (Scheffler et al., 2017) and manually by using the ArcMap software.

At the beginning of the period, only data from Landsat 5 and 7 are available, the number of sensors increases to five in 2017 (Fig. 2). For some years (e.g., 2013) only few data sets are available for the study area, mainly due to frequent cloud cover. The different spatial and spectral resolutions of the sensors (Table 1) are examined in the study for their correlation quality with yield data.

Soil and relief data

Soil type

Soil information is based on the German “Bodenschätzung” (BS) (1:10.000), which contains soil polygons with information about parent material, integrated soil texture to a depth of 1 m, and the soil development stage (AG Boden, 2005). Dobers et al. (2010) elaborate on the development and characteristics of the BS. The parameters “Bodenzahl” (BZ) and “Ackerzahl” (AZ) are quantitative assessments of soil fertility and indicators for potential agricultural productivity. They are given in integers from 0 to 100, where 100 is the reference for the most fertile soil in Germany. The BZ is based on soil type and therefore represents productivity only, while the AZ takes other factors such as morphology and climatic characteristics into account as well.

Organic carbon (C_{org})

A map of soil organic carbon (C_{org}) was also correlated with yield data. This map represents the distribution of organic carbon at the surface and was developed by Blasch et al., (2015a, b). They used a multispectral RapidEye time series from 2009 to 2014 to derive soil organic carbon content based on a regression between spectral reflectance pattern and soil sampling data.

Apparent electrical conductivity (EC_a)

Maps of apparent electrical conductivity (EC_a) from 2009 are available for Farm 1. These measurements of the soil sensor EM38 vary depending on soil type and condition. Consequently, soil properties such as the presence of soil organic matter, CaCO_3 , clay, and gleyic horizons can be derived from the measurements, among others (Kühn et al., 2008). The use of apparent electrical conductivity maps is common in precision agriculture (Corwin & Plant, 2005; Moral et al., 2010), therefore EM38 data serves as an input parameter for the correlation analysis as well. However, this parameter is highly variable as it depends on the current soil moisture during the measurement. The data are available as point data with a partly large point distance. Therefore, a 25 m buffer around the points was created and correlated with yield data located inside these buffers.

Digital elevation model

The digital elevation model (DEM) used in this study has a resolution of 5 m and is based on airborne LIDAR measurements (Amt für Geoinformation Vermessungs- und

Katasterwesen, 2011). The height data was used to calculate the Topographic Positioning Index (TPI) (Weiss, 2001) using the GIS software SAGA (Conrad et al., 2015). The TPI has generally six classes describing land forms such as hilltop, upper slope, etc. and is sensitive to the scales used in the calculation and classification process. The shape of classes changes if the outer radius of annulus in cells according to Weiss (2001) is changed. Therefore, two variations of the TPI were calculated: *TPI fine* with a radius of 25 m and *TPI coarse* with a radius of 50 m. The DEM was also used to calculate the Topographic Wetness Index (TWI) (Beven & Kirkby, 1979), which is a steady state wetness index and is a function of slope and the upstream contribution area.

Yield data

Field boundaries, information about crop types per field and year as well as yield data were provided by two farms located in the study area. The yield data was taken during harvest by a GPS controlled harvester. Yield measure was taken approximately every 1 m (Farm 1) or approximately every 10 m (Farm 2) within a tram line. Farm 1 provided yield data from 315 fields between 2006 and 2018, Farm 2 provided yield data from five large farm-wide data sets from 2012 and between 2014 and 2017. Yield data were cropped to individual field-based yield data sets using field boundaries. This resulted in 632 data sets of various density. The data quality of the yield data from Farm 1 is considered better because of the higher density and the more evenly distribution in the fields. Furthermore, processing and calibration steps of Farm 2 are unknown.

Raw yield data may contain errors due to unknown harvester characteristics, filling and emptying times of the harvester, time delay of the grain through the harvester, positional errors, and rapid speed changes, among others (Sudduth & Drummond, 2007). After acquisition, unreliable yield measurements were removed for the most part by applying filters on harvester speed (discarding of values <2% of all values and >99% of all values) and swath width (discarding of values <4 m and >9 m). Furthermore, statistical outliers were removed by grouping 50 point values and discarding of yield values with a difference of more than 2.5 times the standard deviation of this group. The threshold was tested iteratively to find the right balance between group size and number of outliers to ensure that real outliers are dismissed. Usually only single points and not whole zones are omitted, whereas the distance between the individual points is already smaller than the smallest pixel size of the satellite data used in this study.

Farm 1 operated precision farming over the entire time series with the aim of achieving homogeneous yields and applied fertilization by considering within-field variability. Farm 2, to the authors' knowledge, did not operate any precision farming during the analyzed period.

Phenology

Phenological data was provided by the German National Meteorological Service (DWD) and by GFZ according to the BBCH-Codes (Hack et al., 1992), which is a decimal code system to identify phenological development stages of a plant and the standard phenology-scale in Germany (Table 2). Phenological data of the DWD are part of a long-term observation series and report entry dates of selected phenological growth stages for various crop types at numerous observation points in Germany. They are daily updated by trained volunteers. For the years 2012 to 2018, field observations of BBCH collected by GFZ are

Table 2 Description of BBCH principal growth stages for Cereal and Canola according to Hack et al. (1992)

Principal growth stage	Cereal		Oilseed rape (Canola)	
	BBCH	Description	BBCH	Description
0	00–09	Germination	00–09	Germination
1	10–19	Leaf development	10–19	Leaf development
2	20–29	Tillering	20–29	Formation of side shoots
3	30–39	Stem elongation	30–39	Stem elongation
4	41–49	Booting	–	–
5	51–59	Inflorescence emergence, heading	50–59	Inflorescence emergence
6	61,65,69	Flowering, anthesis	60–69	Flowering
7	71, 73, 75, 77	Development of fruit	71–79	Development of fruit
8	83, 85, 87, 89	Ripening	80–89	Ripening
9	92, 93, 97, 99	Senescence	97, 99	Senescence

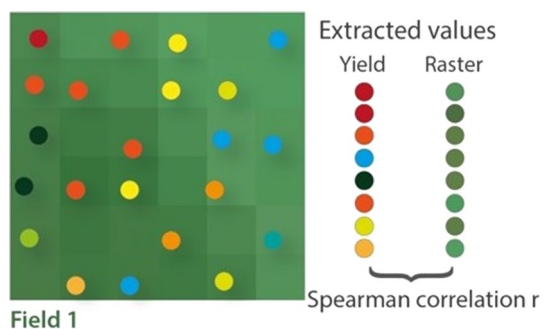
available in the study area (Harfenmeister et al., 2019). The source of the used phenological information was chosen depending on the existence and on the distance of the phenological observation to the respective field. Table 2 lists the BBCH stages and their description.

Method

To understand the relationship between remote sensing data and soil and relief data with yield data, a large number of data sets from 13 years were used for the correlation analysis. The correlation was performed on a field basis. The underlying value of satellite data and soil and relief data was extracted per yield point and the correlation per field was calculated (Fig. 3). Images covered by clouds were excluded beforehand by using a threshold for the blue band within the extent of each field. If the standard deviation of the blue band within the field extent exceeded 150, cloud coverage was likely, and the corresponding satellite data was neglected.

For the correlation analysis, 15 indices that are frequently used both in research and in precision farming practice, were selected (Lilienthal, 2014; Siegmann et al., 2012). In

Fig. 3 Extraction method of satellite data and soil and relief data (raster) per yield point per field to get data pairs for the correlation analysis



accordance with the studies of Georgi et al. (2017) and Vallentin et al. (2019), which are based on the same satellite and yield data sets, the choice of indices was narrowed down to yield-relevant spectral indices (Table 3). The indices were calculated from appropriate spectral bands of the different sensors, which are varying in number and wavelength range. Therefore, 10 of the 15 indices could be calculated with the data from all sensors, whereas five indices could only be calculated from RapidEye and/or Sentinel-2. The collection of vegetation indices does not claim exhaustiveness and could be extended in future studies.

The correlation result is given as Spearman Correlation Coefficient r (Daniel, 1990), also called “Spearman’s Rho”, and always refers to the correlation between the yield data and the respective data source to be analyzed. Earlier studies (Georgi et al., 2017; Vallentin et al., 2019) showed a monotonous but non-linear correlation between satellite data and yield data, which would rule out correlation methods based on linear correlation assumptions. One reason for the non-linearity is the saturation of vegetation indices (Esau et al., 2016; Haboudane, 2004), which in high ranges do not correlate as strongly with the yield as in the lower and middle ranges.

It will be investigated how individual data sets correlate with yield data to find out which data sets are most suitable for potential yield modelling. Because each data set relates to the vitality of the crop in a different way and at least the satellite data are not mutually dependent, a bivariate analysis was chosen. This way, the relevance of individual data sets can be highlighted and examined. The correlation values were therefore calculated as absolute values to enable a ranking.

The resulting correlation coefficients were evaluated in terms of crop, phenology, field heterogeneity, yield amount choice of data source and indices, and extraction method. To compare the results of the various correlation analyses, the median of the correlation coefficients of the analyzed group is used in the following exploration of the results. Nevertheless, outliers are also evaluated to give an assessment of which parameters correlate most with each other.

The analysis and visualisation was done by using R (R Core Team, 2020) with the use of the packages ‘raster’ (Hijmans, 2020), ‘rgdal’ (Bivand et al., 2021), ‘stringr’ (Wickham, 2019), ‘data.table’ (Dowle & Srinivasan, 2021), ‘ggplot2’ (Wickham, 2016), ‘gridExtra’ (Auguie, 2017) and ‘automap’ (Hiemstra et al., 2008).

Results and discussion

The average correlation values of all sensors, all crops, and all satellite data sets are shown in Fig. 4. Correlation coefficients show a broad spectrum: from no correlation up to correlation values of $r=0.94$. However, high correlation values above 0.75 are more outliers than standard. The median of all correlation values is 0.19 with a standard deviation of 0.16. Mild outliers appear as straight lines outside the box and extreme outliers as points above these lines. Differences between years and crops become apparent. Whereas correlation values in 2018 are comparatively high with median correlation coefficients higher than 0.25 for all crop types, years such as 2012 or 2013 show remarkably lower correlation values around 0.15. Only single crops could be analyzed in 2010 and 2011 due to missing satellite or yield data. All four crops have in general similar correlation results, whereas wheat outperforms the other crop types from 2006 to 2009.

The correlation values between soil and relief data and yield are low in most cases (mean $r=0.12$, median $r=0.09$, max $r=0.71$) (Fig. 5). The maximum r value of 0.71 is

Table 3 Overview of the calculated vegetation indices, their formula, and which sensor is available for the calculation

Index	Name	Formula	Reference	Sensors
NDVI	Normalized Difference Vegetation Index	$NDVI = \frac{NIR-Red}{NIR+Red}$	Rouse et al. (1974)	All
GNDVI	Green Normalized Difference Vegetation Index	$GNDVI = \frac{NIR-Green}{NIR+Green}$	Gitelson et al. (1996)	All
SAVI	Soil Adjusted Vegetation Index	$SAVI = \frac{NIR-Red}{(NIR+Red)+0.5} * (1 + 0.5)$	A. Huete (1988)	All
NIR	Absolute reflectance values of the NIR band	NIR	–	All
SR	Simple Ratio	NIR/Red	Jordan (1969)	All
IR/G	Ratio NIR and Green band (simple ratio)	NIR/Green	Jordan (1969)	All
EVI	Enhanced Vegetation Index	$EVI = \frac{NIR-Red}{(NIR+6*Red-7.5*Blue)+1}$	A. Huete et al. (1999)	All
CVI	Chlorophyll Vegetation Index	$CVI = NIR * \frac{Red}{Green^2}$	Vincini et al. (2008)	All
GLI	Green Leaf Index	$GLI = \frac{2*Green-Red-Blue}{2*Green+Red+Blue}$	Viña et al. (2011)	All
NDWI	Normalized Difference Water Index	$NDWI = \frac{Red-NIR}{Red+NIR}$	Gao (1996)	All
MCARI	Modified Chlorophyll Absorption in Reflectance Index	$MCARI = ((NIR - Red Edge) - 0.2 * (NIR - Green)) * (\frac{NIR}{Red Edge})$	Daughtry et al. (2000)	RapidEye & Sentinel-2
NDRE	Normalized Difference Red Edge Index	$NDRE = \frac{NIR-Red Edge}{NIR+Red Edge}$	Barnes et al. (2000)	RapidEye & Sentinel-2
NDRE / NDVI	Ratio of NDRE and NDVI	NDRE/NDVI	See above	RapidEye & Sentinel-2
CCCI	Canopy Chlorophyll Content Index	$CCCI = \frac{NIR-Red Edge}{NIR+Red Edge} * \frac{NIR-Red}{NIR+Red}$	Barnes et al. (2000)	RapidEye & Sentinel-2
REIP	Red Edge Inflection Point	$REIP = 700 + 40 * (\frac{(\frac{Red-Red Edge^2}{2}) - Red Edge^1}{Red Edge^2 - Red Edge^1})$	Guyot (1990)	Sentinel-2

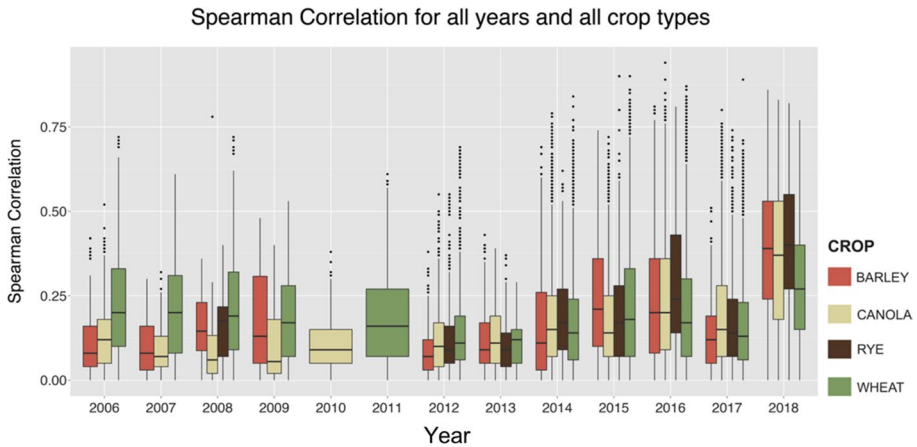


Fig. 4 Combined correlation results of all vegetation indices derived by all satellite data sets per year and per crop type. Crop types are distinguished by different boxplot colors (Color figure online)

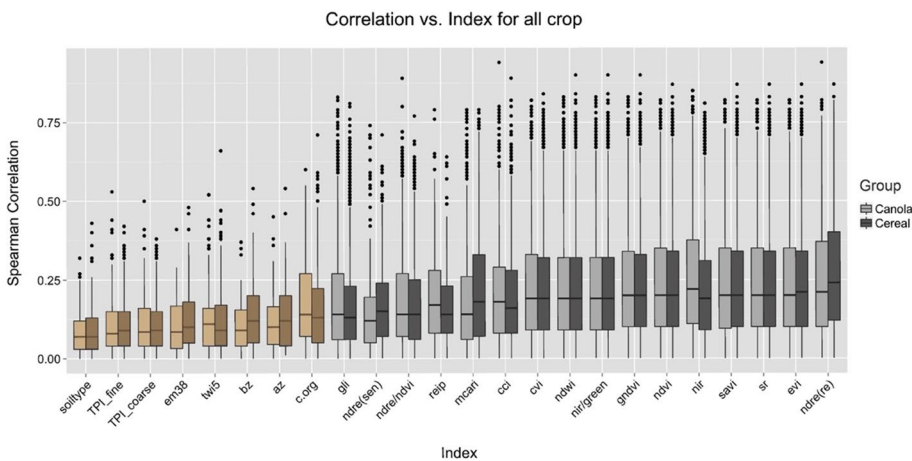


Fig. 5 Spearman correlation values for cereals (light colours) and canola (dark colours) in all years versus vegetation indices (grey colours) and soil and relief data (brown colours) (Color figure online)

attributed to the map of organic carbon, which is a derivation of satellite images itself. The performance difference between soil and relief data (brown boxplots) and satellite data (grey boxplots) can be clearly distinguished.

A reason for the low correlation between soil and relief data and yield might be that the fields in this study are mostly located in regions with flat topography and rather fertile soil compared to other regions. Consequently, the investigated soil and relief parameters often do not vary remarkably within the fields. Furthermore, Farm 1 has been operating precision agriculture for more than a decade, aiming to homogenize the fields despite the heterogeneous inventory.

Another problem concerns the comparison between the point yield data and often polygon-based soil data, which are thus much more coarsely resolved. This is probably

unfavorable as an analytical approach. In future studies, it would be more advantageous to first calculate an average yield per polygon and only then calculate the correlation.

The median of the correlation coefficients summarized for all sensors, vegetation indices and phenological stages show rather low values and exceed 0.25 only in 2018 (Fig. 4). Subsetting the data set by type of satellite sensor, year, crop type, field, vegetation index or phenology and plotting those subsets reveals patterns, which indicate the significance of each of those and other factors. They show certain tendencies, which data sets and which band combinations do best correlate with yield at which phenological stage.

In the following, the correlation results are analyzed regarding differences between sensors, vegetation indices, phenology, field and yield characteristics as well as extraction method. It must be considered that the various factors are influencing each other as well. Due to the low correlation of soil and relief data with yield data and because of their temporal immutability and mono dimensionality, only vegetation indices calculated with satellite images are considered in the following sections.

Sensor

Data from different sensors vary in spatial, spectral, and temporal resolution. All three factors might influence the result of the correlation analysis, which is explained in the following.

The general suitability of a sensor depends additionally on the temporal resolution. The more often an image is available, the higher is the probability to meet important phenological phases in which a high correlation with yield is likely.

The spatial resolution of the six sensors used in this study varies between 30 m (Landsat sensors), 20 m/10 m (Sentinel-2), 5 m (RapidEye), and 3 m (Planetscope) and has different effects on the correlation strength. The difference between sensors is exemplary shown for wheat in 2017 and 2018 (Fig. 6). In 2018, a year with above-average high median correlation values, the high-resolution sensors (RapidEye, Planetscope) perform much better than the poorer resolved sensors (Landsat series) and better than the 10 m-resolved Sentinel-2 sensor (Fig. 6b). The higher correlation coefficients of sensors with higher spatial resolution are also observable in 2017, but to a lower extent (Fig. 6a). Differences between sensors with high and low spatial resolution also appear when comparing correlation coefficients in single BBCH-stages (Fig. 7). RapidEye and Planetscope data often show higher correlation values in more BBCH-stages compared to the other sensors, which is observable for both cereals and canola.

Furthermore, the lower correlation coefficients of sensors with lower spatial resolution could also be a result of the extraction method. Yield data (points) were directly compared with the raster data (pixel) of the satellite data in the analysis. The raster of the Landsat sensors therefore includes several yield data points into one 30 m pixel, whereby the same vegetation index is collected for each yield data point. If the variance of the yield data in small areas is high, this effect can lead to uncertainty. Averaging the yield values per pixel would be possible but would also lead to data loss.

A further reason for the different performance of the sensors might be that their spectral bands (red, blue, green etc.) are covering different wavelength ranges and are consequently variously sensitive to certain wavelengths. The value and accuracy of a calculated vegetation index can differ in dependence of the covered wavelengths (Zhang et al., 2018a) and the type of atmospheric correction model that was used in the preprocessing (Doxani et al., 2018). To answer this question, a correlation analysis

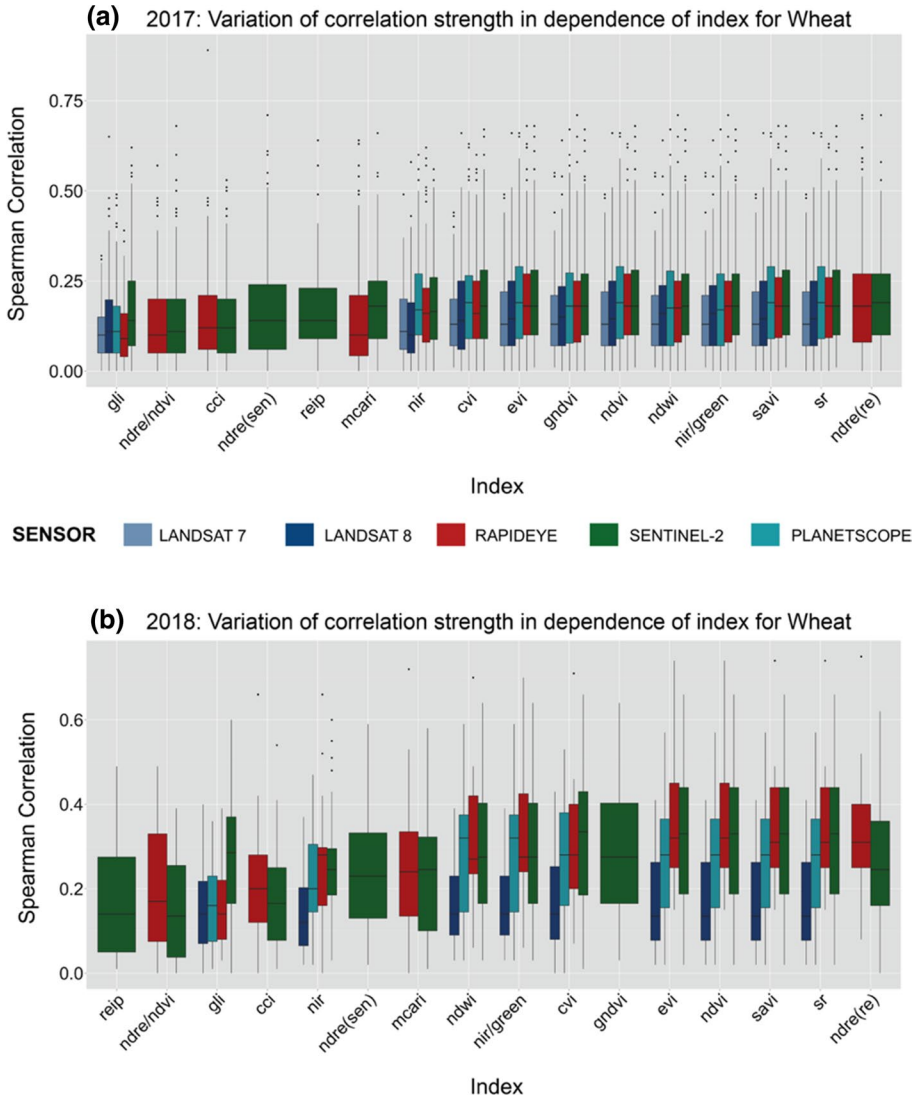


Fig. 6 Correlation per vegetation index for wheat in 2017 (a) and 2018 (b), color-distinguished by remote sensing sensor (Color figure online)

could be performed on PlanetScope and RapidEye data with different resampled products, down to 30 m pixel size. This extra step was not accomplished within this study.

A main spectral difference between sensors is obviously the existence of red edge bands and consequently the suitability to calculate certain vegetation indices, which is explained in the next section.

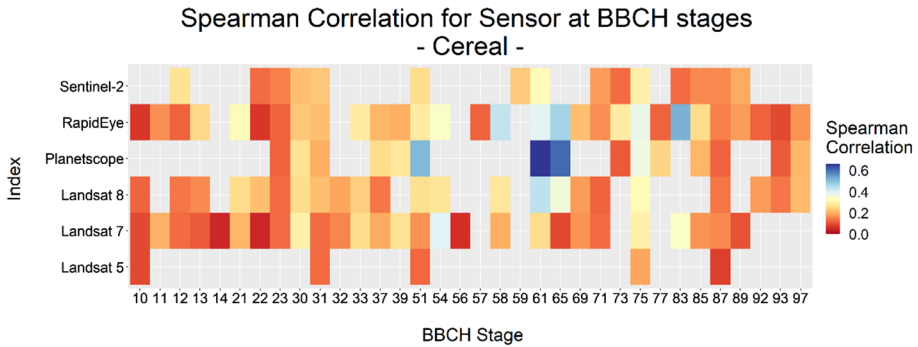


Fig. 7 Heatmap of all Spearman correlation values per sensor and per BBCH stage for cereal. Correlation values are combined for all vegetation indices

Vegetation index

The height of the correlation coefficients is certainly dependent on the calculated vegetation index. To figure out which indices are best correlating with yield data, the varying performance of the calculated indices is analyzed. In general, vegetation indices incorporating both the red and the NIR band achieve highest correlation values when the whole vegetation period is considered (Figs. 5 and 6). This includes the NDVI, the SR, EVI, SAVI, among others.

Among the vegetation indices including red edge bands, only the NDRE of RapidEye images achieves high correlation values, whereas MCARI, CCCI, NDRE/NDVI and REIP are often among the poorer performing indices (Figs. 5 and 6). On the other hand, indices including the red edge band reach higher correlation coefficients at certain phenological stages. Furthermore, if the nitrogen content or the pure biomass or leaf chlorophyll content is modelled, indices working with red edge bands show good results (Barmeier et al., 2017; Cui et al., 2019). This study cannot replicate this for the yield parameter generally, but for certain phenological stages.

Figures 5 and 6 also show a great difference between the median values of NDRE acquired by Sentinel-2 (median $r=0.16$ [Cereal], $r=0.12$ [Canola]) and by RapidEye (median $r=0.25$ [Cereal], $r=0.21$ [Canola]). This might be due to the higher spatial resolution of RapidEye (5 m) compared to Sentinel-2 (10/20 m) as well as to the larger number of available images of RapidEye (179) compared to Sentinel-2 (43). Another reason might be again the different wavelength coverage and width of the red edge and NIR bands of both sensors.

In the following, the correlation between the two indices NDVI and NDRE with cereal yield in single BBCH stages is compared to further investigate the importance of the red edge bands. The median correlation for cereal is $r=0.21$ for the NDVI and $r=0.23$ for the NDRE (taking all sensor types and all BBCH stages into account). The strength of the correlation depends on the cereal type and the respective BBCH stage and there are certainly times when the NDVI performs better (Fig. 8). In the early BBCH stages of barley, the NDVI performs better (BBCH 10, 11, 21, 30) or similarly compared to the NDRE (Fig. 8a). From BBCH 51 the NDRE achieves higher correlation values compared to NDVI. Towards the end of the vegetation phase from BBCH stage 87, both indices perform similarly. The change of the better performing index from heading to ripening might be due

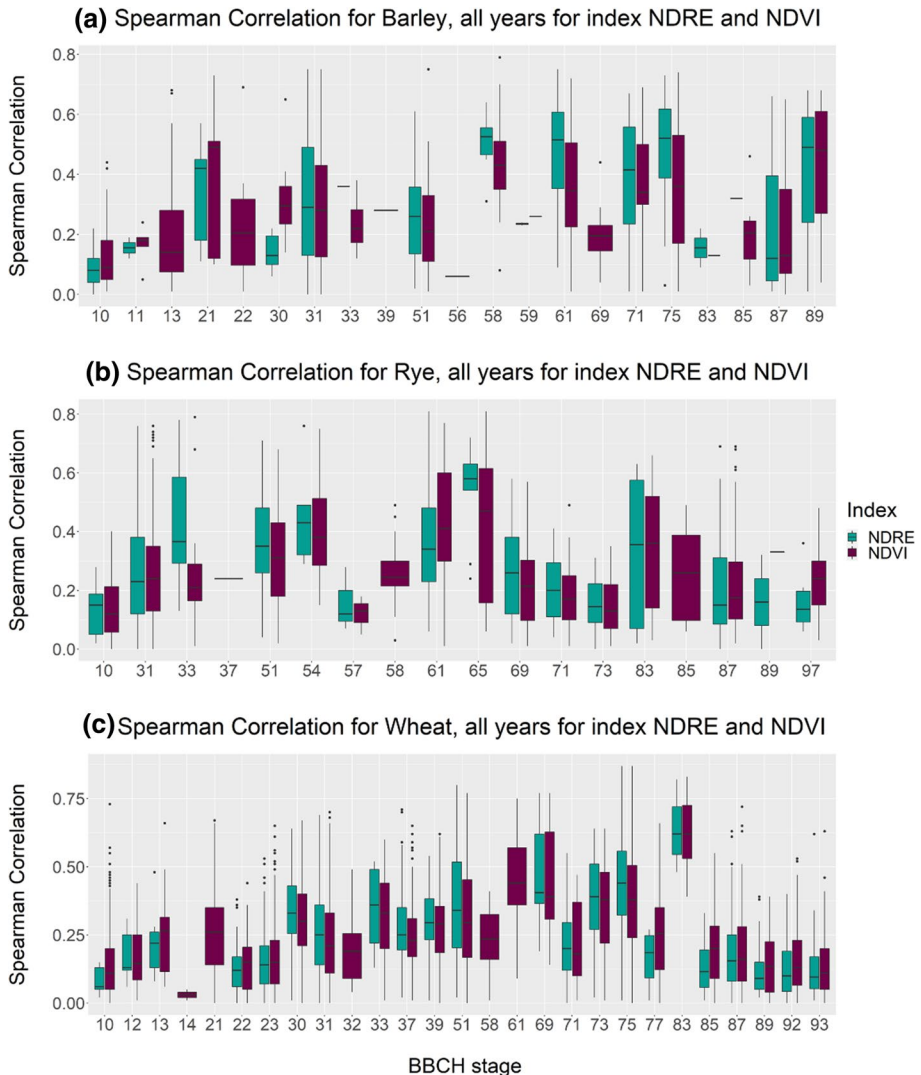


Fig. 8 Spearman correlation of NDRE (purple) and NDVI (green) per BBCH stage for barley (a), rye (b), and wheat (c) (Color figure online)

to the structure of the barley spikes, which form long awns during the vegetation period. The spikes of barley lie down with increasing weight, so that the surface of a barley field has a completely different structure than, for example, a wheat field. Depending on the phenology, this structural difference can be detected with radar satellite data (Harfenmeister et al., 2019).

In the case of rye (Fig. 8b), the NDRE almost always performs better than the NDVI, except for BBCH stages 57 and 61, and at the end of the vegetation phase from BBCH stage 87. Looking at the correlation values of wheat (Fig. 8c), the NDVI performed better than the NDRE at the beginning of the vegetation period from BBCH 10 to 23, and at the

end from BBCH 77. In the phenological phases between (BBCH stages 30–75), NDRE performs better, but differences between NDVI and NDRE are not as strong as for rye and barley.

For wheat and barley, the NDRE performs better than NDVI during BBCH stages connected with a strong increase of biomass and LAI, namely from stem elongation until fruit development. In these growth stages, the influence of the soil decreases, and plants are highly photosynthetically active. Some indices, especially the NDVI, tend to saturate at a certain LAI and are no longer sensitive in the high-density range, which might be a reason for the better performance of NDRE in the respective BBCH stages.

This example indicates that the BBCH stage might have a higher influence on the correlation results than the choice of the index, because many indices perform equally well or similarly, if the optimal BBCH stages are met.

Phenology

Not every phenological phase is suitable for yield prognosis, because the correlation values differ remarkably between different BBCH stages. For both cereals and canola, there are certain phenological phases with consistently high correlation values, mostly regardless of the chosen index (Figs. 9 and 11). Despite the reliable data basis of the phenological data, no weekly recordings of the BBCH stages were available for this study, neither cloud-free satellite images in this frequency. Therefore, this study presents the best correlation values at the times and with the data available for this purpose and does not claim to be exhaustive.

For cereals, particularly BBCH stages 54, 61, 65, and 83 stand out with correlation coefficients around 0.5 for most vegetation indices (Fig. 9). BBCH stages 58 and 75 still reach correlation values around 0.4. In contrast, most of the early BBCH stages from 10 to 23, the late BBCH stages from 85 to 97 as well as BBCH stages 56 and 57 have the lowest correlation values. Best performing BBCH stages can thus be found during flowering (61 and 65), during fruit development at the medium milk stage (75), and at the beginning of

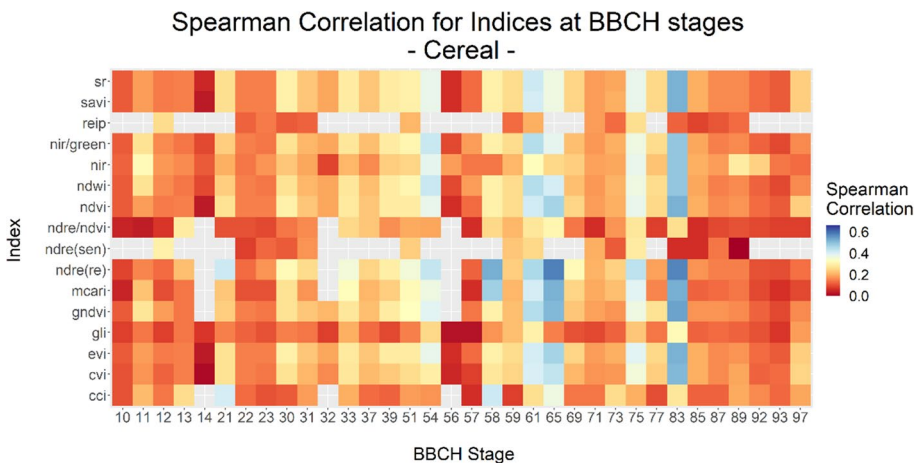


Fig. 9 Heatmap of all Spearman correlation values per vegetation index and BBCH stage for all cereal types based on all satellite sensors

ripening during the early dough stage (83). Furthermore, single BBCH stages during heading (54 and 58) show good results, but also very bad results (56 and 57). The only difference between these stages is their percentage of emerged inflorescence, therefore no causal relationship can be identified here. Consequently, the reason for the different correlation quality must be found elsewhere and is for instance caused by data availability and quality (see section “[Uncertainties](#)”). During stem elongation and until the beginning of heading, correlation values reach values around 0.3. Early BBCH stages in which no significant growth takes place (leaf development and early tillering) show low correlation with yield. Furthermore, correlation values are low when cereals are fully dense (BBCH stages 69 to 73, end of flowering to early milk) because some vegetation indices are saturating at very high biomass and LAI values. During ripening and senescence (BBCH stages 85 to 97), the correlation with final yield is also low. From this time, yield formation takes place in grains, which is not captured by satellite data. The grain yield cannot be measured directly with remote sensing. But the grain yield is closely related to the vegetative components of the cereal plant. The yield itself is produced during the grain filling phase. But the photosynthetically active plant parts such as stem, leaves and roots form the synthetic capacity to enrich the grains. It therefore makes sense that the correlation between vegetation index and final yield in the early stages of the grain filling phase shows a high correlation (Fig. 9), as already discussed by Shanahan et al., 2001 and Marti et al., 2007. The distribution of biomass, which can be indicated by vegetation indices, can also be recorded before the ripening phase, which is why correlation values between satellite image and yield are already present during stem elongation (Fig. 9). However, the uncertainty of a yield prediction depends on the further development of the plant under the given weather conditions and the management.

Differences between phenological stages are more pronounced than differences between vegetation indices. In most cases, vegetation indices perform similarly well. Surprisingly, most indices using the red edge bands (REIP, NDRE (sen), CCI) do not have as high correlation values as the other indices in the best performing BBCH stages 61 or 83. On the other hand, there are BBCH stages (e.g., 21 and 58), when indices using red edge bands perform best.

So that remote sensing sensors provide meaningful values, the crop must be present but not too homogeneous in appearance, particularly in spectral characteristics. In the context of this study, certain phenological phases for yield prediction have proven to be favorable for specific crop species, which also coincides with existing literature (Babar et al., 2006; Gaso et al., 2019; Guo et al., 2018; Peng et al., 2019).

In addition to Fig. 9, Fig. 10 shows the variance of the correlation values within the phenological stages exemplary for NDVI and NDRE. The comparison of NDVI and NDRE (Fig. 10) also shows that the NDRE performs better than or equal as NDVI when the three cereals are taken together. Nevertheless, the NDVI correlation values are also acceptably high in similar phenological phases (except 58 and 65).

Regarding canola, highest correlation values up to around 0.6 are reached in BBCH stages 71 and 77 during fruit development (Fig. 11). BBCH stages 11 (first leaf unfolded) as well as 31 to 64, which corresponds to the beginning of stem elongation to mid flowering, have correlation values around 0.3. Indices in BBCH stages 10, 13 to 16 (leaf development), 65 and 67 (mid flowering), 75 (mid fruit development), as well as 78 to 87 (end of fruit development until end of ripening) correlate worst with yield. In the two best performing BBCH stages 71 and 77, the NDRE of RapidEye, the MCARI as well as the NIR band reach highest correlation values. Furthermore, the NDRE of Sentinel-2 reaches a correlation of 0.55 during BBCH stage 61 (begin flowering), but this is not observable for the

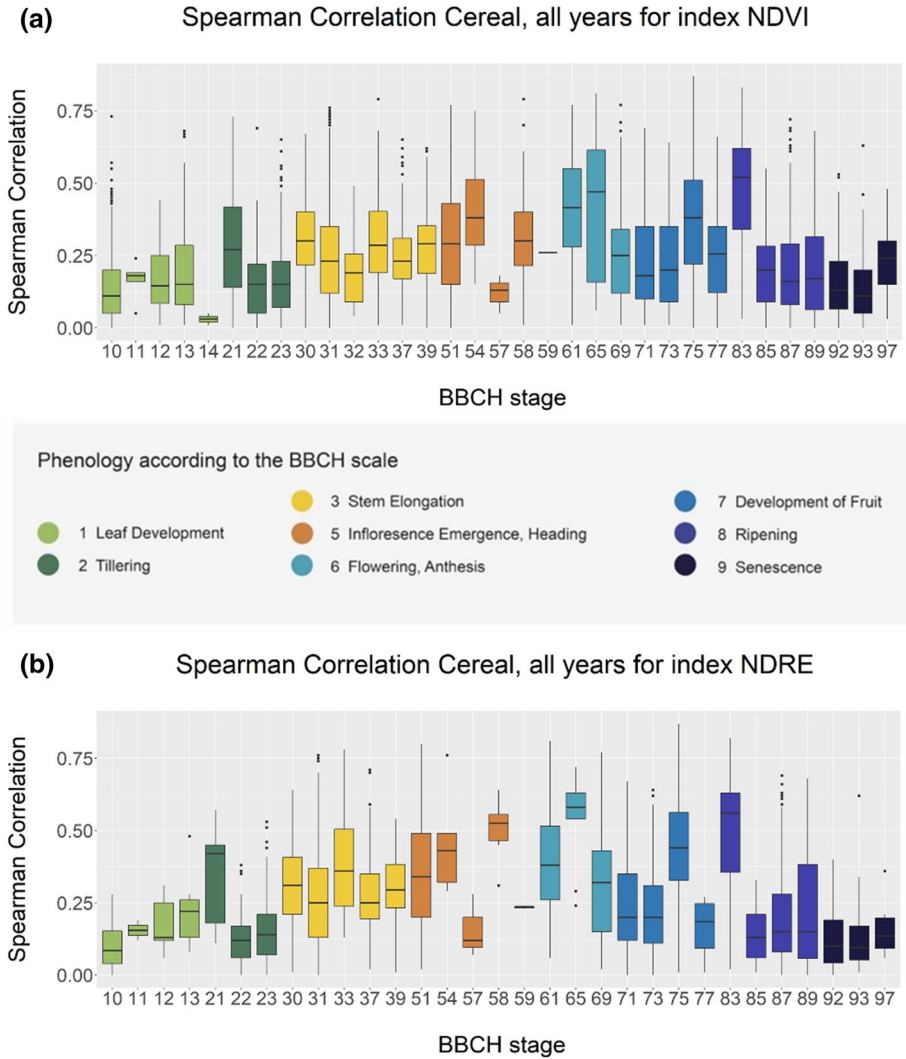


Fig. 10 Spearman correlation values for NDVI (a) and NDRE (b) per BBCH stage for all cereals and all years

remaining indices and must therefore be considered an exception. The higher correlation values in BBCH stage 11 compared to the remaining BBCH stages during leaf development might be the result of the low data availability in this stage (six fields of only one date) and can therefore not be generalized. A reason for the worse correlation values of BBCH stage 75 compared to the good performing stages 71 and 77 cannot be clearly identified. There are, for instance, a higher number of data sets for BBCH stage 75 which lead to more variation between fields.

Previous studies found a decreasing relationship between NDVI and yield with flowering (Piekarczyk et al., 2011). This can also be found in this study: The reflection of the NIR

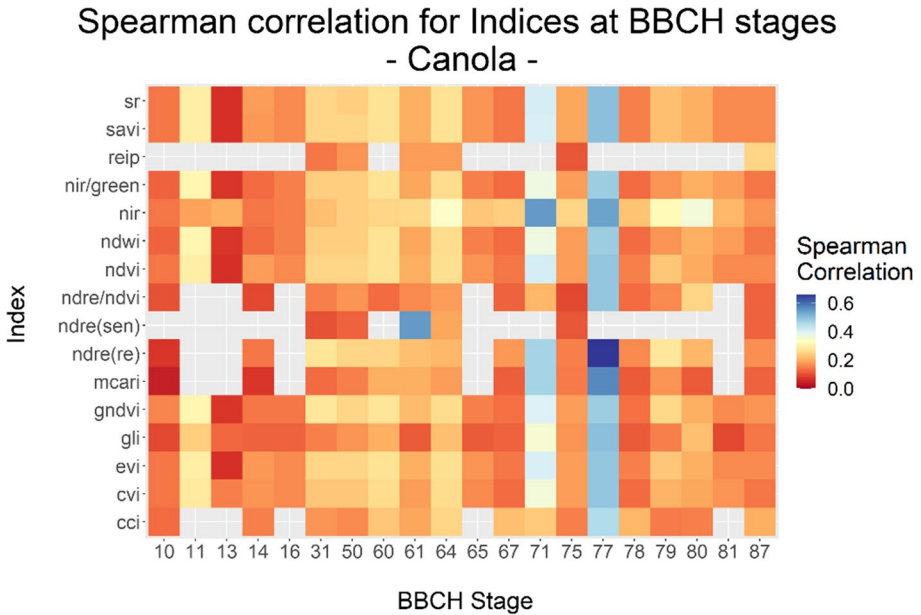


Fig. 11 Heatmap of all Spearman correlation values per vegetation index and BBCH stage for canola based on all satellite sensors

band correlates much better for BBCH stage 61 (beginning of flowering) and later, whereas the NDVI yields better results before stage 61 (Fig. 12). Holzapfel et al. (2009) found that the NDVI data obtained between the six-leaf stage (BBCH 16) and the beginning of flowering (BBCH 60) are correlated with the canola harvest. In this study, correlation results between BBCH 16 and 60 are moderate, but BBCH stages 71 and 77 perform even better.

A yield prognosis for canola is more challenging than for cereals because the vegetative parts of the plant are less significant in the formation of the grain yield (Sulik & Long, 2016). It is assumed that the number of flowers is a suitable proxy, but due to their spectral characteristics they are less sensitive to indices involving the red and infrared bands

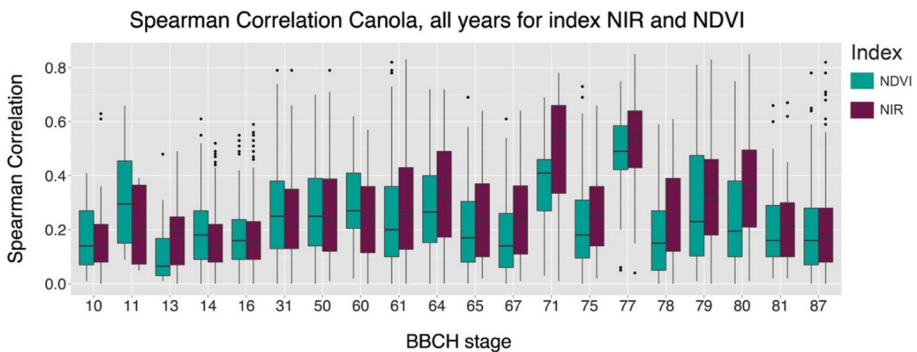


Fig. 12 Spearman correlation of NDRE (purple) and NDVI (green) per BBCH stage for canola (Color figure online)

(Sulik & Long, 2016). Remote sensing indices such as NDVI record vegetative growth, for crops such as canola there is interest in the seed that belongs to reproductive growth. Indices such as NDVI are quite useful in the analysis of crops such as wheat and maize that have inconspicuous flowers and simply "green-up" and then "green-down" after entering the reproductive growth phases. Numerous Brassica oilseeds have "green-up", then "yellow-up" with striking yellow flowers and an overlap of "yellow-down" and "green-down" during ripening (Sulik & Long, 2016). For those reasons, the authors do not feel confident to make similar recommendations for the choice of the "right" satellite images as in the case of cereals. The fact that there are far fewer studies on the correlation of spatial data to canola yield compared to cereals indicates caution in the interpretation of the results of the canola yield analysis and suggests further studies.

Field heterogeneity and density

Even if an optimal index is selected for a suitable BBCH stage, the correlation often varies remarkable between fields. Reasons for this behavior are the within-field heterogeneity as well as the density of the crop and consequently the amount of yield. Comparing the years 2015 to 2018, it is noticeably that mean and maximum correlation values between yield and NDVI achieved for 2015 and 2017 are both below those of 2016 and 2018 (Fig. 13). The availability of satellite data, particularly high-resolution data, of these years is comparable (Fig. 2). In years with high yields, the median correlation values between yield and satellite imagery are lower (median $r=0.19$) compared to years with rather low yields (median $r=0.32$).

The reason for lower correlation values in years with high yields ore fields with a high crop density is the proxy approach. Satellite sensors do not directly measure the grain yield

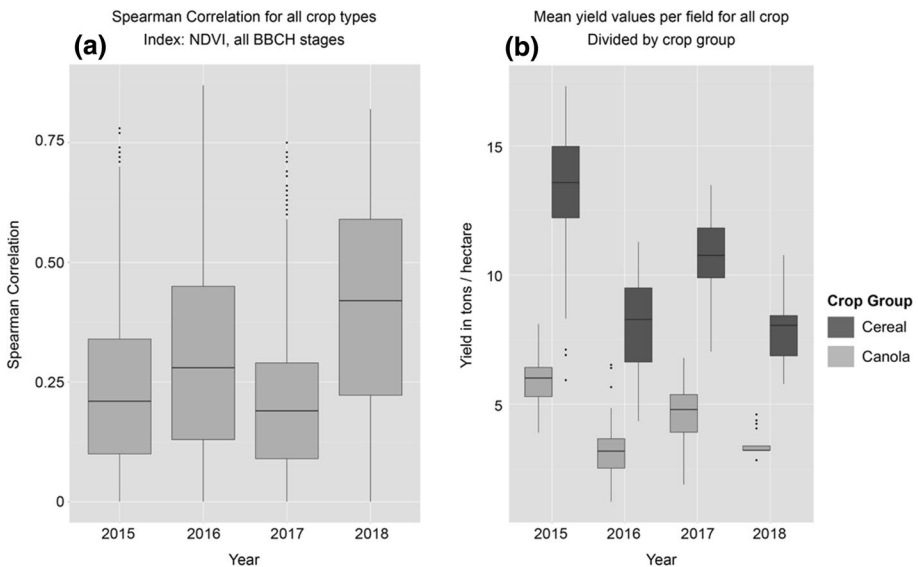


Fig. 13 Spearman correlation per year for NDVI (a). Mean yield per field in tons per hectare for years 2015–2018 (b)

as recorded during yield mapping. Instead, vegetation indices are sensitive to variables such as biomass, vitality, LAI, density and chlorophyll content, which in turn are related to yield (Babar et al., 2006; Ren et al., 2008). If the crop is vital and near the maximum of plant density and has a high LAI and a high biomass, even a sensor from space can only detect few spatial patterns. In addition, saturation occurs in most indices above a certain threshold value, so that nuances in the upper value range can no longer be detected (Haboudane et al., 2004). In contrast, fields with low and average yield perform better in the correlation analysis, because heterogeneous crop cover is more likely and a lower crop density leads to less saturation of vegetation indices at field scale (Esau et al., 2016; Haboudane, 2004). If the crop develops heterogeneously because soil or weather conditions are not in favor for growth, the probability of successfully recognizing yield patterns increases. These patterns can be seen in satellite images (Fig. 14b) and, depending on the field, in the soil map (Fig. 14d). A quantitative indicator for heterogeneity is the histogram of vegetation indices within the field boundaries. If the spread of those values is particularly narrow (Fig. 14e), heterogeneity cannot be assumed. The higher correlation values of the more heterogeneous wheat field 320-01 and the lower correlation values of the more homogeneous wheat field 300-01 are consistent between sensors (Fig. 14g).

Certainly, a maximum of homogeneity does not imply a maximum of yield in any case. For this data set it appears though, that high yields (2015, 2017) concur with more homogeneity of the fields, with more rainfall and increasing soil moisture and with lower correlation values—compared to years with dryer conditions. These findings do not apply though, if the final yield is reduced by extreme weather events or damage caused by animals and diseases, which were not seen in satellite data or only occurred after the recording date.

Extraction method

A field-internal yield analysis is not always required and mean values for a field or regional trends are sufficient. The results discussed so far are based on the correlation values calculated per field (Fig. 3). To test the influence of the extraction method on the correlation result, the correlation analysis was additionally done for the mean values of each individual field as well as for the correlation of all data sets of a crop, independently of the fields (Fig. 15). The performance of the three extraction methods turned out to be quite different.

If only the mean values per field are considered, highest correlation values are achieved (Fig. 16). However, no field-internal patterns are considered using this method. Furthermore, depending on the size of the region, the analyzed crops might be not in the same phenological stage, which makes the selection of satellite images difficult. Nonetheless, a positive relationship can be observed when comparing mean yield values with mean NDRE values, for example for cereal in BBCH stages 21, 33, 75, and 83 (Fig. 17). In general, the NDRE increases with increasing yield, but saturation effects are visible at very high index values, where a high index value covers the whole yield range.

The correlation value extraction per field as well as collectively for all fields perform similar, whereas the field-independent method performs slightly better (Fig. 17). In this case, the pure mass of yield data compared with satellite and soil and relief data could compensate for individual noise within the fields. Following these observations, yield modelling could be promising per field zone to preserve within-field heterogeneity. This is also in line with most precision agriculture methods and has been found in previous studies (Filippi et al., 2020; Mavromatis, 2016).

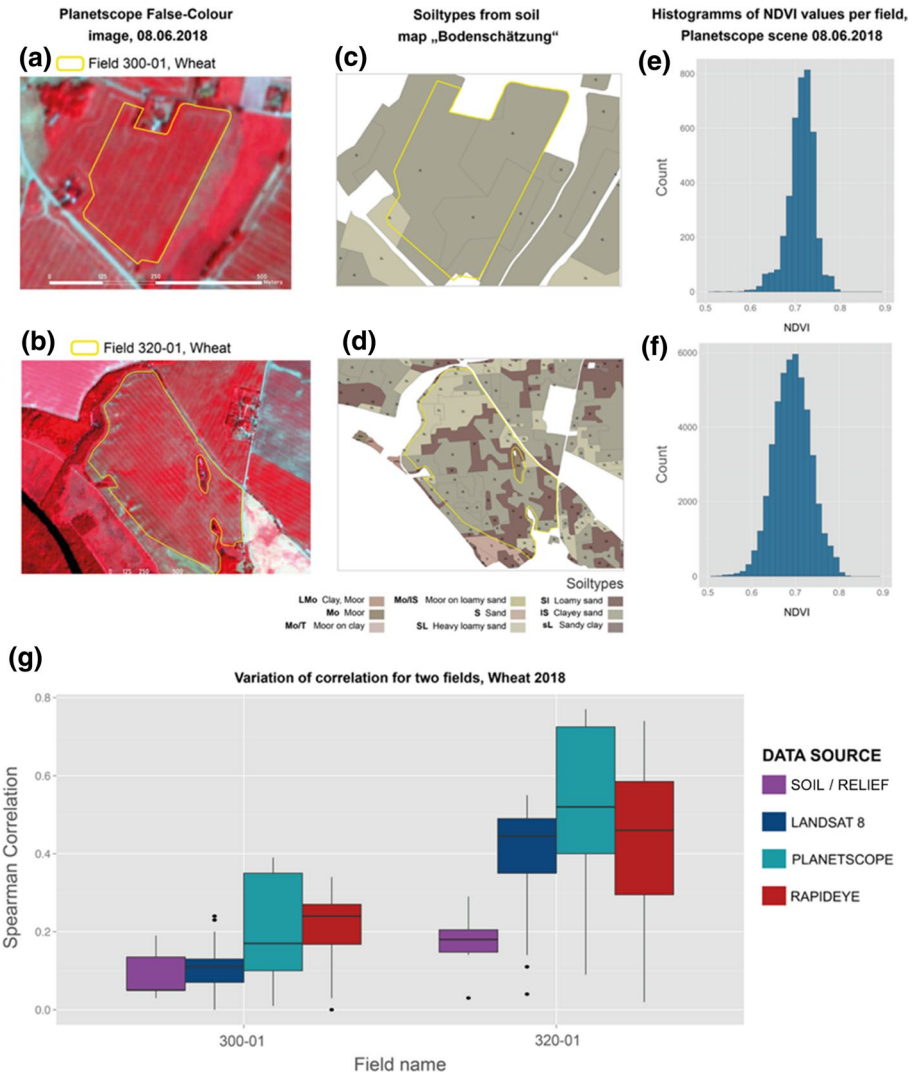


Fig. 14 Comparison of two wheat fields 300-01 (a, c, e) and 320-01(b, d, f). a and b false color image Planetscope, band combination NIR-Red-Green from 08.06.2018. c and d soil types. e and f histograms of the NDVI values of the fields at the time of recording with different spread. g variation of correlation values per field and data source (Color figure online)

Uncertainties

The analysis holds uncertainties in some questions, mainly concerning data availability and data quality and consequently the transferability to other study areas and time periods. The data basis of this study is large, but not complete. The cloud-free coverage of satellite images for every field and every crop at every phenological stage varies and some observations can only be based on thin data density, while others benefit from a large data availability. However, no correlation was found between the number of analyzed data sets and the

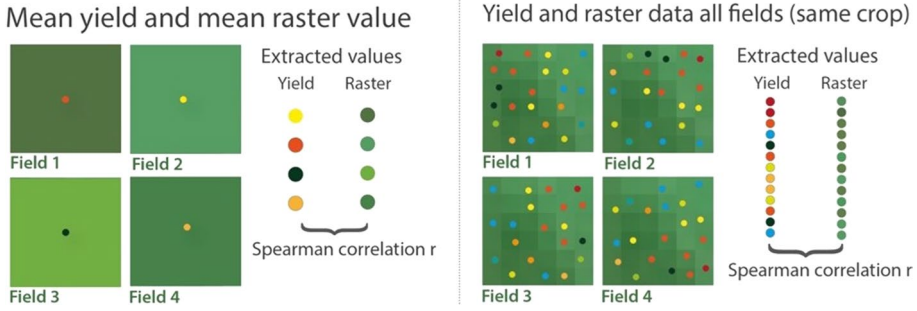


Fig. 15 Additional extraction methods of satellite data and soil and relief data (raster) per yield point per field to get data pairs for the correlation analysis for mean values per field (left) as well as field-independent for all data sets (right)

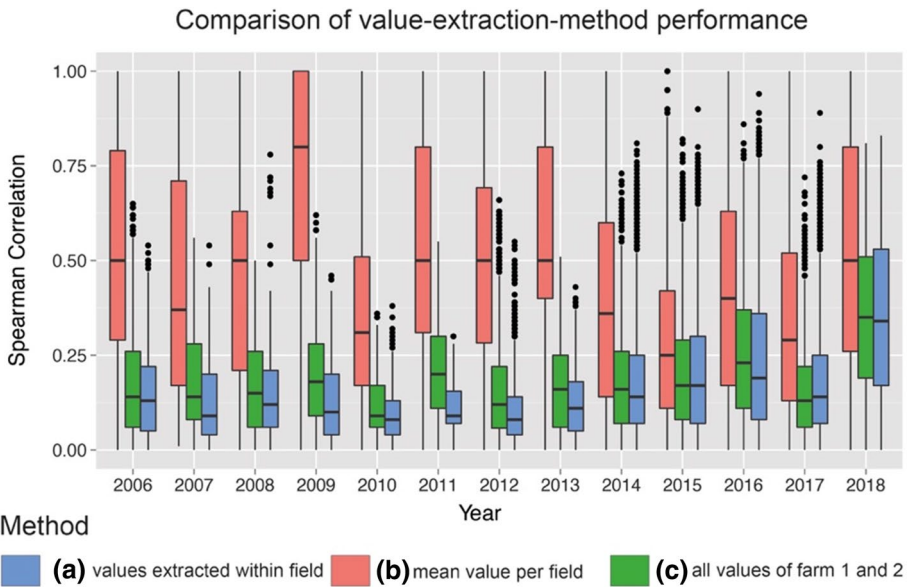


Fig. 16 Comparison of extraction methods as a basis for correlation analysis. (a) extraction of yield and data values per field, (b) extraction of mean values of each individual field, and (c) extraction of yield and data values for all points per crop and year, independent of the fields

resulting correlation coefficients between yield and vegetation index. Nevertheless, very high correlation coefficients occurred nearly exclusively in cases with a low number of available data sets. This might be due to the compensating effect of many data sets. These data sets contain correlation coefficients of a wide range and there are single fields with very high or very low correlation coefficients. Consequently, the resulting median correlations values found in this study are rather low because of the pure mass of data and their mutual balancing. In cases with only few available data sets, the chance to meet very good or very bad performing fields is higher, the resulting median values are more “extreme” and not as balanced as for a large amount of data.

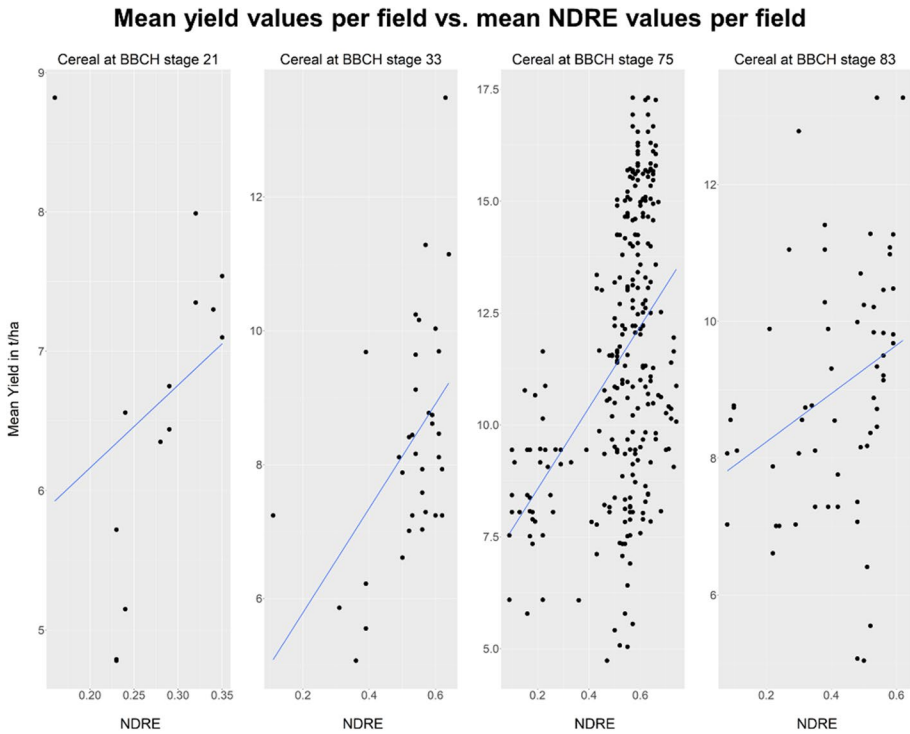


Fig. 17 Dependence of mean NDRE per field on mean yield per field for cereal in phenological BBCH phases 21 (tillering), 33 (stem elongation), 75 (medium milk), and 83 (early dough)

In this context, also the occurrence of specific BBCH stages influences the amount of analyzed data and the different data availability between BBCH stages might also explain some unexpected behavior of the correlation results (e.g., strongly differing behavior of adjacent BBCH stages such as BBCH stage 77 for cereals or 75 for canola). This might be due to the reported BBCH stages. Some BBCH stages may only last for a few days and are not necessarily captured in every year. Consequently, the data availability for these stages is significantly lower. Furthermore, some successive BBCH stages are very similar, and some phases may not have been clearly identified by reporters. There is also often a temporal and spatial shift between reported BBCH stage on a field and the acquisition of the satellite image. Consequently, a mixing between adjacent phases is likely to occur. Additionally, the DWD only reports specific BBCH stages, which are consequently overrepresented in the analysis. In future studies, the summarizing of principal growth stages might prevent confusions between single BBCH stages, but the temporal resolution would be reduced (Table 4).

Conclusions and outlook

This study examined the relation between remote sensing data, soil and relief data and agronomic crop yield data. Remote sensing in agriculture brings the advantage of wide spatial information, which is up to date, but also capable of analyzing the past. It has

Table 4 Short summary of the results and indication of which vegetation index with which sensor has the highest correlation values in which phenological phase

Crop	Top indices	Best BBCH	Best sensor	
Rye	NDRE (RapidEye), NDVI, EVI, SAVI, SR	Early: 33 Late: 65, 61, 83	RapidEye, Planet	Figure 8b
Wheat	NDRE (RapidEye), NDVI, EVI, SAVI, SR, GNDVI	Early: 31 Late: 83, 69, 75, 73	RapidEye, Planet, Sentinel-2	Figures 6, 8c
Barley	NDVI, EVI, SAVI, SR, IR / G, NDWI	Early: 21, 31 Late: 75, 89, 61, 71	RapidEye, Planet, Sentinel-2	Figure 8a
Canola	NIR, NDRE (RapidEye), NDVI, EVI, SAVI, SR, GNDVI	Early: 11 Late: 77, 71	RapidEye Planet, Sentinel-2	Figures 11, 12

shown that not every satellite image equally well correlates with yield data and is consequently suitable for yield estimation. The hypothesis that an optimal combination of sensor, vegetation index and time period exists is confirmed by the analyses of this study and could be extended by further conditions, such as field heterogeneity, that have to be met for a successful yield estimation. A careful selection is necessary depending on the type of crop, the phenology of the crop, the resolution of the sensor and the spectral bands considered in the index calculation.

The relevance of soil and relief data for yield modelling is considered negligible in this study. Since soil and relief data play a minor role in this study compared to satellite data for yield prediction, as they are almost always underperforming, it is important to recognize that up-to-date data sets are needed to map the actual condition of the plants. However, the crop growth system is complex, and soil and relief certainly influence yield. Therefore, it makes sense to use these data in addition to the satellite images for an extended yield prognosis using multiple predictors, possibly with a lower weighting, as they usually do not show the actual condition of the crop and are changing only in space, not in time. Soil and relief data are certainly relevant for specific fields, but because this requires local knowledge, this study cannot give general recommendations.

Satellite images with higher spatial resolution such as RapidEye (5 m) or PlanetScope (3 m) performed better than satellites with a lower spatial resolution such as Landsat (30 m) and even Sentinel-2 (10/20 m). RapidEye and Sentinel-2 have an advantage because of their spectral coverage in the red edge region. Working in Germany and in countries with similar average field size, the use of optical sensors with the spatial and spectral resolution of Sentinel-2, RapidEye and PlanetScope satellites is recommended. Particularly the vegetation index NDRE achieved high correlation values for cereal between stem elongation and fruit development. Vegetation indices covering the NIR and red bands such as NDVI, SR, or EVI perform similarly well as long as suitable phenological growth stages are present. Therefore, phenology plays a more important role than the chosen vegetation index. The acquisition date and therefore the timing in the crops' development is the most important of these factors. For cereals, particularly BBCH stages 54 (middle of heading), 61, 65 (flowering) and 83 (early dough) show highest correlations with yield. Consequently, the crop should be abundant, but not at the maximum of green biomass and density, also not completely ripe. For canola, BBCH stages 71 and 77 (fruit development) showed highest correlation values. When looking at canola in particular, further studies are useful because the correlation values are not yet quite clear in the literature. An analysis of the same data in other regions and natural habitats would provide a valuable basis for the use of satellite data in agriculture.

Additionally, the field heterogeneity influences the correlation between satellite image and yield. Fields with lower yields and consequently lower crop density often resulting in higher field heterogeneity showed higher correlation coefficients. If no patterns are recognizable and the crop is homogeneous, the uncertainty in the yield estimation within the field is high.

Extracting mean field values instead of taking all pixels of a field into account resulted in higher correlation, but on the other hand information about within-field heterogeneity for applying precision agriculture is lost. Therefore, future studies should evaluate yield modelling per field zone to preserve within-field heterogeneity. Furthermore, the analysis of principal growth stages instead of single BBCH stages prevents might be promising to account for differences in data availability and data quality.

Acknowledgements The research is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme, grant number 28DE114A18. The authors would like to thank Climate-KIC for project funding (Fellows Programme 13-1.1.1., Task ID ARED0004_2013-1.1-011_P024-09), Edgar Zabel and the cooperating farms for data and support. We thank the German Aerospace Centre (DLR) for providing the data from the RapidEye Science Archive (RESA 617 FKZ) and the weather stations, the DWD for providing the soilwater model METVER and Planet Labs Inc. for providing Planetscope data for student research purposes. Landsat Data Available from the U.S. Geological Survey.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- AG Boden. (2005). *Bodenkundliche Kartieranleitung*. Bundesanstalt für Geowissenschaften und Rohstoffe und den Geologischen Landesämtern in der Bundesrepublik Deutschland Hannover. Bundesanstalt für Geowissenschaften und Rohstoffe in Zusammenarbeit mit den Staatlichen Geologischen Diensten.
- Ali, A., Martelli, R., Lupia, F., & Barbanti, L. (2019). Assessing multiple years' spatial variability of crop yields using satellite vegetation indices. *Remote Sensing*. <https://doi.org/10.3390/rs11202384>
- Amt für Geoinformation Vermessungs- und Katasterwesen. (2011). DGM 5 - Digitales Geländemodell Gitterweite 5m - Mecklenburg-Vorpommern. Schwerin.
- Auguie, B. (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. <https://cran.r-project.org/package=gridExtra>
- Babar, M. A., van Ginkel, M., Klatt, A. R., Prasad, B., & Reynolds, M. P. (2006). The potential of using spectral reflectance indices to estimate yield in wheat grown under reduced irrigation. *Euphytica*, 150(1–2), 155–172. <https://doi.org/10.1007/s10681-006-9104-9>
- Barmeier, G., Hofer, K., & Schmidhalter, U. (2017). Mid-season prediction of grain yield and protein content of spring barley cultivars using high-throughput spectral sensing. *European Journal of Agronomy*, 90, 108–116. <https://doi.org/10.1016/J.EJA.2017.07.005>
- Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., & Thompson, T. (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 and other resource management*. Bloomington, Madison, USA.
- Baruth, B., Royer, A., Klisch, A., & Genovese, G. (2008). The use of remote sensing within the MARS crop yield monitoring system of the European Commission. *Proceedings ISPRS*, 27, 935–940.
- Battude, M., Al Bitar, A., Morin, D., Cros, J., Huc, M., Marais Sicre, C., Le Dantec, V., & Demarez, V. (2016). Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sensing of Environment*, 184, 668–681. <https://doi.org/10.1016/j.rse.2016.07.030>
- BBC Working Group. (2001). *Growth stages of mono- and dicotyledonous plants*. (U. Meier, Ed.).
- Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24(1), 43–69. <https://doi.org/10.1080/02626667909491834>
- Bivand, R., Keitt, T., & Rowlingson, B. (2021). rgdal: Bindings for the "Geospatial" Data Abstraction Library. R package version 1.5–23. <https://cran.r-project.org/package=rgdal>
- Blasch, G., Spengler, D., Hohmann, C., Neumann, C., Itzerott, S., & Kaufmann, H. (2015a). Multitemporal soil pattern analysis with multispectral remote sensing data at the field-scale. *Computers and Electronics in Agriculture*, 113(April), 1–13. <https://doi.org/10.1016/j.compag.2015.01.012>

- Blasch, G., Spengler, D., Itzerott, S., & Wessolek, G. (2015b). Organic matter modeling at the landscape scale based on multitemporal soil pattern analysis using rapideye data. *Remote Sensing*, 7(9), 11125–11150. <https://doi.org/10.3390/rs70911125>
- Bognár, P., Kern, A., Pásztor, S., Lichtenberger, J., Koronczay, D., & Ferencz, C. (2017). Yield estimation and forecasting for winter wheat in Hungary using time series of MODIS data. *International Journal of Remote Sensing*, 38(11), 3394–3414. <https://doi.org/10.1080/01431161.2017.1295482>
- Campos, I., González-Gómez, L., Villodre, J., Calera, M., Campoy, J., Jiménez, N., Plaza, C., Sánchez-Prieto, S., & Calera, A. (2019). Mapping within-field variability in wheat yield and biomass using remote sensing vegetation indices. *Precision Agriculture*, 20(2), 214–236. <https://doi.org/10.1007/s11119-018-9596-z>
- Clevers, J., Kooistra, L., & van den Brande, M. (2017). Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. *Remote Sensing*, 9(5), 405. <https://doi.org/10.3390/rs9050405>
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, 8(7), 1991–2007. <https://doi.org/10.5194/gmd-8-1991-2015>
- Corwin, D., & Plant, R. (2005). Applications of apparent soil electrical conductivity in precision agriculture. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2004.10.004>
- Cui, B., Zhao, Q., Huang, W., Song, X., Ye, H., & Zhou, X. (2019). Leaf chlorophyll content retrieval of wheat by simulated RapidEye, Sentinel-2 and EnMAP data. *Journal of Integrative Agriculture*, 18(6), 1230–1245. [https://doi.org/10.1016/S2095-3119\(18\)62093-3](https://doi.org/10.1016/S2095-3119(18)62093-3)
- Daniel, W. W. (1990). *Applied nonparametric statistics*. PWS-KENT Pub.
- Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B., & McMurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, 74(2), 229–239. [https://doi.org/10.1016/S0034-4257\(00\)00113-9](https://doi.org/10.1016/S0034-4257(00)00113-9)
- Deutscher Wetterdienst (DWD). (2020). Vieljährige Mittelwerte. Retrieved 9 December, 2020, from https://www.dwd.de/DE/leistungen/klimadatendeutschland/vielj_mittelwerte.html
- Dobers, E. S., Ahl, C., & Stuczynski, T. (2010). Comparison of Polish and German maps of agricultural soil quality using GIS. *Journal of Plant Nutrition and Soil Science*, 173(2), 185–197. <https://doi.org/10.1002/jpln.200800317>
- Dowle, M., & Srinivasan, A. (2021). data.table: Extension of 'data.frame'. R package version 1.14.0. <https://cran.r-project.org/package=data.table>
- Doxani, G., Vermote, E., Roger, J.-C., Gascon, F., Adriaensen, S., Frantz, D., Hagolle, O., Hollstein, A., Kirches, G., Li, F., Louis, J., Mangin, A., Pahlevan, N., Pflug, B., & Vanhellemont, Q. (2018). Atmospheric correction inter-comparison exercise. *Remote Sensing*. <https://doi.org/10.3390/rs10020352>
- Esau, I., Miles, V. V., Davy, R., Miles, M. W., & Kurchatova, A. (2016). Trends in normalized difference vegetation index (NDVI) associated with urban development in northern West Siberia. *Atmospheric Chemistry and Physics*, 16(15), 9563–9577. <https://doi.org/10.5194/acp-16-9563-2016>
- Evans, L. T., & Fischer, R. A. (1999). Yield potential: Its definition, measurement, and significance. *Crop Science*, 39(6), 1544–1551. <https://doi.org/10.2135/cropsci1999.3961544x>
- Filippi, P., Whelan, B. M., Vervoort, R. W., & Bishop, T. F. A. (2020). Mid-season empirical cotton yield forecasts at fine resolutions using large yield mapping datasets and diverse spatial covariates. *Agricultural Systems*, 184, 102894. <https://doi.org/10.1016/j.agsy.2020.102894>
- Finch, H. J. S., Samuel, A. M., & Lane, G. P. F. (2014). 10—Precision farming. In H. J. S. Finch, A. M. Samuel, & G. P. F. Lane (Eds.), *Lockhart & Wiseman's Crop Husbandry Including Grassland (Ninth Edition)* (Ninth Edit., pp. 235–244). Woodhead Publishing. <https://doi.org/10.1533/9781782423928.2.235>
- Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)
- Gasó, D. V., Berger, A. G., & Ciganda, V. S. (2019). Predicting wheat grain yield and spatial variability at field scale using a simple regression or a crop model in conjunction with Landsat images. *Computers and Electronics in Agriculture*, 159, 75–83. <https://doi.org/10.1016/j.compag.2019.02.026>
- Geisler, G. (1988). *Pflanzenbau*. Paul Parey.
- Georgi, C., Spengler, D., Itzerott, S., & Kleinschmit, B. (2017). Automatic delineation algorithm for site-specific management zones based on satellite remote sensing data. *Precision Agriculture*. <https://doi.org/10.1007/s11119-017-9549-y>
- 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 Sensing of Environment*, 58(3), 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7)

- Gómez, J., Kumhálová, J., & Novák, P. (2017). Assessment of the relationship between spectral indices from satellite remote sensing and winter oilseed rape yield. *Agronomy Research*, 15, 55–68.
- Guo, C., Zhang, L., Zhou, X., Zhu, Y., Cao, W., Qiu, X., Cheng, T., & Tian, Y. (2018). Integrating remote sensing information with crop model to monitor wheat growth and yield based on simulation zone partitioning. *Precision Agriculture*, 19(1), 55–78. <https://doi.org/10.1007/s11119-017-9498-5>
- Guyot, G. (1990). Optical properties of vegetation canopies. In M. D. Stevens & J. A. Clark (Eds.), *Applications of remote sensing in agriculture* (pp. 19–43). Butterworths.
- Haboudane, D. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3), 337–352. <https://doi.org/10.1016/j.rse.2003.12.013>
- Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3), 337–352. <https://doi.org/10.1016/j.rse.2003.12.013>
- Hack, H., Bleiholder, H., Buhr, L., Meier, U., Schnock-Fricke, U., Weber, E., & Witzemberger, A. (1992). A uniform code for phenological growth stages of mono- and dicotyledonous plants—Extended BBCH scale, general -. *Nachrichtenblatt Deutscher Pflanzenschutzdienst*, 44(12), 265–270.
- Harfenmeister, K., Spengler, D., & Weltzien, C. (2019). Analyzing temporal and spatial characteristics of crop parameters using Sentinel-1 backscatter data. *Remote Sensing*, 11(13), 1569. <https://doi.org/10.3390/rs11131569>
- Heinrich, I., Balanzategui, D., Bens, O., Blasch, G., Blume, T., Böttcher, F., Borg, E., Brademann, B., Brauer, A., Conrad, C., & Dietze, E. (2018). Interdisciplinary geo-ecological research across time scales in the Northeast German Lowland Observatory (TERENO-NE). *Vadose Zone Journal*. <https://doi.org/10.2136/vzj2018.06.0116>
- Hiemstra, P. H., Pebesma, E. J., Twenhöfel, C. J. W., & Heuvelink, G. B. M. (2008). Real-time automatic interpolation of ambient gamma dose rates from the Dutch Radioactivity Monitoring Network. *Computers & Geosciences*.
- Hijmans, R. J. (2020). raster: Geographic data analysis and modeling. R package version 3.3-13. <https://cran.r-project.org/package=raster>
- Hollstein, A., Segl, K., Guanter, L., Brell, M., & Enesco, M. (2016). Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in Sentinel-2 MSI images. *Remote Sensing*. <https://doi.org/10.3390/rs8080666>
- Holzzapfel, C. B., Lafond, G. P., Brandt, S. A., Bullock, P. R., Irvine, R. B., James, D. C., et al. (2009). Optical sensors have potential for determining nitrogen fertilizer topdressing requirements of canola in Saskatchewan. *Canadian Journal of Plant Science*, 89(2), 411–425. <https://doi.org/10.4141/CJPS08127>
- Huete, A. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- Huete, A., Justice, C., & van Leeuwen, W. (1999). *MODIS Vegetation Index (MOD 13): Algorithm theoretical basis document*. https://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf
- Hunt, M. L., Blackburn, G. A., Carrasco, L., Redhead, J. W., & Rowland, C. S. (2019). High resolution wheat yield mapping using Sentinel-2. *Remote Sensing of Environment*, 233, 111410. <https://doi.org/10.1016/j.RSE.2019.111410>
- Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50(4), 663–666.
- Knoblauch, C., Watson, C., Berendonk, C., Becker, R., Wrage-Mönnig, N., & Wichern, F. (2017). Relationship between remote sensing data, plant biomass and soil nitrogen dynamics in intensively managed grasslands under controlled conditions. *Sensors*, 17(7), 1483. <https://doi.org/10.3390/s17071483>
- Kühn, J., Brenning, A., Wehrhan, M., Koszinski, S., & Sommer, M. (2008). Interpretation of electrical conductivity patterns by soil properties and geological maps for precision agriculture. *Precision Agriculture*, 10(6), 490–507. <https://doi.org/10.1007/s11119-008-9103-z>
- Lilienthal, H. (2014). Optische Sensoren in der Landwirtschaft: Grundlagen und Konzepte. *Journal Für Kulturpflanzen*, 66(2), 34–41. <https://doi.org/10.5073/jfk.2014.02.01>
- Macdonald, J. M., Korb, P., & Hoppe, R. A. (2013). *Farm size and the organization of U.S. Crop Farming*. *Economic research report* (Vol. 152).
- Marti, J., Bort, J., Slafer, G. A., & Araus, J. L. (2007). Can wheat yield be assessed by early measurements of Normalized Difference Vegetation Index? *Annals of Applied Biology*, 150(2), 253–257. <https://doi.org/10.1111/j.1744-7348.2007.00126.x>

- Mavromatis, T. (2016). Spatial resolution effects on crop yield forecasts: An application to rainfed wheat yield in north Greece with CERES-Wheat. *Agricultural Systems*, 143, 38–48. <https://doi.org/10.1016/j.agsy.2015.12.002>
- Moral, F. J., Terrón, J. M., & da Silva, J. R. M. (2010). Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. *Soil and Tillage Research*, 106(2), 335–343. <https://doi.org/10.1016/j.still.2009.12.002>
- Panek, E., Gozdowski, D., Stępień, M., Samborski, S., Ruciński, D., & Buszke, B. (2020). Within-field relationships between satellite-derived vegetation indices, grain yield and spike number of winter wheat and triticale. *Agronomy*. <https://doi.org/10.3390/agronomy10111842>
- Pasqualotto, N., Delegido, J., Van Wittenberghe, S., Rinaldi, M., & Moreno, J. (2019). Multi-crop green LAI estimation with a new simple Sentinel-2 LAI Index (SeLI). *Sensors*. <https://doi.org/10.3390/s19040904>
- Peng, Y., Zhu, T., Li, Y., Dai, C., Fang, S., Gong, Y., Wu, X., Zhu, R., & Liu, K. (2019). Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizer applications. *Agricultural and Forest Meteorology*, 271, 116–125. <https://doi.org/10.1016/j.agrformet.2019.02.032>
- Piekarczyk, J., Sulewska, H., & Szymańska, G. (2011). Winter oilseed-rape yield estimates from hyperspectral radiometer measurements. *Quaestiones Geographicae*, 30(1), 77–84.
- Prey, L., Hu, Y., & Schmidhalter, U. (2020). High-throughput field phenotyping traits of grain yield formation and nitrogen use efficiency: Optimizing the selection of vegetation indices and growth stages. *Frontiers in Plant Science*, 10, 1672. <https://doi.org/10.3389/fpls.2019.01672>
- Punalekar, S. M., Verhoef, A., Quaife, T. L., Humphries, D., Birmingham, L., & Reynolds, C. K. (2018). Application of Sentinel-2A data for pasture biomass monitoring using a physically based radiative transfer model. *Remote Sensing of Environment*, 218, 207–220. <https://doi.org/10.1016/j.RSE.2018.09.028>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.r-project.org/>
- Ren, J., Chen, Z., Zhou, Q., & Tang, H. (2008). Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. *International Journal of Applied Earth Observation and Geoinformation*, 10(4), 403–413. <https://doi.org/10.1016/J.JAG.2007.11.003>
- Ren, J., Li, S., Chen, Z., Zhou, Q., & Tang, H. (2007). Regional yield prediction for winter wheat based on crop biomass estimation using multi-source data. In *IEEE international geoscience and remote sensing symposium, 1–12* (Sensing and understanding our planet) (pp. 805–808).
- Richter, R., & Schläpfer, D. (2005). Atmospheric/topographic correction for satellite imagery. *DLR Report DLR-IB*, 438, 501–565.
- Rouse, J. W. J., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA. Goddard Space Flight Center 3d ERTS-1 Symp., 1*, 309–317.
- Scheffler, D., Hollstein, A., Diedrich, H., Segl, K., & Hostert, P. (2017). AROSICS: An automated and robust open-source image co-registration software for multi-sensor satellite data. *Remote Sensing*. <https://doi.org/10.3390/rs9070676>
- Serrano, L., Filella, I., & Pen, J. (2000). Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Science*, 40, 723–731.
- Shanahan, J. F., Schepers, J. S., Francis, D. D., Varvel, G. E., Wilhelm, W. W., Tringe, J. M., et al. (2001). Use of remote-sensing imagery to estimate corn grain yield. *Agronomy Journal*, 93(3), 583. <https://doi.org/10.2134/agronj2001.933583x>
- Siegmann, B., Jarmer, T., Lilienthal, H., Richter, N., Selige, T., & Sensing, R. (2012). Comparison of narrow band vegetation indices and empirical models from hyperspectral remote sensing data for the assessment of wheat nitrogen concentration.
- Spengler, D., Förster, M., & Borg, E. (2018). Editorial. *PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science*. <https://doi.org/10.1007/s41064-018-0052-5>
- Statistische Ämter des Bundes und der Länder. (2011). *Agrarstrukturen in Deutschland Einheit in Vielfalt: Regionale Ergebnisse der Landwirtschaftszählung 2010*. <https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Landwirtschaft-Forstwirtschaft-Fischerei/Landwirtschaftliche-Betriebe/Publikationen/Downloads-Landwirtschaftliche-Betriebe/agrarstrukturen-in-deutschland-5411203109005.html>
- Sudduth, K. A., & Drummond, S. T. (2007). Yield editor: Software for removing errors from crop yield maps. *Agronomy Journal*, 99(6), 1471. <https://doi.org/10.2134/agronj2006.0326>
- Sulik, J. J., & Long, D. S. (2016). Spectral considerations for modeling yield of canola. *Remote Sensing of Environment*, 184, 161–174. <https://doi.org/10.1016/J.RSE.2016.06.016>

- Thenkabail, P. S., Smith, R. B., & De Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, *71*(2), 158–182. [https://doi.org/10.1016/S0034-4257\(99\)00067-X](https://doi.org/10.1016/S0034-4257(99)00067-X)
- Vallentin, C., Dobers, E. S., Itzerott, S., Kleinschmit, B., & Spengler, D. (2019). Delineation of management zones with spatial data fusion and belief theory. *Precision Agriculture*. <https://doi.org/10.1007/s11119-019-09696-0>
- Viña, A., Gitelson, A. A., Nguy-Robertson, A. L., & Peng, Y. (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment*, *115*(12), 3468–3478. <https://doi.org/10.1016/J.RSE.2011.08.010>
- Vincini, M., Frazzi, E., & D'Alessio, P. (2008). A broad-band leaf chlorophyll vegetation index at the canopy scale. *Precision Agriculture*, *9*(5), 303–319. <https://doi.org/10.1007/s11119-008-9075-z>
- Weiss, A. (2001). Topographic position and landforms analysis. Retrieved 7 January, 2019, from http://www.jennessent.com/downloads/tpi-poster-tnc_18x22.pdf
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. <https://ggplot2.tidyverse.org>
- Wickham, H. (2019). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0. <https://cran.r-project.org/package=stringr>
- Zhang, H. K., Roy, D. P., Yan, L., Li, Z., Huang, H., Vermote, E., Skakun, S., & Roger, J.-C. (2018a). Characterization of Sentinel-2A and Landsat-8 top of atmosphere, surface, and nadir BRDF adjusted reflectance and NDVI differences. *Remote Sensing of Environment*, *215*, 482–494. <https://doi.org/10.1016/j.rse.2018.04.031>
- Zhang, Y., Qin, Q., Ren, H., Sun, Y., Li, M., Zhang, T., & Ren, S. (2018b). Optimal hyperspectral characteristics determination for winter wheat yield prediction. *Remote Sensing*. <https://doi.org/10.3390/rs10122015>
- Zhao, Y., Potgieter, A. B., Zhang, M., Wu, B., & Hammer, G. L. (2020). Predicting wheat yield at the field scale by combining high-resolution sentinel-2 satellite imagery and crop modelling. *Remote Sensing*. <https://doi.org/10.3390/rs12061024>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.