University of Reading

# *In-field detection and quantification of Septoria tritici blotch in diverse wheat germplasm using spectral-temporal features*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Anderegg, J., Hund, A., Karisto, P. and Mikaberidze, A. (2019) In-field detection and quantification of Septoria tritici blotch in diverse wheat germplasm using spectral-temporal features. Frontiers in Plant Science, 10. 1355. ISSN 1664-462X doi: 10.3389/fpls.2019.01355 Available at https://centaur.reading.ac.uk/86585/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01355/abstract
To link to this article DOI: http://dx.doi.org/10.3389/fpls.2019.01355

Publisher: Frontiers

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

University of Reading

www.reading.ac.uk/centaur

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 October 2019
doi: 10.3389/fpls.2019.01355

# In-Field Detection and Quantification of *Septoria Tritici* Blotch in Diverse Wheat Germplasm Using Spectral–Temporal Features

*Jonas Anderegg[1]*, Andreas Hund[1], Petteri Karisto[2] and Alexey Mikaberidze[2]*

[1] Crop Science Group, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland, [2] Plant Pathology Group, Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

Hyperspectral remote sensing holds the potential to detect and quantify crop diseases in a rapid and non-invasive manner. Such tools could greatly benefit resistance breeding, but their adoption is hampered by i) a lack of specificity to disease-related effects and ii) insufficient robustness to variation in reflectance caused by genotypic diversity and varying environmental conditions, which are fundamental elements of resistance breeding. We hypothesized that relying exclusively on temporal changes in canopy reflectance during pathogenesis may allow to specifically detect and quantify crop diseases while minimizing the confounding effects of genotype and environment. To test this hypothesis, we collected time-resolved canopy hyperspectral reflectance data for 18 diverse genotypes on infected and disease-free plots and engineered spectral–temporal features representing this hypothesis. Our results confirm the lack of specificity and robustness of disease assessments based on reflectance spectra at individual time points. We show that changes in spectral reflectance over time are indicative of the presence and severity of *Septoria tritici* blotch (STB) infections. Furthermore, the proposed time-integrated approach facilitated the delineation of disease from physiological senescence, which is pivotal for efficient selection of STB-resistant material under field conditions. A validation of models based on spectral–temporal features on a diverse panel of 330 wheat genotypes offered evidence for the robustness of the proposed method. This study demonstrates the potential of time-resolved canopy reflectance measurements for robust assessments of foliar diseases in the context of resistance breeding.

**Keywords: high-throughput phenotyping, field-based phenotyping, feature engineering, feature selection, spectral vegetation index**

OPEN ACCESS

***Edited by:***
*Julie A. Dickerson, Iowa State University, United States*

***Reviewed by:***
*Maria Rosa Simon, Universidad Nacional de La Plata, Argentina*
*Sameh Selim, UniLaSalle, France*

****Correspondence:***
*Jonas Anderegg*
*jonas.anderegg@usys.ethz.ch*

***Specialty section:***
*This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science*

***Received:*** *21 May 2019*
***Accepted:*** *02 October 2019*
***Published:*** *25 October 2019*

***Citation:***
*Anderegg J, Hund A, Karisto P and Mikaberidze A (2019) In-Field Detection and Quantification of Septoria Tritici Blotch in Diverse Wheat Germplasm Using Spectral–Temporal Features. Front. Plant Sci. 10:1355. doi: 10.3389/fpls.2019.01355*

## INTRODUCTION

Hyperspectral remote sensing has shown significant potential for the rapid, non-invasive assessment of crop diseases at different scales, ranging from single leaves (*e.g.,* Mahlein et al., 2010; Ashourloo et al., 2014) to the canopy (*e.g.,* Cao et al., 2013; Yu et al., 2018) to fields and regions (Wakie et al., 2016). Applications have been proposed primarily in the context of precision agriculture, but resistance breeding may equally benefit (Mahlein, 2016). The identification of novel sources of durable, quantitative disease resistance requires screening large and diverse germplasm collections

under field conditions. Reflectance-based techniques hold the potential to reduce associated costs and allow for the screening of more genetic variation, if deployed on adequate phenotyping platforms (Kirchgessner et al., 2017; Aasen et al., 2018; Aasen and Bolten, 2018). This may enable indirect selection in early breeding generations and facilitate the identification of novel sources of resistance.

However, to benefit crop breeding, new methods must accurately estimate phenotypes for large numbers of diverse genotypes under field conditions (Furbank and Tester, 2011; Araus and Cairns, 2014; Araus et al., 2018). This represents a significant challenge because genotypic diversity and contrasting environmental conditions are major sources of variation in spectral reflectance. This variation arises mostly from i) genotype morphology, canopy cover, and canopy 3-D structure (Haboudane et al., 2002; Zarco-Tejada et al., 2005; Jacquemoud et al., 2009; Gutierrez et al., 2015); ii) differences in genotype phenology and the timing of developmental transitions such as heading, flowering, and senescence (Pimstein et al., 2009; Stuckens et al., 2011; Kipp et al., 2014); and iii) reactions to other biotic or abiotic stresses, which may result in similar spectral responses as the disease of interest (Zhang et al., 2012). At present, effects of diseases on canopy reflectance are often investigated at specific points in time (*see e.g.,* Yang, 2010; Cao et al., 2013; Yu et al., 2018). Such investigations have provided valuable but highly context-specific insights (i.e., specific to the genotype, growth stage and/or site and environment under study; *see e.g.*, Delalieux et al., 2007; Zhang et al., 2012; Zheng et al., 2019). Accordingly, identified spectral features and corresponding thresholds or calibration curves are not sufficiently robust (i.e., universally applicable) for use in resistance breeding. Largely due to such difficulties, high throughput phenotyping of disease resistance under field conditions using hyperspectral reflectance is still elusive (Araus et al., 2018).

*Septoria tritici* blotch (STB) caused by the fungal pathogen *Zymoseptoria tritici* is a serious threat to wheat production in major wheat growing areas around the world (Orton et al., 2011; Torriani et al., 2015). The development of cultivars with improved resistance to this disease has become a significant objective in wheat breeding and constitutes a key component of STB management strategies (O'Driscoll et al., 2014; Brown et al., 2015; McDonald and Mundt, 2016). Several major resistance genes conferring near-complete resistance to certain *Z. tritici* isolates have been identified and used in commercial cultivars (*reviewed by* Brown et al., 2015). However, these genes are frequently overcome within a few years of their introduction due to the high evolutionary potential of *Z. tritici* populations (McDonald and Mundt, 2016). Genetic loci conferring broadly effective partial resistance are thought to be more durable than major resistance genes (McDonald and Linde, 2002; McDonald and Mundt, 2016). However, sources of partial resistance are much more difficult to identify, as subtle differences in disease severity must be accurately quantified under field conditions, ideally over time.

Automated image analysis has shown great potential to accurately quantify STB resistance and characterize different components of resistance in genetically diverse breeding material (Stewart et al., 2016; Karisto et al., 2018). However, such measurements are more labor-intensive than visual scorings and do not provide the necessary throughput to routinely screen large breeding trials over time. Some recent work has investigated the potential of reflectance-based techniques to detect and quantify STB non-destructively at the leaf and canopy levels (Odilbekov et al., 2018; Yu et al., 2018). At the canopy level, the above-mentioned challenges are particularly pronounced in the case of STB, because epidemics frequently reach damaging levels and affect crop performance most during the grain filling phase (Bancal et al., 2007). Consequently, detection and quantification of STB must be achieved in fully developed canopies with a complex architecture, and a clear delineation of STB and physiological senescence is essential for efficient selection.

Recently, efforts have been made to increase the specificity of reflectance-based methods. For example, new spectral vegetation indices (SVIs) with improved specificity to diseases have been developed by several authors for various patho-systems (*e.g.,* Mahlein et al., 2013; Ashourloo et al., 2014). Other work has demonstrated that SVI combinations may allow to differentiate between diseases (Mahlein et al., 2010) and to delineate disease symptoms and nitrogen deficiency in wheat (Devadas et al., 2015). Yu et al. (2018) investigated the potential of different spectral features to robustly estimate STB severity at the canopy level in a large population of genetically diverse wheat genotypes. Other work has demonstrated that the sequence of temporal changes in hyperspectral reflectance signatures at the leaf level may be disease-specific, allowing to differentiate between sources of biological stress (Mahlein et al., 2010; Mahlein et al., 2012; Wahabzada et al., 2015).

Here, we aimed to achieve robust reflectance-based detection and quantification of STB under field conditions by exploiting changes in hyperspectral canopy reflectance over time. The basic rationale of the proposed approach is that pathogenesis consists in a specific and fixed sequence of events producing a constant outcome (i.e., disease symptoms). Accordingly, these events and outcomes should result in a specific and fixed sequence of changes in canopy spectral reflectance over time, irrespective of the genotype or environment under study. It seems highly likely that relying exclusively on this type of information increases the robustness of resulting estimations.

To test the feasibility of this approach, we engineered spectral–temporal features based on hyperspectral time series measurements. These features are designed to capture relevant changes in reflectance over time while minimizing the effect of the known confounding factors discussed above. We put forward the following hypotheses: ($H_1$) Confounding effects of contrasting morphology, canopy cover, and canopy 3-D structure are strongly reduced, if only relative changes in reflectance over time at the level of individual plots are analyzed. ($H_2$) Confounding effects of phenology can be reduced by using combinations of STB-sensitive and STB-insensitive spectral features. Specifically, we hypothesize that several plant traits are relatively unaffected in their temporal dynamics by STB. Thus, related spectral features can be used as a baseline of changes in spectral reflectance over time, arising primarily from advancing crop phenology. This baseline can then be used to correct temporal patterns observed

in STB-sensitive features for variation in phenology. Finally, we hypothesized ($H_3$) that the sequence and the dynamics of STB-sensitive features are to a certain extent specific to this disease and not related to other biotic or abiotic stresses.

Thus, the objective of this study was i) to evaluate the potential of time-resolved hyperspectral reflectance measurements to detect and quantify STB infections, ii) to delineate STB and physiological senescence, and iii) to estimate the robustness of the proposed method and hence its potential for breeding applications.

# MATERIALS AND METHODS

## Plant Materials, Experimental Design, Phenology, and Meteorological Data

A field experiment was carried out in the field phenotyping platform (FIP, Kirchgessner et al., 2017) at the ETH Research Station for Plant Sciences Lindau-Eschikon, Switzerland (47.449N, 8.682E, 520 m a.s.l.; soil type: eutric cambisol), in the wheat growing season of 2017–2018. A subset of 18 bread wheat (*Triticum aestivum*) genotypes was selected from the genom-analyse in biologischen system pflanze (GABI) wheat panel (Kollers et al., 2013; complemented with Swiss cultivars) for contrasting levels of resistance to STB and for contrasting stay-green properties based on previous experiments at the same location (Anderegg et al., 2019, *submitted to this issue*; Karisto et al., 2018). The set comprised morphologically diverse genotypes (e.g., awned and unawned), and there were obvious differences in canopy structural parameters (e.g., flag leaf angle) among the selected genotypes (**Supplementary Figure 1**). The study was conducted as a two-factorial experiment in a split-plot design with the presence/absence of artificial pathogen inoculation as a whole-plot factor and genotype as a split-plot factor.

Artificial inoculation with *Z. tritici* spore suspension was done on May 21, 2018. *Z. tritici* strains ST99CH_1A5, ST99CH_1E4, ST99CH_3D1, and ST99CH_3D7 were used (Zhan et al., 2002; see also http://www.septoria-tritici-blotch.net/isolate-collections.html). Spores were grown for 6 days in 200 ml of yeast sucrose broth liquid media (10 g yeast extract and 10 g sucrose in 1 L of water) in several flasks for each strain. The spore suspension was filtered and pooled together for each strain. Spore concentration was adjusted, and spore suspensions of each strain were mixed to achieve 150 ml of inoculum for each plot containing in total $10^6$ spores/ml (2.5 x $10^5$ sp/ml of each strain). Inoculum was sprayed in the evening into wet canopy of each plot.

There were two replications for the whole-plot factor. On the same site, the entire GABI panel was also grown in two replicates (two spatially separated lots). One replication of the inoculated plots each was located in a row adjacent to a lot of the GABI panel, separated by one row sown with the resistant cultivar CH NARA (DSP, Delley, Switzerland). One replication of the non-inoculated control plots each was spatially randomized within a lot of the GABI panel. The experiments were sown with a sowing density of 400 plants $m^{-2}$ on Oct 18, 2017. The plots sown with the GABI panel (and thus control plots within it) were treated with fungicides on three occasions: i) Input, Bayer (a mixture of sprioxamine at 300 g/L and prothioconazole at 150 g/L), was applied with a dose of 1.25 L/ha on 23 April 2018 (BBCH 31); ii) Aviator Xpro, Bayer (a mixture of bixafen at 75 g/L and metconazole at 41.25 g/L), was applied with a dose of 1.25 L/ha on 14 May 2018 (BBCH 51); and iii) Osiris, BASF (a mixture of epoxiconazole at 56.25 g/L and metconazole at 41.25 g/L), was applied with a dose of 2.5 L/ha on 28 May 2018 (BBCH 65). The inoculated control plots did not receive fungicide treatment. Temperature data was obtained from an on-site weather station. Rainfall data was obtained from a nearby weather station of the federal Swiss meteorological network Agrometeo (www.agrometeo.ch) located at ca. 250 m distance to the field trial. The temperature data was used to calculate growing degree-days (GDD) following

$$Tmean_d = \frac{\sum \frac{maxT_{d,h} + minT_{d,h}}{2} - baseT}{24}$$

$$GDD = \sum_{d=1}^{n} Tmean_d$$

where $Tmean_d$ is the mean temperature for day d after sowing, $maxT_{d,h}$ and $minT_{d,h}$ are hourly maximum and minimum temperatures for day d after sowing, respectively, and $baseT$ is the base temperature, set to 0°C. Heading date was recorded when 50% of the spikes were fully emerged from the flag leaf sheath (BBCH 59, Lancashire et al., 1991). BBCH scores within the main growth stages were linearly interpolated between assessment dates. Stay-green was assessed visually as described previously (Anderegg et al., 2019, *submitted*), separately for the flag leaf and the whole canopy, following guidelines provided by Pask et al. (2012). Flag leaf stay-green was scored based on the portion of green leaf area on a scale from 0 (0% green leaf area) to 10 (100% green leaf area). An integer mean value was estimated for plants located in a central region of about 0.5 m × 0.5 m of each plot. Canopy stay-green was scored on the same scale by estimating the overall greenness of the plot when inspected at a view angle of approximately 45° considering the entire plot area. All scorings were done by the same person in 2–3 day intervals. An overview of measurements, scorings, and samplings is given in **Table 1**. Growth stages were recorded until physiological maturity following the BBCH scale (Lancashire et al., 1991).

## Hyperspectral Reflectance Measurements

Canopy hyperspectral reflectance in the optical domain from 350 to 2500 nm was measured using a passive non-imaging spectroradiometer (ASD FieldSpec® 4 spectroradiometer, ASD Inc., USA) equipped with an optic fiber with a field of view of 25°. Five spectra were recorded as the average of 15–25 separate spectral records while moving the fiber optic once along the diagonal of each plot at a height of approximately 0.4 m above the canopy. A Spectralon® white reference panel was used for calibration before measuring canopy reflectance. The calibration was repeated after measuring one-half of a replicate (i.e., after nine plots, approximately every 3–5 min).

**TABLE 1 |** Overview of wheat phenology, canopy reflectance measurements, visual scorings, and samplings.

| Date | dai | dah | Reflectance | STB scoring | Leaf samplings | GS | Stg Cnp | Stg Fl0 |
|---|---|---|---|---|---|---|---|---|
| 25.05.2018 | 4 | 0 | | | | 59 | | |
| 30.05.2018 | 9 | 5 | x | | | 65 | | |
| 02.06.2018 | 11 | 7 | x | | | 73 | | |
| 05.06.2018 | 14 | 10 | x | | | | | |
| 06.06.2018 | 15 | 11 | | t1 | | | | |
| 10.06.2018 | 19 | 15 | x | | | | | |
| 14.06.2018 | 23 | 19 | | t2 | t2 (Fl1) | | 10 | 10 |
| 16.06.2018 | 25 | 21 | x | | | | 10 | 10 |
| 19.06.2018 | 28 | 24 | x | t3 | t3 (Fl0) | 77 | 10 | 10 |
| 20.06.2018 | 29 | 25 | x | | | | 9 | 10 |
| 22.06.2018 | 31 | 27 | x | | | | 8 | 9 |
| 26.06.2018 | 35 | 31 | x | t4 | t4 (Fl0) | 79 | 8 | 8 |
| 30.06.2018 | 39 | 35 | x | | | 83 | 3 | 3 |
| 02.07.2018 | 41 | 37 | x | t5 | t5 (Fl0) | | 1 | 1 |
| 07.07.2018 | 46 | 42 | x | | | 87 | 0 | 0 |
| 09.07.2018 | 48 | 44 | x | | | | 0 | 0 |
| 12.07.2018 | 51 | 47 | x | | | 89 | 0 | 0 |

*Visual incidence scorings of Septoria tritici blotch (STB), leaf samplings (Fl1 = second leaf, Fl0 = flag leaf), average growth stage (GS), canopy stay-green (Stg Cnp), and flag leaf stay-green (Stg Fl0) are reported. Days after inoculation (dai) and days after heading (dah) are indicated for each date.*

Measurements were carried out on 14 dates between heading and physiological maturity (i.e., between May 30 and July 12, 2018), resulting in an average of one measurement every 3 days. The maximum distance between two consecutive measurement dates was 6 days. In parallel, one lot of the GABI panel was measured on 13 dates. Here, the sensor calibration was repeated approximately every 10 min after completion of measurements on two rows.

## STB Disease Assessment

The amount of STB in each plot was assessed on five dates (t1–t5) between 16 days after inoculation (dai, June 6, 2018) and 42 dai (July 2, 2018). STB was quantified by combined assessments of disease incidence (i.e., the proportion of leaves showing visible symptoms of STB) and conditional disease severity (i.e., the amount of disease on symptomatic plants). Disease incidence was assessed visually for 30 plants per plot by inspecting the leaves of one tiller per plant. Incidence scorings were obtained per leaf layer. Conditional disease severity was then measured using automated analysis of scanned leaves exhibiting obvious disease symptoms. To this end, eight infected leaves were sampled per plot, transported to the laboratory, and imaged on flatbed scanners following the method described by Stewart et al. (2016) and Karisto et al. (2018). However, to avoid interfering excessively with the development of the disease epidemic, leaf samples from inoculated plots were taken only if disease incidence was at least 1/3 (i.e., if at least 10 out of 30 examined leaves exhibited symptoms of STB infection). Thus, no leaf samples were taken at t1, while at t2, second leaves from the top were sampled from a subset of plots. Starting at t3, all plots were sampled at the flag leaf layer. In contrast, from non-inoculated control plots, eight leaves were sampled without reference to their disease status due to very low disease incidence. Automated image analysis was then used to extract the percent of leaf area covered by lesions (PLACL) from the generated leaf scans using thresholds in the HSV color space and functions of the Python application program interface of OpenCV V3.0.0 (https://opencv.org/). The precision of the automated image analysis method used here to assess STB has been demonstrated repeatedly in greenhouse and field studies (Lendenmann et al., 2014; Zhong et al., 2017; Krishnan et al., 2018; Meile et al., 2018; Stewart et al., 2018). The procedure was optimized to minimize the effect of insect damage, powdery mildew infections, and physiological senescence, particularly leaf-tip necrosis, on the derivation of PLACL. PLACL was extracted only from t2 and t3 scans, as leaves increasingly displayed physiological senescence at later time-points. Finally, overall disease severity was calculated by multiplying disease incidence with conditional disease severity for inoculated plots, whereas it was directly extracted from the leaf scans for control plots.

## Data Analysis

All data analyses were done in the R environment for statistical computing (R version 3.5.2; R Core Team, 2018). Raw spectra were smoothed using the Savitzky–Golay smoothing filter (Savitzky and Golay, 1964) with a window size of 11 spectral bands and a third order polynomial, using the R package "*prospectr*" V0.1.3. (Ramirez-Lopez and Stevens, 2014). Spectral regions comprising the wavelengths from 1,350 to 1,475 nm, from 1,781 to 1,990 nm, and from 2,400 to 2,500 nm were removed because of the very low signal-to-noise ratio resulting from high atmospheric absorption. Spectra were averaged for each experimental plot. Pre-processed spectra, consisting of reflectance values at 1,709 wavelengths, were then used for time-point specific analysis as well as for time-integrated analyses, as described in the next sections. For ease of notation, the reflectance at a specific wavelength will be abbreviated by R followed by the wavelength in nm (e.g., R750).

## Benchmark Time-Point Specific Analysis

The relationship between spectral reflectance and STB was studied on a diverse panel of wheat genotypes by Yu et al. (2018), but the analysis was limited to single time-points. We performed a comparable analysis for each measurement time-point as a benchmark and to estimate model transferability across time. Yu et al. (2018) reported improved prediction of STB severity metrics and classification accuracy when using the full spectrum rather than single SVIs. Therefore, our analysis focuses on these approaches. We tested two parametrically structured linear models [partial least squares (PLS) regression and ridge regression] and two tree-based ensemble models (random forest regression and cubist regression) for their capability to predict STB severity metrics. For classification, PLS discriminant analysis (PLSDA) was used (*for details on these methods we refer to* Kuhn and Johnson, 2013 *and citations therein*). Prior to modeling, spectral resolution was reduced to 6 nm by binning (i.e., by computing average values for six adjacent wavelengths) due to very high correlation of reflectance values at neighboring wavelengths. Following a standard procedure (Kuhn and Johnson, 2013), model hyperparameters were tuned using 10 times repeated 10-fold cross-validation. Thus, training and test datasets comprised 90 and 10% of the original dataset, respectively. The root mean square error of predictions (RMSE) and overall classification accuracy were used as performance metrics for regression and classification, respectively. The overall accuracy reflects the agreement between the predicted and the observed classes. This agreement can also be expressed in terms of sensitivity and specificity of the model, with

$$Sensitivity = \frac{\#\ plots\ \ correctly\ \ predicted\ \ as\ \ diseased}{\#\ diseased\ \ plots} \quad \text{and}$$

$$Specificity = \frac{\#\ plots\ \ correctly\ \ predicted\ \ as\ \ healthy}{\#\ healthy\ \ plots}.$$

The simplest model with a performance within one standard error of the absolute best model was chosen as the final model. Variable importance for the projection (VIP) was computed for PLSDA models to estimate the importance of individual wavebands to predict the class (i.e., "healthy" or "diseased"). In the regression setting, two different training datasets were used for model fitting: the full dataset, including all control plots, and a dataset consisting of the inoculated plots only. When all control plots were used for model fitting, the RMSE and $R^2$ of the resulting models were adjusted by removing the predicted and observed values for the control plots again, in order to avoid overly optimistic performance estimates resulting from a good prediction of disease severity in control plots. The R packages "*caret*" V6.0.80 (Kuhn, 2008), "*mixOmics*" V6.3.2 (Rohart et al., 2017), "*pls*" V2.7.0 (Mevik et al., 2018), "*Cubist*" V0.2.2 (Kuhn et al., 2018), "*ranger*" V0.10.1 (Wright and Ziegler, 2017), and "*elasticnet*" V1.1.1 (Hastie, 2018) were used for the analysis.

## Time-Integrated Analysis

Summarizing $H_1$–$H_3$, we hypothesized that the analysis of temporal dynamics in hyperspectral reflectance signatures may facilitate a robust detection and quantification of STB across diverse wheat genotypes under field conditions. To evaluate these hypotheses, we condensed the hyperspectral time series into time series of SVIs, similar to the procedure described previously (Anderegg et al., 2019, *submitted*). Thereby, we obtained a comprehensive summary representation of the hyperspectral dataset collected over time, interpretable in terms of plant physiology and canopy characteristics. The smoothness of SVI values over time was evaluated graphically, and only SVIs showing a clear and interpretable temporal trend were maintained for further analyses. Values of the selected SVIs were scaled to range from 0 to 10, representing the minimum and maximum values recorded during the assessment period for the corresponding experimental plot, respectively. To simplify subsequent steps in the analysis, the scale for SVIs with increasing values over time was inverted. Measurement dates were converted to thermal time after heading by subtracting the plot-specific accumulated thermal time at heading from the accumulated thermal time at each measurement date. The scaled SVI values were then fitted against thermal time after heading for each experimental plot and SVI using a Gompertz model with asymptotes constrained to 0 and 10 (eq. 1).

$$S = 10\exp\left[-\exp\left[-b*\left(t-M\right)\right]\right] \tag{1}$$

where $S$ represents the scaled SVI value, $t$ is the accumulated thermal time after heading, $b$ is the rate of change at time $M$, and $M$ is the accumulated thermal time after heading when the rate is at its maximum (Gooding et al., 2000). Eq. (1) was fitted using the R package "*nls.multstart*" V1.0.0 (Padfield and Matheson, 2018). Two types of dynamics parameters for each experimental plot and SVI were extracted from the resulting model fits: 1) "key time-points," which are specific points in time when a certain criterion (e.g., a threshold) is met, and 2) "change parameters," which represent the rate or duration of a process (**Figure 1**). We extracted two key time-points: the $M$ parameter of the Gompertz model and the time when fitted values decreased to 8.5 (t85). As change parameters, we used the rate parameter $b$ of the Gompertz model and the duration between t85 and M (dur). While the b and M parameters of the Gompertz model fully describe the fitted curve, the t85 and dur parameters are affected by both Gompertz model parameters, thus representing a mix of both. The threshold was set to 8.5 because i) this level efficiently captured observed variation during the late stay-green phase (**Figure 2**), ii) it was little affected by somewhat unstable values during the stay-green phase observed for some SVIs, and iii) for some SVIs, the initial highest values were not optimally represented by the Gompertz model.

Next, SVIs least affected in their temporal dynamics by the presence or absence of STB infections were selected separately for each dynamics parameter as follows: for the key time-points (t85, M) by selecting SVIs with the smallest average difference between the parameter values of the inoculated and non-inoculated control plots and for change parameters (b, dur) by selecting SVIs with the smallest average deviance from 1 of the ratio of the change parameter values. For each dynamics parameter, a subset

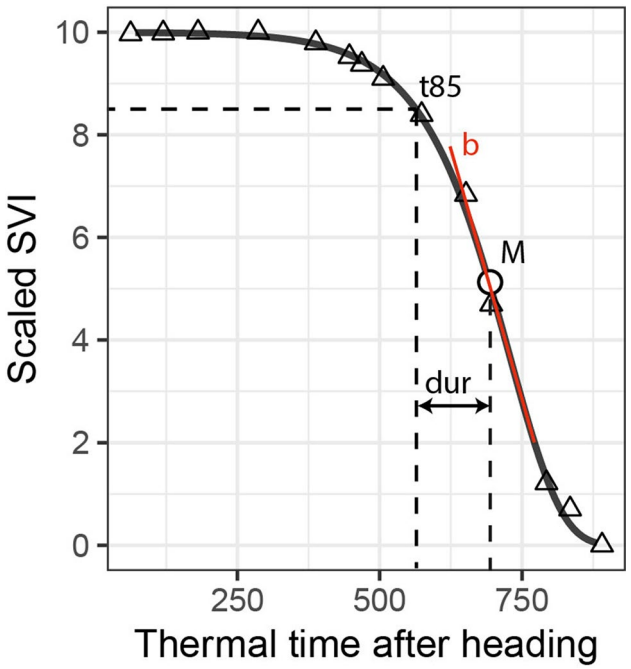


**FIGURE 1 |** Extraction of dynamics parameters for one spectral vegetation index (SVI; here scaled values of the plant senescence reflectance index) and one experimental plot (here a non-inoculated control plot). The t85 parameter is the time point when fitted scaled SVI values decrease to 8.5; M is a parameter of the Gompertz model, representing the time point when the rate of decrease is at its maximum; the dur parameter represents the duration in thermal time between t85 and M; and b is a parameter of the Gompertz model, representing the maximum rate of decrease. M and t85 are labeled "key points"; dur and b are labeled "change parameters."

of eight SVIs with the smallest difference or ratio was selected. All other SVIs were considered to be significantly affected by STB infection. **Figure 2** shows an example of an STB-sensitive and an STB-insensitive SVI.

We then performed unsupervised subset selection (i.e., without considering the response) on both sets of SVIs (i.e., the STB-sensitive and STB-insensitive SVIs) with the aim of removing redundant SVIs. For each dynamics parameter (t85, M, dur, b), pairwise Pearson correlation coefficients between the parameter values derived from all used SVIs were computed. For change parameters, the maximum linear correlation allowed was set to r = 0.9, whereas for the key time-points, it was set to r = 0.95, as these were generally highly collinear. In cases where pair-wise correlations were higher than these threshold values, only one of the two SVIs was retained, preferring narrow-band SVIs over broad-band SVIs, SVIs with a specific physiological interpretation, and SVIs developed specifically for use in wheat or barley canopies over more generic SVIs. Additionally, the goodness of the Gompertz model fit was evaluated qualitatively (i.e., graphically) and used as an additional selection criterion. The parameters were then combined by calculating differences between the key time-points derived from selected STB-sensitive and STB-insensitive SVIs and the ratios of the change parameters derived from STB-sensitive and STB-insensitive SVIs (**Figure 3**). These differences and ratios were calculated for all possible pairs of STB-sensitive and STB-insensitive SVIs and were then used as features for 1) the classification of plots into non-inoculated healthy control plots and inoculated, diseased plots and 2) the prediction of STB severity in each plot. This final step was performed primarily to estimate whether a combination of features outperform single features in predicting STB severity (relevant with respect to $H_3$) and to identify the most predictive features (see section 2.4.3). Models and model fitting procedures were identical to the time-point specific analysis.

### Selection of Spectral–Temporal Features

While tree-based models are considered naturally resistant to non-informative predictors and some perform feature selection intrinsically, the presence of highly correlated predictors makes the interpretation of resulting variable

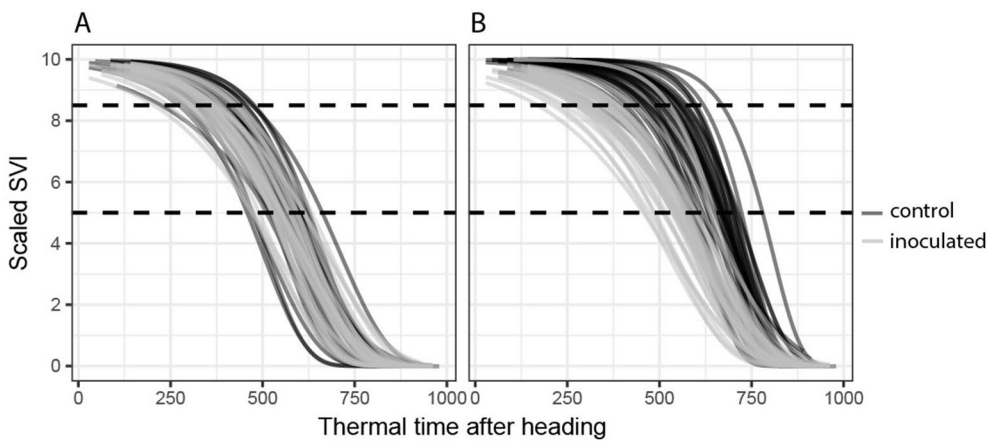


**FIGURE 2 |** Temporal dynamics of spectral vegetation indices (SVIs). Gompertz model fits for all 72 experimental plots are shown. **(A)** A disease-insensitive SVI (here the flowering intensity index, FII) displays the same temporal patterns for both treatments (inoculated and control). **(B)** A disease-sensitive SVI (here the modified chlorophyll absorption ratio index, MCARI2) displays contrasting temporal patterns for control and inoculated plots.

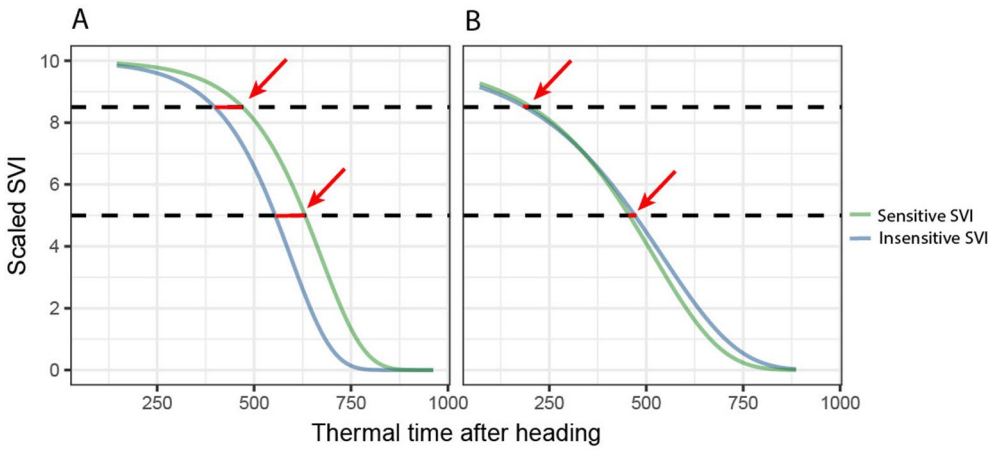


**FIGURE 3 |** Derivation of the final key time points based predictors for disease classification and quantification. Key time points extracted from disease-sensitive and disease-insensitive spectral vegetation indices (SVI) are combined to isolate the effect of the disease from other effects (e.g., contrasting stay-green) by calculating the differences (highlighted by red arrows). **(A)** Control, "healthy" plot. **(B)** Inoculated, "diseased" plot. For change parameters, the ratio, rather than the difference, was calculated.

importance measures challenging (Strobl et al., 2007). Hence, we performed supervised feature selection by recursive feature elimination with cubist and random forest regression as base-learners, using a nested cross-validation approach. The dataset was resampled 30 times with an 80:20 split using stratified sampling. Samples were binned into eight classes based on percentiles of STB severity to ensure balanced evaluation datasets. Thus, for each resample, feature elimination was carried out on 80% of the data, and model performance was evaluated on the remaining 20% in 28 decreasing steps. Eliminated features were assigned a rank corresponding to the iteration after which they were excluded (i.e., those eliminated first had rank 28, whereas the feature retained as the last had rank 1). In each iteration, the base-learner hyperparameters were tuned using 10-fold cross-validation (*see* Ambroise and McLachlan, 2002; Guyon et al., 2002; Granitto et al., 2006 *for a detailed discussion of the methodology*).

## Independent Model Validation

Due to the relatively small size of the experiment (n = 72 experimental plots), we did not rely exclusively on the cross-validated training performance estimates for model evaluation. The performance and robustness of the developed models were further evaluated using data of 360 wheat plots sown with 330 registered varieties obtained from one replication of the GABI panel. Low to intermediate disease incidence and very low conditional disease severity as well as late appearance of symptoms in all 36 control plots spatially randomized within the two replications of the panel suggested that STB disease should not have reached damaging levels in the vast majority of these plots. Therefore, these plots were considered as essentially disease-free. For all of these plots, spectral–temporal features were extracted from the 13 measurement time-points as described previously and were then used to generate a class label and class probabilities from the classification models as well as a prediction of disease severity from the regression models. To distinguish the performance measures obtained for held-out samples of the main experiment (i.e., the cross-validated training performance) from those obtained for the independent plots, these are referred to as the internal accuracy ($acc_{int}$) and the generalized accuracy ($acc_{gen}$), respectively. It is important to note that accuracy represents only model specificity in this case, as no independent plots with significant levels of STB were available. In a final validation step, the spatial distribution of class labels and severity predictions were examined by creating plots of the experimental design. Thus, we aimed to test the robustness of the models to heterogeneous field conditions. Field heterogeneity may affect plant physiology and thus hyperspectral reflectance over time (e.g., through the development of local drought stress). The presence or absence of spatial patterns in model predictions can therefore be interpreted as an additional measure of model robustness.

## Validation of the Most Predictive Feature in a Contrasting Environment

Spectral–temporal features were engineered specifically to minimize effects of genotype and environment. However, it is still conceivable that relative changes in spectral reflectance over time are also affected by environmental conditions. This may result in the extraction of environment-specific relationships between spectral–temporal features and disease severity. We therefore evaluated the relationship between the most predictive spectral–temporal feature and STB severity using data from a separate year. This dataset enables a rigorous validation of the spectral–temporal features as predictors of STB for several reasons: i) it originates from a strongly contrasting environment with wet and cool weather conditions; ii) sowing parameters were different, likely affecting canopy structural parameters; iii) only natural variance in disease resistance was observed, as

artificial inoculations were not performed; iv) a large number of morphologically, phenologically, and structurally diverse genotypes not contained in the training population were assessed; v) reflectance measurements were not carried out with a sufficient frequency to allow fitting of parametric models, and we had to use linear interpolations of individual measurement time points instead, losing the advantage of the smoothing effect; and finally, vi) sampling procedures to quantify STB severity were not optimal for our purpose. As parametric models could not be fitted, the time point when interpolated values decreased below 50% of their initial value ($t_{50}$) was extracted as an equivalent to the *M* parameter of the Gompertz model. For more details on datasets and experiments, we refer to Karisto et al. (2018) and Anderegg et al. (2019, *submitted*).

# RESULTS

## Development of STB Disease

Toward the end of the vegetation period, all inoculated plots had substantial levels of STB. In contrast, non-inoculated control plots were essentially disease-free until late in the vegetation period. Thus, artificial inoculations were effective in all plots, and the dataset was suitable for testing the feasibility of the classification of plots into healthy and diseased canopies based on reflectance spectra or spectral–temporal features. Furthermore, large variation in the levels of STB could be observed among the inoculated plots, probably attributable to different levels of resistance, with the largest variation observed during late stay-green (i.e., at t3, June 19, 2018, compare with **Figure 4B**). Thus, the dataset was also suitable for testing the feasibility of disease quantification using reflectance spectra or spectral–temporal features.

High levels of STB in inoculated plots were the result of both high incidence and high conditional severity, particularly at t4 and t5 (**Figure 5A**, **Table 2**). Visual assessments of scanned leaves suggested a high conditional severity in all inoculated plots at these late stages. In contrast, the non-inoculated control plots displayed very low levels of STB even at late stages. STB incidence increased in some plots at t4 and t5, but visual assessments of the sampled leaves demonstrated very low conditional disease severity even at t5. Since the subset of genotypes used for the experiment also included some highly susceptible genotypes, this suggests that natural infections did not cause agronomically significant levels of STB in this experiment. This was likely the result of fungicide applications and the very low rainfall in the period from May to July. Rainfall in this period totaled 178 mm, which represents 52% of the average of 343 mm in the reference period 1981–2010 (MeteoSwiss, 2019).

STB incidence was low at t1, both in inoculated and in control plots. Symptoms were apparent at significant levels only on lower leaf layers of inoculated plots, whereas flag leaves were essentially disease-free in both treatments. At t2, there was approximately a five-fold increase in STB incidence at the flag leaf and subtending leaf layer in many inoculated plots (**Figure 5A**, **Table 2**). At t3, STB incidence on flag leaves reached very high levels in most inoculated plots, and PLACL reached an average of 17%,

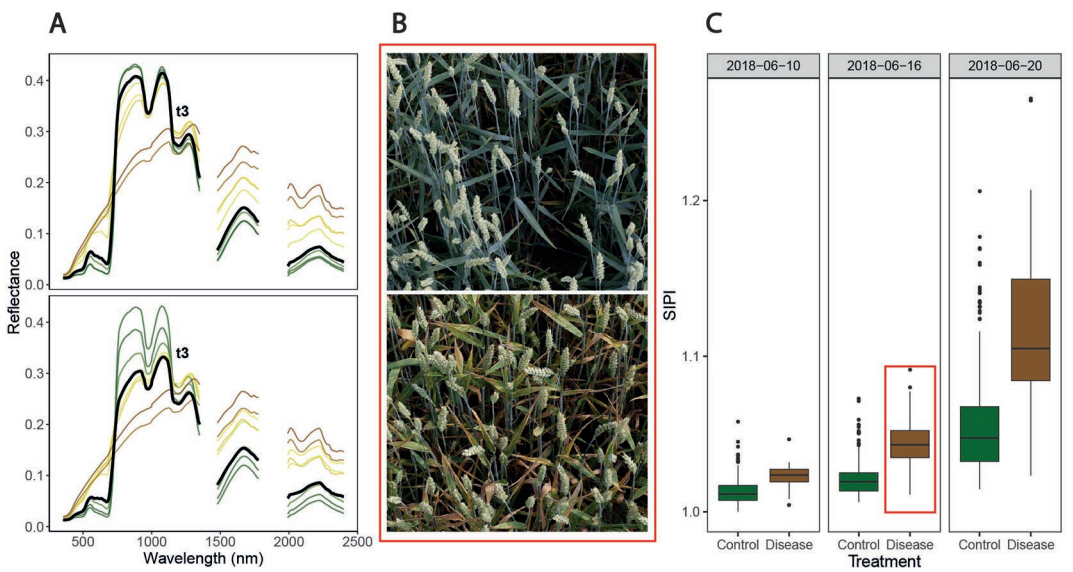


**FIGURE 4 |** Symptoms of *Septoria tritici* blotch (STB) and associated spectral reflectance characteristics over time. **(A)** Date-wise averaged reflectance spectra of healthy canopies (upper panel) and of inoculated, diseased canopies (lower panel). Colors approximate the average color of the vegetation on the corresponding measurement date (estimated based on average visual canopy senescence scorings). The thick black spectra mark the average reflectance spectra measured at t3 (i.e., June 19, 2018). **(B)** Images of two inoculated plots, taken on June 16, 2018. The genotype in the upper panel was highly resistant to STB, developing visible symptoms only later, whereas the genotype in the lower panel was highly susceptible and displays severe symptoms of STB. Images were captured using the field phenotyping platform (FIP, Kirchgessner et al., 2017). **(C)** Values of the structure insensitive pigment index (SIPI) for both treatments on three measurement dates. The plots shown in panel B are contained in this boxplot (indicated by the red box). No obvious signs of apical senescence were visible by June 20, 2018, in any of the healthy control plots, but senescence started shortly after.

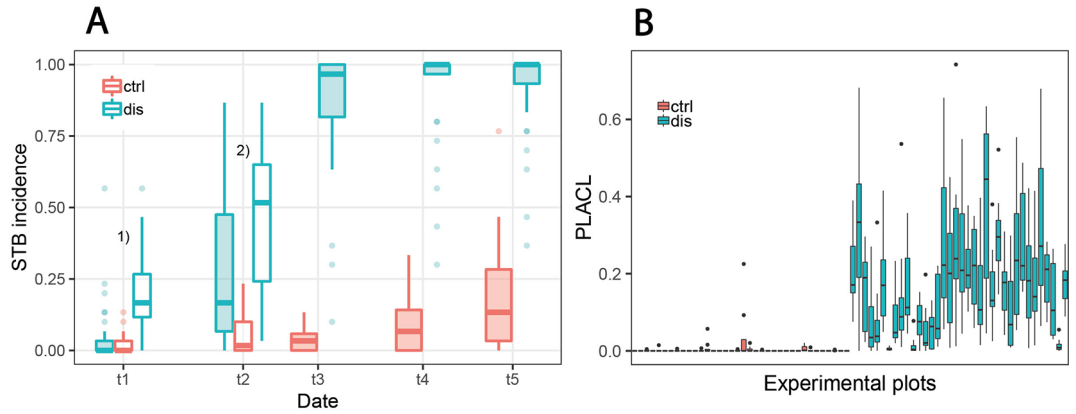


**FIGURE 5 |** Development of *Septoria tritici* blotch (STB) disease. **(A)** STB incidence at five different assessment dates for non-inoculated "healthy" control plots (ctrl) and for artificially inoculated "diseased" plots (dis). In diseased plots, STB incidence on flag leaves was assessed at all time points, whereas for the control plots, it was assessed only from t3 onward. Filled boxes represent STB incidence on flag leaves; open boxes represent STB incidence on lower leaf layers. [1)]Open boxes represent STB incidence on third leaves from the top; [2)]open boxes represent STB incidence on second leaves from the top. **(B)** Conditional disease severity measured as percent leaf area covered by lesions (PLACL) at t3 (June 19, 2018) for eight flag leaves per plot for all 72 experimental plots.

**TABLE 2 |** Summary of *Septoria tritici* blotch (STB) assessments. STB incidence and conditional severity were assessed at the level of individual leaf layers, namely, the flag leaf (Fl0), the penultimate leaf (Fl1), and the ante-penultimate leaf (Fl2). Conditional severity was measured as percent leaf area covered by lesions (PLACL). Severity was calculated as the product of STB incidence and conditional severity. Values are reported separately for non-inoculated control plots and inoculated plots, separated by a slash. Mean values across all plots are reported, with minima and maxima in brackets. Disease assessments were carried out on five dates (t1–t5) covering the growth phases of 15 days after inoculation (dai) to 41 dai.

| Time point | dai | STB incidence | | | PLACL | | Severity | |
|---|---|---|---|---|---|---|---|---|
| | | Fl0 | Fl1 | Fl2 | Fl0 | Fl1 | Fl0 | Fl1 |
| t1 | 15 | -/ | -/ | 0.02 (0, 0.13)/ | -/ | -/ | -/ | -/ |
| | | 0.05 (0, 0.57) | 0.10 (0, 0.53) | 0.20 (0, 0.57) | - | - | - | - |
| t2 | 23 | -/ | 0.05 (0, 0.23)/ | -/ | -/ | -/ | -/ | -/ |
| | | 0.26 (0, 0.87) | 0.45 (0.03, 0.87) | - | - | 0.05 (0, 0.25) | - | 0.03 (0, 0.20) |
| t3 | 28 | 0.03 (0, 0.13)/ | 0.09 (0, 0.50)/ | -/ | -/ | -/ | **0 (0, 0.04)/** | -/ |
| | | 0.87 (0.10, 1) | 0.91 (0.10, 1) | - | 0.17 (0, 0.38) | - | **0.16 (0, 0.38)** | - |
| t4 | 35 | 0.09 (0, 0.33)/ | -/ | -/ | -/ | -/ | -/ | -/ |
| | | 0.92 (0.30, 1) | - | - | - | - | - | - |
| t5 | 41 | 0.17 (0, 0.77)/ | -/ | -/ | -/ | -/ | -/ | -/ |
| | | 0.92 (0.37, 1) | - | - | - | - | - | - |

indicating a moderate conditional STB severity on average. Thus, the observed differences in STB severity among inoculated plots are primarily a result of differences in conditional severity, with PLACL ranging from 0 to 38% (**Figure 5B**). In control plots, almost no lesions were detected. There were some signs of physiological senescence on sampled flag leaves at t3, but these were mostly limited to yellowing of the entire leaves and/or leaf tip necrosis. As there was ample variation for disease severity at t3 among the inoculated plots and physiological senescence did not significantly affect extraction of PLACL from leaf scans, this time-point was chosen as a measure of overall disease severity and used as response variable in the time-integrated analysis.

## Effects of STB and Phenology on Canopy Spectral Reflectance

Over the assessment period, observed changes in spectral reflectance were similar for diseased and healthy canopies (**Figure 4A**), showing the typical pattern of senescing canopies. However, an obvious effect of STB infections consisted in an early marked decrease in reflectance in the near-infrared region (NIR) not observable in healthy canopies. This decrease preceded the appearance of physiological senescence and was observable in the pre-symptomatic phase of STB infections. Furthermore, an early increase in reflectance in the VIS, especially for wavelengths greater than 535 nm, was observed. An early increase in short-wave infrared region (SWIR) reflectance in diseased canopies compared to healthy canopies was also discernible. However, these differences were small compared to changes in reflectance over time.

Canopy spectral reflectance seemed to remain relatively constant throughout the stay-green phase in healthy canopies (**Figure 4A**, upper panel). However, the examination of SVI values over time revealed significant changes in canopy reflectance during this period (**Figure 4C**). Importantly, variation caused by advancing phenology (i.e., within-treatment variation

in **Figure 4C**) was prominent with respect to STB-induced variation (i.e., between-treatment variation in **Figure 4C**). This is true even for the structure insensitive pigment index (SIPI), which has been proposed as a potential surrogate for crop disease under field conditions (Yu et al., 2018; **Figure 4C**). For several other spectral indices, initial variation as well as variation over time was even larger (data not shown).

## Time-Point Specific Full-Spectrum Analysis

### Binary Classification Into Healthy and Diseased Canopies Using Reflectance Spectra

PLSDA models correctly classified all held-out samples in most resampling iterations, resulting in classification accuracies $acc_{int} \geq 0.96$ for all time-points (**Figure 6**). The optimal number of components used by the PLSDA models (determined *via* repeated CV) was between 5 and 17, depending on the time-point. Correct class labels were obtained for all held-out samples even for the first time-point at 9 dai, when no visual symptoms of STB were present in most experimental plots. However, prediction accuracies for the independent GABI plots were distinctly lower, particularly for models calibrated with data from early and late measurement time-points (data not shown). A satisfactory performance on independent plots was observed for models calibrated with data from the late stay-green phase (i.e., between 2018-06-10 and 2018-06-22), which correctly predicted the independent plots as disease-free in most cases ($acc_{gen} \geq 0.88$ in all cases).

VIP scores quantify the importance of wavebands to predict the response, i.e., to generate the class label ("healthy" or "diseased") or to predict STB disease severity here (Yu et al., 2018). VIP scores for the first three components showed some general patterns across time-points (**Figure 6**). The NIR (750–1,300 nm) and the SWIR (1,475–1,781 and 1,991–2,400 nm) had a relatively high importance (**Figure 6**). However, the

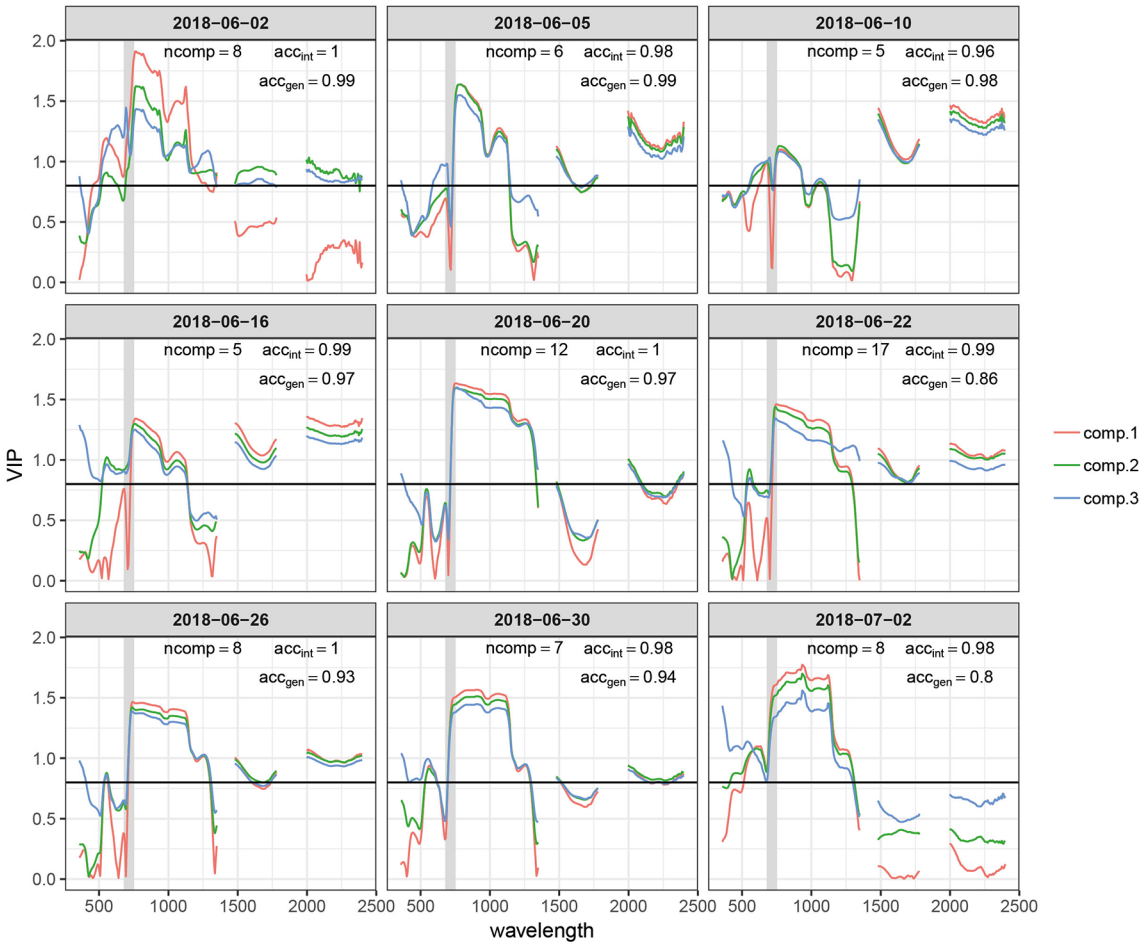


**FIGURE 6 |** Variable importance for the projection (VIP) of the time point specific partial least squares discriminant analysis (PLSDA) models for the first three components (comp.1–comp.3). The total number of components used by the model (ncomp), the prediction accuracy for plots included in the experiment ($acc_{int}$), and prediction accuracy for the independent test set comprising >300 plots of an adjacent experiment measured on the same day or with a maximum delay of one day ($acc_{gen}$) are also reported for each model. The earliest and latest time points are not represented. The gray shaded area represents the spectral range between 680 and 750 nm, i.e., the red edge. The horizontal black line marks a commonly used threshold value for an important contribution (VIP = 0.8).

relative importance of the SWIR compared to the NIR drastically changes over time. The importance of the SWIR is comparably low during early stay-green, but its importance greatly increases and exceeds the importance of the NIR during late stay-green. Furthermore, the red edge (RE, 680–750 nm) had a low importance at the beginning but is increasingly important at later stages, as indicated by a gradual left-shift of the peak in VIP at the NIR for later time-points. Finally, at early time-points, there is a significant contribution of wavebands in the visible range (VIS, 400–700 nm). This feature is somewhat transformed over time, resulting in a narrow peak in VIP at wavelengths around 535 nm at intermediate time-points. Toward later time-points, this feature broadens again.

Given the common patterns but also significant differences across time-points, we aimed to evaluate the robustness of the developed models to temporal changes in reflectance induced by advancing crop phenology, as differences in crop phenology are typically present among genotypes in breeding programs. Model performance across time is shown in **Figure 7**. The performance of models calibrated with data from early and late time-points quickly deteriorates. In contrast, models calibrated with data of intermediate time-points show a higher stability over time, although the performance of some models still decreases rather fast. The models created using data from 2018-06-10 and 2018-06-19 were most robust over time and produce accurate class predictions over a period of about 10 days. It is essential to note that these performance estimates are derived from the same plots used to calibrate the models, although at different time-points. Given the lower $acc_{gen}$ compared to the $acc_{int}$ (see above), significantly decreased performance should be expected on entirely independent plots (different genotypes).

### Regression Models to Quantify Disease Severity Using Reflectance Spectra

Cubist regression models performed best in predicting disease severity. The smallest RMSE was obtained for models trained on data from inoculated plots only (RMSE = 0.061, $R^2$ = 0.67; **Figure 8**). The underlying model was simple, building on only four variables (R748, R766, R892, and R1084) in a single model tree. Model performance was slightly decreased when all available data was used for model fitting ($RMSE_{adj}$ = 0.066, $R^2_{adj}$ = 0.55). Here, significant improvements were achieved by increasing model complexity. Validation on the largely disease-free plots of the GABI panel suggested a high specificity of the model (i.e., disease levels on healthy plots were predicted to be virtually zero for almost all plots). This was true irrespective of whether the control plots were included in the training dataset or not. PLS and ridge regression performed similarly (RMSE = 0.063, $R^2$ = 0.64), whereas random forest regression performed comparably poorly (RMSE = 0.087, $R^2$ = 0.46). A strong systematic bias was observed in predictions of the random forest, with low values of disease severity overestimated and high values underestimated.

## Time-Integrated Analysis Using Combinations of Spectral Vegetation Indices

### Engineering of Spectral–Temporal Features as New Predictors

Fifty-seven of the tested SVIs were deemed amenable for analysis of their temporal dynamics in the proposed framework (i.e., they displayed a clear and interpretable temporal trend, which could be represented using a Gompertz-type model). For seven of these, the last three measurements were excluded prior to

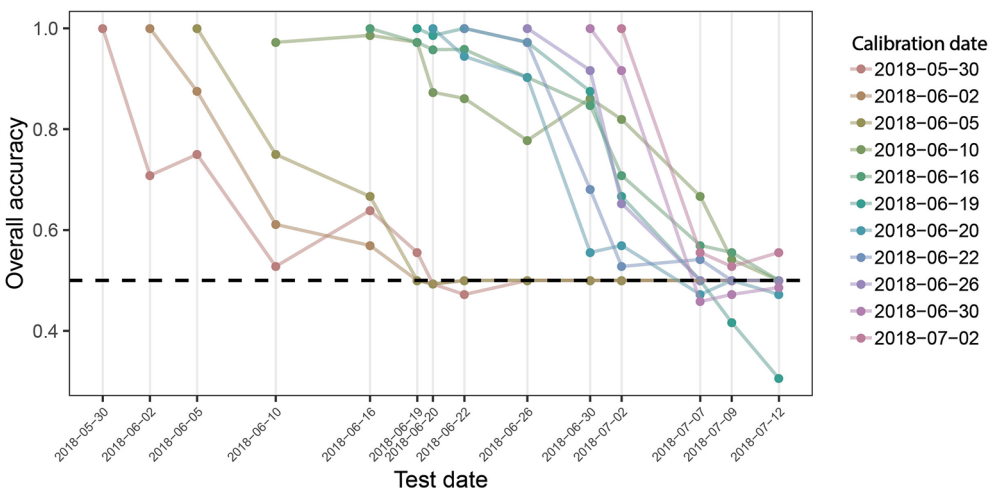


**FIGURE 7 |** Overall prediction accuracy of PLSDA models across time. Binary classification models were calibrated for each measurement time point, using reflectance spectra collected on this date as predictors and the known class label (i.e., "diseased" or "healthy") as response. The performance of these models was evaluated on held out samples of the same date as well as on all plots of the subsequent measurement time points as the overall accuracy of classification. Colored lines track the performance of each date-specific model across time (e.g., the left-most red line represents the performance of the model calibrated with spectra obtained on May 30, 2018, when tested on the same date, and for all subsequent measurement dates). The broken black line indicates a performance of 0.5, i.e., the performance of a random guess of the class label.

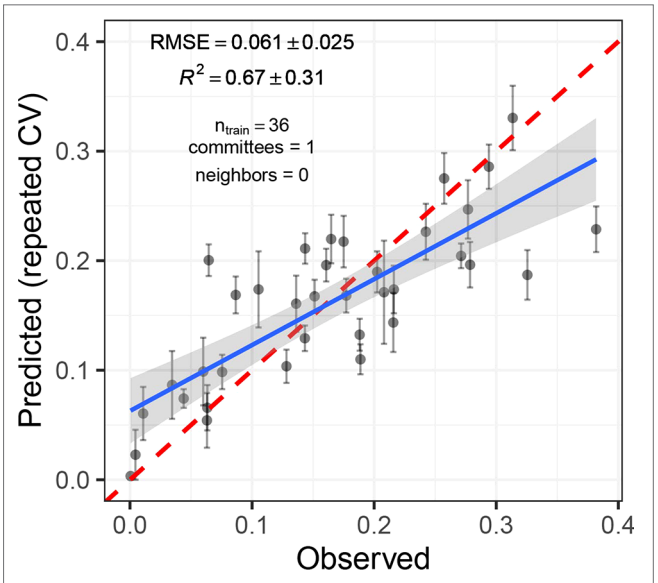


**FIGURE 8 |** Predicted *vs*. observed *Septoria tritici* blotch (STB) severity levels of 36 artificially inoculated wheat plots. Mean and standard error of predictions are shown. STB severity was measured on flag leaves using a combination of visual incidence scorings and scans of flag leaves exhibiting disease symptoms. Predictions were obtained from a cubist regression model based on the reflectance spectrum of the canopies measured on June 19, 2018. The broken red line represents the 1:1 line, the blue line represents the least squares line of the linear regression of predicted *vs*. observed values, and the gray area represents the 95% confidence interval of the least squares line.

modeling their temporal dynamics, as values increased again in later stages. After subset selection, 7 and 13 SVIs were retained as insensitive and sensitive SVIs, respectively, for the key time-points. For change parameters, 4 and 12 SVIs were retained as insensitive and sensitive SVIs, respectively. In total, 24 distinct SVIs were retained, of which 10 sensitive and 14 insensitive SVIs. From their fitted dynamics, a total of 278 (i.e., 4 insensitive SVIs * 12 sensitive SVIs * 2 parameters + 7 insensitive SVIs * 13 sensitive SVIs * 2 parameters) spectral–temporal features were generated as pairwise combinations of dynamics parameters obtained from sensitive and insensitive SVIs. These features were then used for classification and regression, as described below.

## Binary Classification Into Healthy and Diseased Canopies Using Spectral–Temporal Features

A PLSDA model using four components achieved a classification accuracy $acc_{int}$ = 1.00, suggesting correct classification of each experimental plot as healthy or diseased canopy based on spectral–temporal features. In the external validation, the model achieved $acc_{gen}$ = 0.86, thus correctly classifying 304/353 plots as healthy. This is slightly less accurate than the best time-point specific models (**Figure 6**).

## Regression Models to Quantify Disease Severity Using Spectral–Temporal Features

Overall, disease severity predictions from spectral–temporal features were similarly accurate as those obtained from time-point specific models based on reflectance spectra. The lowest RMSE was achieved when no control plots were used as training data using the PLS algorithm (RMSE = 0.068, $R^2$ = 0.71). Differences in performance among algorithms were smaller than in time-point specific analyses, but the random forest performed relatively poorly also in this case (RMSE = 0.076, $R^2$ = 0.62). Both tree-based models, and particularly the random forest, produced systematically biased predictions, which was not observable for PLS and ridge regression. In contrast to the time-point specific analyses, validation on the largely disease-free plots of the GABI panel suggested the necessity to include the control plots in the training data in order to obtain accurate predictions for low levels of disease, except for ridge regression. When these were included, tree-based models produced good estimates of the low disease-levels, whereas PLS and ridge regression still predicted disease severity of >0.05 in a significant number of plots. The inclusion of the control plots resulted in a lower systematic bias of the tree-based predictions, while only marginally decreasing model performance ($RMSE_{adj}$ = 0.074 and $RMSE_{adj}$ = 0.076 for cubist and random forest, respectively; **Figure 9A**). Importantly, the inclusion of the control plots also strongly reduced or eliminated spatial patterns in predictions of the GABI panel, except for ridge regression (**Figure 9B**). Thus, cubist regression seemed to perform best when taking all evaluated aspects of model performance into account.

## Feature Selection and Validation

Feature selection was performed to identify the most important spectral–temporal features and to estimate the benefit of adding additional features. The difference between *M* derived from the modified chlorophyll absorption ratio index (MCARI2) and the SIPI was consistently the most informative spectral–temporal feature (**Supplementary Table 1**). The MCARI2 was designed to estimate green leaf area index in crop canopies, whereas the SIPI measures pigment concentrations and ratios in leaves. This feature was retained as the last in all 30 resamples by the random forest and in 29 resamples by cubist. Following features had much increased ranks. The most influential features were all based on the *M* parameter of the Gompertz model, while other parameters were clearly less important. In particular, differences in change parameters did not seem to be informative of disease severity. Most selected features used the SIPI, R780/R740, and PRInorm indices as STB-insensitive index, even though they seemed to be somewhat more affected by the presence of disease than the flowering intensity index (FII) on average (**Table 3**). There was little evidence for the existence of complementary information among the spectral–temporal features, as model performance was affected little by the sequential removal of features (**Figure 10**). However, the small sample size resulted in very high variance of the performance estimates obtained from the test set and contrasting patterns between the cross-validated training and the test performance estimates (**Figure 10**). The top-selected spectral–temporal feature was found to be informative of STB severity in the separate experiment carried out under contrasting environmental conditions (Pearson r = 0.53, $p < 0.001$; **Figure 11**).

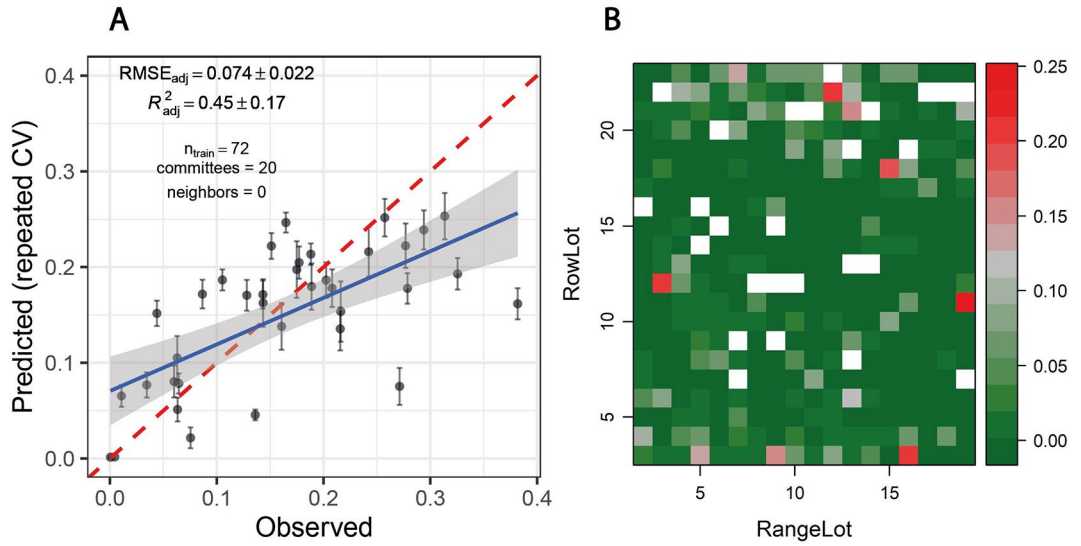


**FIGURE 9 | (A)** Predicted *vs*. observed *Septoria tritici* blotch (STB) severity levels of 36 artificially inoculated wheat plots. Mean and standard error of predictions are shown. Data from all experimental plots (n = 72) was used to tune/train the model, but reported performance estimates are based only on artificially inoculated plots (n = 36) in order to avoid overly optimistic performance estimates resulting from a good prediction of disease severity in control plots. STB severity was measured on flag leaves using a combination of visual incidence scorings and scans of flag leaves exhibiting disease symptoms. Predictions were obtained from a cubist regression model based on spectral–temporal features for the same 36 plots and 36 non-inoculated control plots sown with the same genotypes. The broken red line represents the 1:1 line, the blue line represents the least squares line of the linear regression of predicted *vs*. observed values, and the gray area represents the 95% confidence interval of the least squares line. **(B)** Spatial distribution of predicted STB severity levels of ~360 largely disease-free plots of the GABI wheat panel, grown next to the plots used as training dataset. White fields correspond to the plots contained in the training dataset.

**TABLE 3 |** Spectral vegetation indices (SVI) identified to be insensitive in their temporal dynamics to the presence or absence of *Septoria tritici* blotch (STB) disease. For each dynamics parameter, a subset of eight SVI with the smallest difference (diff, GDD) or ratio (dimensionless) of the parameters between treatments was selected. The reported SVI were retained after subset selection based on pairwise correlations. The mean pairwise correlation (corr) is reported per dynamics parameter. Values in brackets report the minimum and maximum pairwise correlations.

| SVI | DynPar | diff/ratio | Pairwise corr |
|---|---|---|---|
| R780/R740 | t85 | 75.3 | 0.66 (0.26, 0.95) |
| CHLRE | t85 | 55.6 | |
| DCNI | t85 | 45.6 | |
| FII | t85 | 33.2 | |
| GNDVI | t85 | 81.3 | |
| PRInorm | t85 | 58.2 | |
| SIPI | t85 | 45.3 | |
| R780/R740 | M | 37 | 0.82 (0.60, 0.95) |
| CHLRE | M | 55.6 | |
| DCNI | M | 25.7 | |
| FII | M | 4.4 | |
| GNDVI | M | 40.5 | |
| PRInorm | M | 21.6 | |
| SIPI | M | 22.7 | |
| R780/R750 | b | 1.03 | 0.73 (0.48, 0.88) |
| CHLRE | b | 0.98 | |
| MTCI | b | 0.95 | |
| SR | b | 0.99 | |
| R780/R750 | dur | 1.03 | 0.75 (0.53, 0.88) |
| CHLRE | dur | 1 | |
| MTCI | dur | 1.02 | |
| SR | dur | 0.99 | |

# DISCUSSION

## Limitations of Time-Point Specific Analyses

Current reflectance-based approaches to high throughput phenotyping of crop diseases under field conditions suffer from a lack of specificity and from insufficient robustness to genotypic diversity and environmental variability (i.e., context specificity). This problem has previously been described in detail with regard to growth stages of the crop, different phases in the pathogenesis, and the presence of other stresses (Zhang et al., 2012; Devadas et al., 2015; Zheng et al., 2019).

Our results prominently illustrate context-specificity of the relationship between spectral reflectance and disease. Firstly, variation on a specific date in potentially disease-sensitive spectral features, such as the SIPI (*see* Bajwa et al., 2017; Yu et al., 2018), is quickly overridden by variation caused by advancing phenology (**Figure 4C**), illustrating the difficulty in defining thresholds or calibration curves. Secondly, unstable VIP values of single wavelengths in PLSDA models, systematic shifts in VIP patterns (**Figure 6**), and limited model applicability over time, even for the plots contained in the training dataset (**Figure 7**), illustrate marked short-term changes in the relationship between STB and spectral reflectance. Thirdly, the decreased classification accuracy on independent test plots (**Figure 6**) indicates context-specificity related to the effect of genotypes and, possibly, field heterogeneity.

There was a short period during late stay-green when classification models were transferable between time-points to

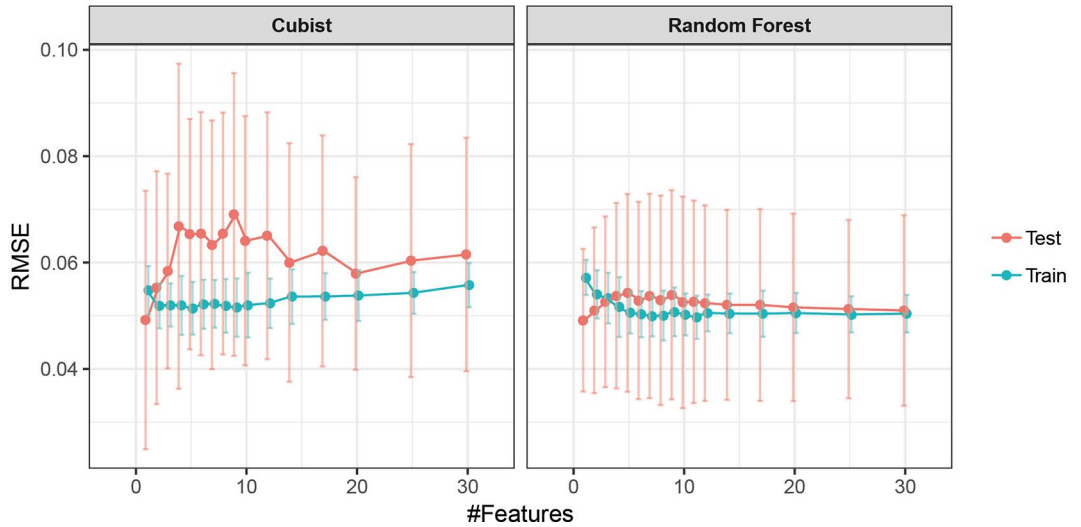


**FIGURE 10 |** Performance profile of models based on spectral–temporal features to predict STB severity as a function of the number of spectral–temporal features used. Mean performance and standard deviation are shown based on 30 resamples of the data.

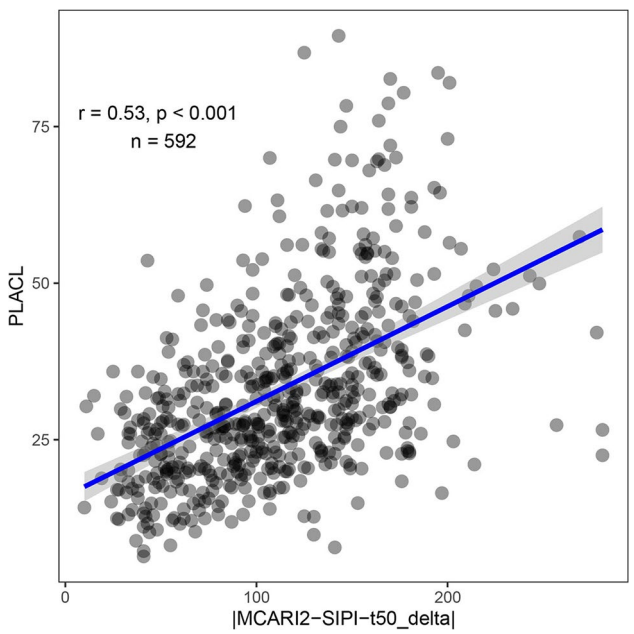


**FIGURE 11 |** STB severity as measured by the percent leaf area covered by lesions (PLACL) plotted against the spectral–temporal feature identified as most predictive of STB severity (the difference between $t_{50}$ extracted from the MCARI2 and SIPI spectral vegetation indices, expressed in growing degree days). The Pearson product moment correlation coefficient and the p-value of the linear correlation are based on 592 experimental plots. Data was collected in a separate experiment conducted in 2016 without artificial inoculation. In this experiment, 330 genotypes were grown in two replicates, but some plots were excluded from the analysis due to heavy lodging.

some extent (**Figure 6**). This can be explained by the relatively synchronized appearance of moderate to high levels of STB in front of the relatively stable background of a stay-green canopy. In this intermittent phase, the signal caused by STB is strong compared to the noise caused by genotypic diversity and in-field measurements (see also **Figure 4A**). Nonetheless, the regression model for STB-severity based on reflectance spectra is still context-specific, as reflectance in the NIR (used as predictors) gradually decreases during the stay-green and senescence phase irrespective of the presence of STB (**Figure 4A**). NIR reflectance is also strongly affected by genotype morphology, canopy 3-D structure, and canopy cover (Jacquemoud et al., 2009; Gutierrez et al., 2015) and is therefore not specific to STB if analyzed on a particular point in time. In addition, time-point specific models highlight the potential of detecting STB in different phases, using different spectral features. This potential would be left unused if only a short period would be targeted.

## Potential of Temporal Changes in Reflectance to Detect and Quantify STB

Due to the strong limitations of models based on reflectance spectra, we evaluated the potential of exploiting temporal changes in reflectance for disease detection and quantification instead. Models based on spectral–temporal features were characterized by a somewhat lower performance compared with models trained on reflectance spectra of a specific time-point (**Figure 7**, **Figure 8**). Nevertheless, classification accuracies were similar to the time-point specific PLSDA models, and regression models suggested that spectral–temporal features were also informative of disease severity. This is encouraging, particularly given the strongly contrasting morphological, canopy structural, and stay-green properties of the genotypes comprised in the experiment.

## Selected Spectral Indices and Resulting Spectral–Temporal Features

Even though high levels of STB developed during the stay-green phase in most artificially inoculated plots (**Figure 4**), several SVIs could be identified which displayed similar temporal patterns across treatments (**Table 3**). In particular, the FII (Stuckens et al.,

2011), i.e., the normalized difference of R475 and R365, was found to be almost unaffected by STB (**Figure 2A**). In a previous study, we found that early physiological senescence of wheat canopies results in only (proportionally) small increases in reflectance at wavelengths shorter than 500 nm (unpublished data). Strong increases were observable only toward later stages of senescence. The observed insensitivity of the FII to STB likely results from the fact that STB affects only parts of the vegetation, initially mostly lower leaf layers, while significant amounts of healthy green tissue remain. Thus, FII values should change significantly only with the onset of rapid apical senescence, encompassing a generalized loss of green leaf area. It has been suggested that STB does not accelerate or anticipate apical senescence under a range of environmental conditions (Bancal et al., 2016). This is in line with the observed insensitivity of the FII to STB. Interestingly, the dynamic pattern of the SIPI (Penuelas et al., 1995) was also found to be highly insensitive to STB. In contrast, this SVI was previously suggested as a potential surrogate for crop disease under field conditions (Bajwa et al., 2017; Yu et al., 2018). This index was developed at the leaf scale to maximize sensitivity to the ratio between carotenoid and chlorophyll *a* concentrations (Car/Chl *a* ratio), while minimizing the effect of leaf surface and mesophyll structure (Penuelas et al., 1995). Provided the principles underlying the SIPI hold also for canopy level reflectance, a low sensitivity of the dynamic pattern to the presence of STB would be expected, as there seems to be no reason to expect a significant change at canopy level of the Car/Chl *a* ratio due to STB. STB causes the appearance of localized necrotic lesions; however, a general increase in the Car/Chl *a* ratio is not expected, unless STB accelerates or anticipates apical senescence, which does not seem to be the case (Bancal et al., 2016). The PRInorm was also among the most STB-insensitive SVIs. This SVI is based on the photochemical reflectance index (PRI), initially employed to measure changes in the relative levels of pigments in the xanthophyll cycle (Gamon et al., 1992). Over larger temporal scales, the PRI was shown to be strongly responsive to the Car/Chl ratio (Sims and Gamon, 2002). Zarco-Tejada et al. (2013) modified this SVI to decrease the effect of reduced canopy leaf area resulting from water stress. Thus, its insensitivity to STB can probably be explained in an analogous manner as for the SIPI.

The temporal patterns of water-sensitive SVIs such as the water index (WI; Peñuelas and Inoue, 1999) and the normalized difference water index (NDWI; Gao, 1996) were found to be highly sensitive to STB. Similarly, the disease water stress index (DSWI; Apan et al., 2004), which uses information from the water-sensitive SWIR and the NIR, was strongly affected in its temporal dynamics. In particular, water sensitive SVIs decreased much earlier for inoculated than for control plots, and the decrease occurred more gradually than in healthy plots (data not shown). This is in line with findings by Yu et al. (2018), who reported both the WI and NDWI to discriminate best between STB-diseased and healthy canopies in early stages of disease development. Several SVIs using reflectance in the RE and NIR also showed strongly contrasting temporal patterns [e.g., DSWI, normalized difference vegetation index, plant senescence reflectance index (PSRI), and VI700]. Similar to the SIPI, the PSRI (Merzlyak et al., 1999) is highly sensitive to the Car/Chl ratio at the leaf level. However, NIR reflectance is used to normalize the difference between R677 and R500. It seems highly questionable whether the PSRI is particularly sensitive to the Car/Chl ratio in diverse germplasm at the canopy level. Variation in the PSRI seems to arise primarily from differences in canopy structure among genotypes and from canopy structural changes over time (Anderegg et al., 2019, *submitted*). The modified chlorophyll absorption ratio index (MCARI2; Haboudane et al., 2004), sensitive to green leaf area index, also showed strongly contrasting dynamic patterns (**Figure 2B**). The healthy index (HI; Mahlein et al., 2013) was developed to separate healthy sugar beet leaf tissue from tissues affected by various foliar diseases. The prominent use of the RE by this index suggests that in our case, HI values are mostly driven by canopy structure and to a lesser extent chlorophyll absorption. Overall, our results thus suggest that SVIs sensitive to leaf internal structure and canopy structure are strongly affected by the presence of STB. This effect has been previously described for various patho-systems (*e.g.,* Zhang et al., 2012; Yu et al., 2018; Zheng et al., 2019).

In summary, it seems that many of the derived spectral–temporal features can be interpreted as robust measures of STB-induced temporal changes in leaf internal structure, canopy structural parameters, and canopy water content. These are obtained by normalizing the temporal dynamics of corresponding SVIs *via* the estimation of temporal changes in pigment ratios and reflectance at short wavelengths centered around 465 nm, likely representing physiological apical senescence. Thus, spectral–temporal features seem to well represent our hypothesis $H_2$, as STB-affected plant and canopy traits are expressed relative to phenology-related traits.

## Robustness of Spectral–Temporal Features

As far as quantifiable in the framework of this experiment, models based on spectral–temporal features were robust to variation in genotype morphology, phenology, canopy cover, and canopy 3D structure as well as genotype-specific temporal changes thereof. Accurate predictions were obtained also for 330 diverse genotypes comprised in the GABI panel and grown in a large field experiment, suggesting robustness of the method to varying growing conditions arising from field heterogeneity (**Figure 9B**). The most predictive feature was also among the best predictors of STB severity in a different experiment. The observed correlation is very similar to the correlations reported by Yu et al. (2018) for simple SVI measured at individual time points, however with the advantages discussed throughout the manuscript. The results of this final validation likely underestimate the power of the developed approach for several reasons as described in section 2.4.5 but offer evidence for the transferability of our results across sites, environments, genotypes, and agricultural practices.

Results from feature selection suggested that a single spectral–temporal feature (i.e., the difference between *M* derived from MCARI2 and SIPI), relating structural changes in leaves and canopies to senescence-induced changes in pigment composition, was sufficient to achieve the performance illustrated above. Other stresses occurring during grain filling such as terminal drought stress and nitrogen shortages are likely to result in a similar decrease in green leaf area index. However, these stresses

are also known to accelerate physiological senescence (Martre et al., 2006; Bogard et al., 2011; Distelfeld et al., 2014). Therefore, we speculate that the developed models may be moderately robust against the effect of common other stresses despite their simplicity. Yet, we conclude that our hypothesis $H_3$ (i.e., that the combination of several spectral–temporal features representing the unique sequence and dynamics of separate events during pathogenesis could increase the specificity of the method) remains to be confirmed in larger experiments including other stress factors. In particular, other diseases causing similar symptoms and prevalent in the same developmental stage of the crop may have similar effects on the temporal evolution of hyperspectral reflectance. Large multifactor experiments will be required to judge the potential of the proposed approach to detect, quantify, and delineate individual necrotrophic foliar diseases. Finally, it should be noted that some effects of fungicide applications on canopy reflectance characteristics cannot be excluded in our experiment. However, a fungicide formulation without greening effect was used for the last treatment at BBCH 65. It seems unlikely that this or earlier fungicide applications significantly affected the temporal dynamics of canopy reflectance.

## Multiple Spectral Vegetation Indices to Exploit Temporal Dynamics in Reflectance

A key component of the proposed approach consists in summarizing hyperspectral data in terms of multiple SVIs and modeling of their temporal dynamics. Though this may result in the loss of some relevant information contained in reflectance spectra (Pauli et al., 2016), the use of SVIs presented a number of advantages here: i) noise in temporal patterns was much reduced compared to reflectance values at single wavelengths, facilitating the fitting of parametric models; ii) the inevitable subset selection step preceding feature combination could be based on objective criteria related to the form and purpose of SVIs; iii) many of the used SVIs were designed specifically to maximize responsiveness to certain vegetation properties while minimizing the effect of common confounding factors, which is likely to also increase the robustness of derived spectral–temporal features (*see e.g.*, Penuelas et al., 1995; Haboudane et al., 2004); and finally, iv) the procedure results in a summary of the hyperspectral dataset that is interpretable in terms of plant physiology and canopy characteristics, which also holds true for derived spectral–temporal features. Fitting parametric models to scaled SVI values may smooth out measurement errors related to single measurement dates, resulting for example from varying sun angle at measurement or short-term variation in illumination conditions. Thus, scaling SVI values and modeling their temporal dynamics reduces the effect of confounding factors on initial reflectance spectra and minimizes the effect of errors related to single measurements in the series.

## Context and Scope

In this study, we used a non-imaging spectroradiometer and manual feature engineering for disease detection and quantification. A high spatial resolution of imaging sensors has been deemed critical for disease detection, identification, and quantification by others (Mahlein et al., 2010; Mahlein et al., 2012; Mahlein, 2016). The high potential of 2-D information in combination with deep learning methods for disease identification has been demonstrated recently (Mohanty et al., 2016; Fuentes et al., 2017). However, changes in spectral reflectance over time have also been shown to be highly informative at the leaf level (Mahlein et al., 2010; Wahabzada et al., 2015). To make use of the spatial and temporal dimensions under field conditions, individual lesions would arguably have to be tracked across time. Some solutions to this problem have been presented for close-range hyperspectral measurements (Behmann et al., 2018). However, similar solutions at the canopy level may be technically extremely challenging to implement and require extensive studies due to problems in tracking individual pixels or organs over time and in obtaining a clean spectral signal from objects with varying orientation. Existing approaches to make use of spectral, spatial, and temporal information rely on automated and data-driven extraction of characteristic spectral features for diseased plants under controlled conditions (Wahabzada et al., 2015; Wahabzada et al., 2016; Thomas et al., 2018). Here, promising results were achieved using a non-imaging sensor and manual feature extraction. This highlights that an improved understanding of potential confounding factors arising under field conditions may equally boost the potential of remote sensing methods for applications in crop breeding.

We developed and validated the presented method to facilitate robust in-field detection and quantification of STB. However, the underlying concepts should be transferrable to different problems, such as the detection and quantification of other foliar diseases. Several features of the proposed approach (e.g., exploiting plot-based relative changes in reflectance over time, combining sensitive and insensitive features, or the SVI-based parameterization of temporal dynamics) may also be valuable in quantifying other breeding-relevant traits, such as the timing and dynamics of nitrogen remobilization.

# CONCLUSION

Here, we tested the possibility to detect and quantify STB relying exclusively on relative changes in spectral reflectance over time, which is expected to minimize confounding effects on spectral reflectance arising from genotypic diversity and environmental conditions. Our results demonstrated the feasibility of the proposed approach and suggested that resulting models were robust against variation in several common nuisance factors. Specifically, it appears that the temporal dynamics in green leaf area index when set in relation to the dynamics of physiological apical senescence is highly indicative of the presence of STB infections and of STB severity. Time-resolved measurements of the MCARI2 and the SIPI SVIs could allow to assess these traits at very high throughput, facilitating time-resolved large-scale screenings of breeding nurseries.

Larger calibration experiments will offer the opportunity to evaluate the inclusion of additional spectral–temporal features that better capture relevant information in different phases of pathogenesis. This is likely to improve sensitivity and specificity of resulting models, which should also be tested in more detail.

Furthermore, the evaluation of the scalability to unmanned aerial vehicles will represent a crucial step toward application of such methods in breeding programs.

## DATA AVAILABILITY STATEMENT

The datasets generated and analyzed for this study can be found in the ETH Zürich publications and research data repository (https://www.research-collection.ethz.ch/) and can be downloaded from the following link: https://doi.org/10.3929/ethz-b-000370027. All analysis scripts are publicly available. Development repositories: https://github.com/and-jonas/Andereggetal2019b and https://github.com/and-jonas/stb_placl. Programming language: R, Python. License: GNU General Public License, version 3 (GPL-3.0).

## AUTHOR CONTRIBUTIONS

All authors planned the experiment. AH created the experimental design. PK prepared and carried out pathogen inoculation. AM and PK designed the STB assessment strategy and supervised STB scorings, collections and scanning of leaves. JA performed reflectance measurements. JA performed leaf scan analysis, conceived the proposed data analysis approach, analyzed the data and drafted the manuscript. AM and AH supervised the project. All authors contributed to, read and approved the final version of the manuscript.

## FUNDING

The project was partially funded by the Swiss Federal Office of Agriculture (FOAG) and by the Swiss National Science Foundation through Ambizione grant PZ00P3_161453.

## ACKNOWLEDGMENTS

We thank Prof. Dr. Achim Walter for input to the overall conception of the study and Mr. Hansueli Zellweger for field management and crop husbandry. The authors gratefully acknowledge fruitful discussions with Mr. Philipp Baumann and Dr. Helge Aasen. Pathogen inoculation was supported by Mr. Reto Zihlmann. Measurements in the field and STB assessment campaigns were supported by Ms. Kelbet Nagymetova, Ms. Delphine Piccot, Mr. Michel Colombo, Mr. Mario Zurfluh and Mr. Pablo Bovy. The project was partially funded by the Swiss Federal Office of Agriculture (FOAG). AM and PK gratefully acknowledge financial support by the Swiss National Science Foundation through Ambizione grant PZ00P3_161453.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01355/full#supplementary-material

## REFERENCES

Aasen, H., and Bolten, A. (2018). Multi-temporal high-resolution imaging spectroscopy with hyperspectral 2D imagers – From theory to application. *Remote Sens. Environ.* 205, 374–389. doi: 10.1016/j.rse.2017.10.043

Aasen, H., Honkavaara, E., Lucieer, A., and Zarco-Tejada, P. J. (2018). Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows. *Remote Sens.* 10, 1091. doi: 10.3390/rs10071091

Ambroise, C., and McLachlan, G. J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. *PNAS* 99, 6562–6566. doi: 10.1073/pnas.102102699

Anderegg, J., Yu, K., Aasen, H., Walter, A., Liebisch, F., and Hund, A. (2019). Spectral vegetation indices to track senescence dynamics in diverse wheat germplasm. *Front. Plant Sci.*

Apan, A., Held, A., Phinn, S., and Markley, J. (2004). Detecting sugarcane 'orange rust' disease using EO-1 Hyperion hyperspectral imagery. *Int. J. Remote Sens.* 25, 489–498. doi: 10.1080/01431160310001618031

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., and Cairns, J. E. (2018). Translating high-throughput phenotyping into genetic gain. *Trends Plant Sci.* 23, 451–466. doi: 10.1016/j.tplants.2018.02.001

Ashourloo, D., Mobasheri, M. R., and Huete, A. (2014). Developing two spectral disease indices for detection of wheat leaf rust (Pucciniatriticina). *Remote Sens.* 6, 4723–4740. doi: 10.3390/rs6064723

Bajwa, S. G., Rupe, J. C., and Mason, J. (2017). Soybean disease monitoring with leaf reflectance. *Remote Sens.* 9, 127. doi: 10.3390/rs9020127

Bancal, M.-O., Ben Slimane, R., and Bancal, P. (2016). Zymoseptoria tritici development induces local senescence in wheat leaves, without affecting their monocarpic senescence under two contrasted nitrogen nutrition. *Env. Exp. Bot.* 132, 154–162. doi: 10.1016/j.envexpbot.2016.09.002

Bancal, M.-O., Robert, C., and Ney, B. (2007). Modelling wheat growth anduield losses from late epidemics of foliar diseases using loss of green leaf area per layer and pre-anthesis reserves. *Ann. Bot.* 100, 777–789. doi: 10.1093/aob/mcm163

Behmann, J., Bohnenkamp, D., Paulus, S., and Mahlein, A.-K. (2018). Spatial referencing of hyperspectral images for tracing of plant disease symptoms. *J. Imaging* 4, 143. doi: 10.3390/jimaging4120143

Bogard, M., Jourdan, M., Allard, V., Martre, P., Perretant, M. R., Ravel, C., et al. (2011). Anthesis date mainly explained correlations between post-anthesis leaf senescence, grain yield, and grain protein concentration in a winter wheat population segregating for flowering time QTLs. *J. Exp. Bot.* 62, 3621–3636. doi: 10.1093/jxb/err061

Brown, J. K. M., Chartrain, L., Lasserre-Zuber, P., and Saintenac, C. (2015). Genetics of resistance to Zymoseptoria tritici and applications to wheat breeding. *Fungal Genet. Biol.* 79, 33–41. doi: 10.1016/j.fgb.2015.04.017

Cao, X., Luo, Y., Zhou, Y., Duan, X., and Cheng, D. (2013). Detection of powdery mildew in two winter wheat cultivars using canopy hyperspectral reflectance. *Crop Prot.* 45, 124–131. doi: 10.1016/j.cropro.2012.12.002

Delalieux, S., van Aardt, J., Keulemans, W., Schrevens, E., and Coppin, P. (2007). Detection of biotic stress (Venturia inaequalis) in apple trees using hyperspectral data: Non-parametric statistical approaches and physiological implications. *Eur. J. Agron.* 27, 130–143. doi: 10.1016/j.eja.2007.02.005

Devadas, R., Lamb, D. W., Backhouse, D., and Simpfendorfer, S. (2015). Sequential application of hyperspectral indices for delineation of stripe rust infection and nitrogen deficiency in wheat. *Precision Agric.* 16, 477–491. doi: 10.1007/s11119-015-9390-0

Distelfeld, A., Avni, R., and Fischer, A. M. (2014). Senescence, nutrient remobilization, and yield in wheat and barley. *J. Exp. Bot.* 65, 3783–3798. doi: 10.1093/jxb/ert477

Fuentes, A., Yoon, S., Kim, S. C., and Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* 17, 2022. doi: 10.3390/s17092022

Furbank, R. T., and Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Gamon, J. A., Peñuelas, J., and Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* 41, 35–44. doi: 10.1016/0034-4257(92)90059-S

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

Gooding, M. J., Dimmock, J. P. R. E., France, J., and Jones, S. A. (2000). Green leaf area decline of wheat flag leaves: the influence of fungicides and relationships with mean grain weight and grain yield. *Ann. Appl. Biol.* 136, 77–84. doi: 10.1111/j.1744-7348.2000.tb00011.x

Granitto, P. M., Furlanello, C., Biasioli, F., and Gasperi, F. (2006). Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. *Chemometr. Intell. Lab. Syst.* 83, 83–90. doi: 10.1016/j.chemolab.2006.01.007

Gutierrez, M., Reynolds, M. P., and Klatt, A. R. (2015). Effect of leaf and spike morphological traits on the relationship between spectral reflectance indices and yield in wheat. *Int. J. Remote Sens.* 36, 701–718. doi: 10.1080/01431161.2014.999878

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learn.* 46, 389–422. doi: 10.1023/A:1012487302797

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., and 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 Sens. Environ.* 90, 337–352. doi: 10.1016/j.rse.2003.12.013

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., and Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416–426. doi: 10.1016/S0034-4257(02)00018-4

Hastie, H. Z. T. (2018). elasticnet: Elastic-Net for Sparse Estimation and Sparse PCA.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., et al. (2009). PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sens. Environ.* 113, S56–S66. doi: 10.1016/j.rse.2008.01.026

Karisto, P., Hund, A., Yu, K., Anderegg, J., Walter, A., Mascher, F., et al. (2018). Ranking quantitative resistance to septoria tritici blotch in elite wheat cultivars using automated image analysis. *Phytopathology*. 108(5), 568—581. doi: 10.1094/PHYTO-04-17-0163-R

Kipp, S., Mistele, B., and Schmidhalter, U. (2014). Identification of stay-green and early senescence phenotypes in high-yielding winter wheat, and their relationship to grain yield and grain protein concentration using high-throughput phenotyping techniques. *Funct. Plant Biol.* 41, 227–235. doi: 10.1071/FP13221

Kirchgessner, N., Liebisch, F., Yu, K., Pfeifer, J., Friedli, M., Hund, A., et al. (2017). The ETH field phenotyping platform FIP: a cable-suspended multi-sensor system. *Funct. Plant Biol.* 44, 154–168. doi: 10.1071/FP16165

Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., et al. (2013). Whole genome association mapping of fusarium head blight resistance in european winter wheat (Triticum aestivum L.). *PLoS One* 8, e57500. doi: 10.1371/journal.pone.0057500

Krishnan, P., Meile, L., Plissonneau, C., Ma, X., Hartmann, F. E., Croll, D., et al. (2018). Transposable element insertions shape gene regulation and melanin production in a fungal pathogen of wheat. *BMC Biology* 16, 78. doi: 10.1186/s12915-018-0543-2

Kuhn, M. (2008). Building Predictive Models in *R* Using the caret Package. *J. Stat. Softw.* 28, 1–26. doi: 10.18637/jss.v028.i05

Kuhn, M., and Johnson, K. (2013). *Applied Predictive Modeling*. New York, NY: Springer New York. doi: 10.1007/978-1-4614-6849-3

Kuhn, M., Weston, S., Keefer, C., Coulter, N., and code), R.Q. (Author of imported C., code), R.R.P.L. (Copyright holder of imported C. (2018). Cubist: Rule- And Instance-Based Regression Modeling.

Lancashire, P. D., Bleiholder, H., Boom, T. V. D., Langelüddeke, P., Stauss, R., Weber, E., et al. (1991). A uniform decimal code for growth stages of crops and weeds. *Ann. Appl. Biol.* 119, 561–601. doi: 10.1111/j.1744-7348.1991.tb04895.x

Lendenmann, M. H., Croll, D., Stewart, E. L., and McDonald, B. A. (2014). Quantitative trait locus mapping of melanization in the plant pathogenic fungus zymoseptoria tritici. *Genes Genomes Genet.* 4, 2519–2533. doi: 10.1534/g3.114.015289

Mahlein, A.-K. (2016). Plant disease detection by imaging sensors – parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis.* 100, 241–251. doi: 10.1094/PDIS-03-15-0340-FE

Mahlein, A.-K., Oerke, E.-C., Steiner, U., and Dehne, H.-W. (2012). Recent advances in sensing plant diseases for precision crop protection. *Eur. J. Plant Pathol.* 133, 197–209. doi: 10.1007/s10658-011-9878-z

Mahlein, A.-K., Rumpf, T., Welke, P., Dehne, H.-W., Plümer, L., Steiner, U., et al. (2013). Development of spectral indices for detecting and identifying plant diseases. *Remote Sens. Environ.* 128, 21–30. doi: 10.1016/j.rse.2012.09.019

Mahlein, A.-K., Steiner, U., Dehne, H.-W., and Oerke, E.-C. (2010). Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. *Precision Agric.* 11, 413–431. doi: 10.1007/s11119-010-9180-7

Martre, P., Jamieson, P. D., Semenov, M. A., Zyskowski, R. F., Porter, J. R., and Triboi, E. (2006). Modelling protein content and composition in relation to crop nitrogen dynamics for wheat. *Eur. J. Agron.* 25, 138–154. doi: 10.1016/j.eja.2006.04.007

McDonald, B. A., and Linde, C. (2002). The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica* 124, 163–180. doi: 10.1023/A:1015678432355

McDonald, B. A., and Mundt, C. C. (2016). How Knowledge of pathogen population biology informs management of septoria tritici blotch. doi: 10.1094/PHYTO-03-16-0131-RVW

Meile, L., Croll, D., Brunner, P. C., Plissonneau, C., Hartmann, F. E., McDonald, B. A., et al. (2018). A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance to septoria tritici blotch. *New Phytol.* 219, 1048–1061. doi: 10.1111/nph.15180

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant* 106, 135–141. doi: 10.1034/j.1399-3054.1999.106119.x

MeteoSwiss. (2019). Climate Normals Zürich/Kloten. Reference period 1981–2010. Federal Department of Home Affairs FDHA, Federal Office of Meteorology and Climatology, MeteoSwiss. Available online at: http://www.meteoswiss.admin.ch.

Mevik, B.-H., Wehrens, R., Liland, K. H., and Hiemstra, P. (2018). pls: Partial Least Squares and Principal Component Regression.

Mohanty, S. P., Hughes, D. P., and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 7, 1419. doi: 10.3389/fpls.2016.01419

Odilbekov, F., Armoniené, R., Henriksson, T., and Chawade, A. (2018). Proximal phenotyping and machine learning methods to identify septoria tritici blotch disease symptoms in wheat. *Front. Plant Sci.* 9, 685. doi: 10.3389/fpls.2018.00685

O'Driscoll, A., Kildea, S., Doohan, F., Spink, J., and Mullins, E. (2014). The wheat–Septoria conflict: a new front opening up? *Trends Plant Sci.* 19, 602–610. doi: 10.1016/j.tplants.2014.04.011

Orton, E. S., Deller, S., and Brown, J. K. M. (2011). Mycosphaerella graminicola: from genomics to disease control. *Mol. Plant Pathol.* 12, 413–424. doi: 10.1111/j.1364-3703.2010.00688.x

Padfield, D., and Matheson, G. (2018). nls.multstart: Robust Non-Linear Regression using AIC Scores.

Pask, A., Pietragalla, J., Mullan, D., and Reynolds, M. P. (2012). Physiological breeding II : a field guide to wheat phenotyping iv, 132 pages.

Pauli, D., Chapman, S. C., Bart, R., Topp, C. N., Lawrence-Dill, C. J., Poland, J., et al. (2016). The quest for understanding phenotypic variation *via* integrated approaches in the field environment. *Plant Physiol.* 172, 622–634. doi: 10.1104/pp.16.00592

Penuelas, J., Baret, F., and Filella, I. (1995). Semiempirical indexes to assess carotenoids chlorophyll-a ratio from leaf spectral reflectance. *Photosynthetica* 31, 221–230.

Peñuelas, J., and Inoue, Y. (1999). Reflectance indices indicative of changes in water and pigment contents of peanut and wheat leaves. *Photosynthetica* 36, 355–360. doi: 10.1023/A:1007033503276

Pimstein, A., Eitel, J. U. H., Long, D. S., Mufradi, I., Karnieli, A., and Bonfil, D. J. (2009). A spectral index to monitor the head-emergence of wheat in semi-arid conditions. *Field Crops Res.* 111, 218–225. doi: 10.1016/j.fcr.2008.12.009

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria. Vienna, Austria: R Foundation for Statistical Computing.

Ramirez-Lopez, L., and Stevens, A. (2014). prospectr: Miscellaneous functions for processing and sample selection of vis-NIR diffuse reflectance data.

Rohart, F., Gautier, B., Singh, A., and Cao, K.-A. L. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* 13, e1005752. doi: 10.1371/journal.pcbi.1005752

Savitzky, A., and Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* 36, 1627–1639. doi: 10.1021/ac60214a047

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. doi: 10.1016/S0034-4257(02)00010-X

Stewart, E. L., Hagerty, C. H., Mikaberidze, A., Mundt, C. C., Zhong, Z., and McDonald, B. A. (2016). An improved method for measuring quantitative resistance to the wheat pathogen zymoseptoria tritici using high-throughput automated image analysis. *Phytopathology* 106, 782–788. doi: 10.1094/PHYTO-01-16-0018-R

Stewart, E., Croll, D., Lendenmann, M. H., Sanchez-Vallet, A., Hartmann, F. E., Palma-Guerrero, J., et al. (2018). Quantitative trait locus mapping reveals complex genetic architecture of quantitative virulence in the wheat pathogen Zymoseptoria tritici. *Mol. Plant Pathol.* 19, 201–216. doi: 10.1111/mpp.12515

Strobl, C., Boulesteix, A.-L., Zeileis, A., and Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* 8, 25. doi: 10.1186/1471-2105-8-25

Stuckens, J., Swennen, R. L., Coppin, P., Dzikiti, S., Verreynne, S., and Verstraeten, W. W. (2011). Extracting physiological info from a hyperspectral time series of a citrus orchard. *Acta Hortic.* 11–18. doi: 10.17660/ActaHortic.2011.919.1

Thomas, S., Behmann, J., Steier, A., Kraska, T., Muller, O., Rascher, U., et al. (2018). Quantitative assessment of disease severity and rating of barley cultivars based on hyperspectral imaging in a non-invasive, automated phenotyping platform. *Plant Methods* 14, 45. doi: 10.1186/s13007-018-0313-8

Torriani, S. F. F., Melichar, J. P. E., Mills, C., Pain, N., Sierotzki, H., and Courbot, M. (2015). Zymoseptoria tritici: a major threat to wheat production, integrated approaches to control. *Fungal Genet. Biol.* 79, 8–12. doi: 10.1016/j.fgb.2015.04.010

Wahabzada, M., Mahlein, A.-K., Bauckhage, C., Steiner, U., Oerke, E.-C., and Kersting, K. (2016). Plant phenotyping using probabilistic topic models: uncovering the hyperspectral language of plants. *Sci. Rep.* 6, 22482. doi: 10.1038/srep22482

Wahabzada, M., Mahlein, A.-K., Bauckhage, C., Steiner, U., Oerke, E.-C., and Kersting, K. (2015). Metro maps of plant disease dynamics—automated mining of differences using hyperspectral images. *PLoS One* 10, e0116902. doi: 10.1371/journal.pone.0116902

Wakie, T. T., Kumar, S., Senay, G. B., Takele, A., and Lencho, A. (2016). Spatial prediction of wheat septoria leaf blotch (Septoria tritici) disease severity in Central Ethiopia. *Ecol. Inform.* 36, 15–30. doi: 10.1016/j.ecoinf.2016.09.003

Wright, M. N., and Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* 77, 1–17. doi: 10.18637/jss.v077.i01

Yang, C.-M. (2010). Assessment of the severity of bacterial leaf blight in rice using canopy hyperspectral reflectance. *Precision Agric.* 11, 61–81. doi: 10.1007/s11119-009-9122-4

Yu, K., Anderegg, J., Mikaberidze, A., Karisto, P., Mascher, F., McDonald, B. A., et al. (2018). Hyperspectral canopy sensing of wheat septoria tritici blotch disease. *Front. Plant Sci.* 9, 1195. doi: 10.3389/fpls.2018.01195

Zarco-Tejada, P. J., Berjón, A., López-Lozano, R., Miller, J. R., Martín, P., Cachorro, V., et al. (2005). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sens. Environ.* 99, 271–287. doi: 10.1016/j.rse.2005.09.002

Zarco-Tejada, P. J., González-Dugo, V., Williams, L. E., Suárez, L., Berni, J. A. J., Goldhamer, D., et al. (2013). A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. *Remote Sens. Environ.* 138, 38–50. doi: 10.1016/j.rse.2013.07.024

Zhan, J., Kema, G. H. J., Waalwijk, C., and McDonald, B. A. (2002). Distribution of mating type alleles in the wheat pathogen Mycosphaerella graminicola over spatial scales from lesions to continents. *Fungal Genet. Biol.* 36, 128–136. doi: 10.1016/S1087-1845(02)00013-0

Zhang, J., Pu, R., Huang, W., Yuan, L., Luo, J., and Wang, J. (2012). Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. *Field Crops Res.* 134, 165–174. doi: 10.1016/j.fcr.2012.05.011

Zheng, Q., Huang, W., Cui, X., Dong, Y., Shi, Y., Ma, H., et al. (2019). Identification of Wheat Yellow Rust Using Optimal Three-Band Spectral Indices in Different Growth Stages. *Sensors* 19, 35. doi: 10.3390/s19010035

Zhong, Z., Marcel, T. C., Hartmann, F. E., Ma, X., Plissonneau, C., Zala, M., et al. (2017). A small secreted protein in Zymoseptoria tritici is responsible for avirulence on wheat cultivars carrying the Stb6 resistance gene. *New Phytol.* 214, 619–631. doi: 10.1111/nph.14434

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2019 Anderegg, Hund, Karisto and Mikaberidze. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*