ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

TECHNICKÁ FAKULTA

ANALYSIS AND APPLICATION OF THE HYPERSPECTRAL REFLECTANCE TECHNIQUE FOR EARLY DETECTION OF FIELD CROP STRESS FACTORS

ANALÝZA A VYUŽITÍ TECHNIKY HYPERSPEKTRÁLNÍ ODRAZIVOSTI PRO VČASNOU DETEKCI STRESOVÝCH FAKTORŮ POLNÍCH PLODIN

Katedra zemědělských strojů

Disertační práce

WIKTOR RAFAŁ ŻELAZNY

2022

Prohlášení

Prohlašuji, že jsem tuto disertační práci vypracoval samostatně pod vedením školitele a uvedl jsem veškerou použitou literaturu. Tištěná a elektronická verze práce se doslovně shodují.

JMÉNO A PŘÍJMENÍ: Wiktor Żelazny

PODPIS:

DATUM: 13. listopadu 2022

Declaration

I declare to have prepared this dissertation thesis myself under the supervisor's supervision, and that I have stated all employed literature. The printed and the electronic versions of the work agree to the word.

NAME AND SURNAME: Wiktor Żelazny

SIGNATURE:

DATE: November 13, 2022

Poděkování

Rád bych poděkoval svému školiteli docentovi Milanovi Kroulíkovi a také spoluřešitelům: doktorom Janovi Lukášovi a Pavlovi Hamouzovi, inženýrce Janě Chrpové a inženýrovi Tomášovi Šimonovi.

Acknowledgements

I kindly thank my supervisor associate professor Milan Kroulík, and research partners: doctors Jan Lukáš, Jana Chrpová, Pavel Hamouz, and Tomáš Šimon.

Abstrakt

Včasná detekce stresů polních plodin vyžaduje sběr dat v mnoha časových bodech a zohlednění prostorové variability pole. Měření hyperspektrální odrazivosti lze provádět s velkým časoprostorovým rozlišením, což není praktický postup při použití tradičních, náročných a nákladných referenčních laboratorních metod. Tato dizertace průřezově zkoumá potenciál spektroskopického přístupu hodnocení stresových faktorů polních plodin. Je studován vliv sucha u dvou odrůd juvenilní řepky olejné s různými strategiemi hospodaření vodou, které jsou hyperspektrálně snímány z blízké vzdálenosti. Bylo pozorováno konzistentní zvětšení hodnot směrodatných odchylek mnoha vegetačních indexů v reakci na zhoršení vodního režimu. Byla snaha predikovat na základě spektrálního signálu intenzitu nákazy sbírky odrůd pšenice ozimé s rozmanitou rezistancí vůči fuzarióze klasů. Pokud byla přijatá racionální tolerance chyby, přesnost hodnocení se blížila k 100 %. Byla aplikována selekce příznaků za účelem minimalizovat počet hyperspektrálních pásem potřebných k odhadu obsahu listového chlorofylu u ovsa na základě hyperspektrálního zobrazování. Filtrování pásem zvětšilo kvalitu predikce a získaná redukce byla dostačující pro hypotetický vývoj cenově přístupného specializovaného snímače. Měření odrazivosti ve středním infračerveném spektrálním regionu je hodnoceno jako nástroj pro rychlou diagnozu obsahu a lability půdního uhlíku. Studie poskytuje důkazy o pozitivním vlivu spike sampling v situaci, kdy je potřeba predikovat vlastnosti půdního uhlíku ve vzorcích, pocházejících z lokality, která není zastoupena ve spektrální knihovně.

Klíčová slova: stres abiotický; stres biotický; hyperspektrální zobrazování; strojové učení; Bayesova statistika; Open Science

x

Abstract

Early crop stress detection requires data collection at multiple time points, while accounting for spatial variability of a field. Unlike resource-intensive traditional laboratory reference methods, hyperspectral reflectance measurements can be performed with high spatiotemporal frequency. This dissertation provides a cross-sectional exploration of the potential of the spectroscopic approach to assess crop stress factors. The influence of drought is studied in two juvenile oilseed rape cultivars with different water management strategies subjected to proximal hyperspectral imaging. A consistent increase in multiple vegetation index standard deviations to worsening of the hydric regime was observed. Prediction of fusarium head blight infection intensity is attempted from spectral response induced in a collection of winter wheat cultivars with varied disease resistance. With reasonable error tolerance, the rating accuracies approached 100 %. Feature selection is applied to minimize the number of bands needed to estimate oat leaf chlorophyll from airborne hyperspectral imagery. Filtering of the bands improved the prediction performance, and the obtained reduction was sufficient to envision the development of an affordable single-purpose imager. Mid-infrared diffuse reflectance spectroscopy is evaluated as a tool for rapid soil carbon content and lability diagnosis. The study provides evidence of positive influence of spike sampling while predicting soil carbon properties in a locality that is not represented in a spectral library.

Key words: abiotic stress; biotic stress; hyperspectral imaging; machine learning; Bayesian statistics; Open Science

Table of contents

List of abbreviations

CPLSR canonical partial least squares regression

ES effect size

FHB fusarium head blight

HS hyperspectral

HWC hot-water extractable carbon

ilr isometric log-ratio

LAI leaf area index

MIR mid-infrared

MIR-DRIFTS mid-infrared diffuse reflectance infrared Fourier transform spectroscopy

MRMR minimum redundancy, maximum relevance

MSC multiplicative scatter correction

NIR near-infrared

NRMSE normalized root mean square error

OC organic carbon

OSR oilseed rape

PC1 first principal component

PCA principal component analysis

PLS partial least squares

PLSR partial least squares regression

PLSR2 partial least squares regression 2

RBF radial basis function

RGB red, green, blue

RMSE root mean square error

RMSEP root mean square error of prediction

RPD ratio of prediction to deviation

RPIQ ratio of prediction to interquartile range

SC soil carbon

SNV standard normal variate

SPAD soil-plant analysis development

SVM support vector machine

SWIR short-wave infrared

TC total carbon

UAV unmanned aerial vehicle

VSS visual symptom score

VIS visible

visNIR visible and near-infrared

Chapter 1

Introduction

1.1 Crop stress and the challenge of its early detection

Excessive biotic and abiotic stress factors impair the crop yield and quality (Mahlein; Oerke, et al., 2012; Yang; Duan, et al., 2013), which brings negative consequences for profitability of the farming enterprise. The farmer can prevent or mitigate crop stress by fertilization, irrigation, or protection measures (Peteinatos et al., 2016). Advanced diagnostic methods and modeling techniques are crucial for the optimal timing and deciding on the intensity of these interventions (Stafford, 2000; West; Bravo, et al., 2003; Mahlein, 2016; Virnodkar et al., 2020). Effective control should, in particular, take into account spatial heterogeneity of the field (Ihuoma et al., 2017).

Early stress detection reduces the costs of its mitigation, and decreases the odds of permanent crop damage (Ihuoma et al., 2017; Lowe et al., 2017). On the other hand, it requires continuous monitoring performed at fine spatial scales (Mahlein; Oerke, et al., 2012; Gholizadeh; Kopačková, 2019). Such an undertaking is unfeasible in terms of labor and financial expenditure if traditional methods are involved. These include: visual assessment (Martinelli et al., 2015; Mahlein, 2016; Lowe et al., 2017), taking measurements with hand-held field instruments (Govender et al., 2009; Ihuoma et al., 2017; Virnodkar et al., 2020), and performing reference laboratory analyzes of collected plant tissue samples (Blackburn, 2007; Govender et al., 2009).

1.2 Plant reflectance spectra as a source of stress information

Stressed and healthy vegetation differ with respect to their spectral characteristics, and the symptoms of stress can be detectable before they become apparent for the naked eye (Knipling, 1970; Carter; Knapp, 2001; Kim et al., 2011; Zovko et al., 2019). Laboratory and field spectroscopy aimed at measuring of crop reflectance spectra can offer a more cost-effective alternative to the conventional methods (Mar-

tinelli et al., 2015). Among the reflective spectral regions for which dedicated equipment is readily available on the market, visible (VIS, 400–700 nm wavelength range) reflectance is affected primarily by plant pigments, near-infrared (NIR, 750–1000 nm) by mesophyll structure of the cells, and short-wave infrared (SWIR, 1000–2500 nm) — also by their water content (Jacquemoud et al., 2001; Govender et al., 2009; Ihuoma et al., 2017; Lowe et al., 2017; Mishra et al., 2017; Morin et al., 2017). However, according to Carter (1993), a SWIR response is detectable at a relatively late stage, making it less suitable for early stress detection than the visible and near-infrared (visNIR) region. This view was contested by Zovko et al. (2019).

An array of both biotic and abiotic stress factors alter the visNIR spectral profile of a crop in a similar way (Carter, 1993; Stafford, 2000; Carter; Knapp, 2001; West; Bravo, et al., 2003). Chloroplasts deteriorate in leaf tissues, leading to a decrease in the content and activity of chlorophyll (Jacquemoud et al., 2001; Din et al., 2011; Morin et al., 2017). On the other hand, anthocyanins can be synthesized and carotenoid concentrations increase to protect the photosynthetic apparatus from oxidative stress (Barton, 2001; Gaspar et al., 2002; Gill et al., 2010; Ashraf et al., 2013). These changes are detectable in the VIS spectrum, typically in the form of overall reflectance increase (Knipling, 1970; Martinelli et al., 2015; Morin et al., 2017), especially for longer wavelengths (Carter, 1993; Carter; Knapp, 2001).

As the red reflectance increases under the stress (Govender et al., 2009; Yang; Duan, et al., 2013), the red-edge, located at the boundary of the VIS and NIR region, shifts towards shorter wavelengths (Carter; Knapp, 2001; Govender et al., 2009; Martinelli et al., 2015). At the same time, after an initial increase, the NIR reflectance decreases due to structural changes related to cell shrinkage (Knipling, 1970; Behmann; Steinrücken, et al., 2014), making the red-edge less steep (Werff et al., 2008; Govender et al., 2009).

1.3 The value of hyperspectra in the context of early stress assessment

The magnitude of stress can be estimated using vegetation indexes, which are derived from small subsets of reflectance bands (Mahlein; Oerke, et al., 2012; Mulla, 2013; Xue et al., 2017; Virnodkar et al., 2020). These are typically measured using simple hand devices, on-the-go sensors mounted on agricultural machinery, or derived from multispectral airborne and satellite sensors (Perry; Davenport, 2007; Gnyp et al., 2015). This traditional approach can be insufficient for early stress detection (Römer et al., 2012). Accordingly, it has been complemented by methods

based on multivariate predictive models, which can exploit the information contained in a whole spectral region (Mulla, 2013; Amigo et al., 2015; Martinelli et al., 2015; Mishra et al., 2017).

A hyperspectrum offers a satisfactory approximation of a continuous spectral signature (Lu et al., 2019; Thenkabail et al., 2021). A number of studies compared full-spectrum and index-based approaches or predictive models developed using hyperspectral (HS) and multispectral data to estimate plant traits. Consistent improvements were noted (Hernandez et al., 2015; Marshall et al., 2015), even if slight (Capolupo et al., 2015; Lu et al., 2019) or restricted in magnitude in certain developmental phases or experimental conditions (Aguate et al., 2017; Montesinos-López et al., 2017). Under the traditional approach, high-resolution spectral data can foster estimation accuracy by enabling flexible formulation of custom vegetation indexes (Blackburn, 2007; Gnyp et al., 2015). The Mariotto et al. (2013) study on crop biophysical parameters is illustrative in this regard. The advantage of hyperspectra over multispectral data was recently summarized by Thenkabail et al. (2021).

1.4 The added value of spatial data provided by imaging spectroscopy

A hyperspectrum can be readily obtained by using a spectroradiometer, which integrates the reflectance captured in its field of view into a single measurement (Steiner; Bürling, et al., 2008; Milton et al., 2009; Behmann; Steinrücken, et al., 2014; Mahlein, 2016). However, the resulting spectra have limited spatial information (Mac Arthur et al., 2015; Adão et al., 2017) and resolution (Thomas; Kuska, et al., 2018). For precision agriculture, which involves differentiated interventions (Stafford, 2000; Zhang; Kovacs, 2012; Mulla, 2013), it is crucial to know the distribution of the crop status across the field (Bullock et al., 2000; Mahlein, 2016; Peteinatos et al., 2016). Because of that, there has been an increasing interest in applying HS imaging, also known as imaging spectroscopy, to crop stress detection (Adão et al., 2017; Lowe et al., 2017; Khan et al., 2018).

HS imaging overcomes limitations of traditional point spectrometry by offering minute georeferencing possibilities (Adão et al., 2017) and availability of segmentation methods (Mishra et al., 2017), for example for masking of background (Ge; Bai, et al., 2016) or identification of areas affected by unfavorable illumination effects (Asaari et al., 2019). When mounted on unmanned aerial vehicles (UAVs), visNIR HS cameras can provide readily available, high-quality data at the spatial scale of one

or more cultivated fields (Zhang; Kovacs, 2012; Gago et al., 2015; Sankaran et al., 2015). In proximal applications, individual plant organs can be discerned, and the resulting images enable fine distinctions — for example, between parts bearing severe disease symptoms and those less affected (Mahlein; Kuska, et al., 2017).

1.5   The role of soil spectroscopy in crop stress assessment

Crop health and productivity are closely linked to soil fertility. Just as certain aspects of soil quality can be deducted from vegetation health (Gholizadeh; Kopačková, 2019), occurrence of crop stress factors, drought for instance (Ihuoma et al., 2017; Virnodkar et al., 2020), can be indirectly predicted from the soil status.

Traditional methods of pedological laboratory evaluations are subject to similar limitations as for conventional plant stress assessment (Kuang et al., 2012). As an alternative, soils can be characterized with spectroscopic methods indoors and in the field (Ge; Thomasson, et al., 2011; Kuang et al., 2012; Nocita et al., 2015; Meer, 2018).

1.6   Impediments towards wider adoption of HS stress detection

Unlike numerous modern technologies that are commonly employed in precision agriculture — such as guidance systems enabled by satellite navigation, mapping of yield spatial variation, or variable rate technology (Stafford, 2000; West; Bravo, et al., 2003; Pedersen et al., 2017) — the potential of spectroscopy for site-specific farming is far from being fully exploited. Insufficient theoretical grounding is one contributing factor. In particular, the current number of findings linking stressors to crop characteristics captured by HS patterns is small compared to other techniques (Gago et al., 2015; Martinelli et al., 2015; Sankaran et al., 2015; Adão et al., 2017) — even in controlled conditions despite the long history of laboratory spectroscopy (Mac Arthur et al., 2015).

Another notable problem is the high purchase cost of both spectroradiometers (Milton et al., 2009) and HS cameras (Deery et al., 2014). However, the devices are predicted to become more affordable in the future (Pedersen et al., 2017). Adoption of spectroscopic stress assessment methods does not entirely eliminate the need for conventional data collection. Traditional measurements are still relied on, to the extent necessary for predictive model calibration and validation purposes (Govender et al., 2009; Morin et al., 2017; Virnodkar et al., 2020). The laboriousness

and cost of their acquisition can reduce the utility of spectroscopic methods for a farmer.

Finally, there is a challenge involved in analysis of spectral measurements, stemming from morphological, anatomical, and physiological differences between crops (Jacquemoud et al., 2001; Ihuoma et al., 2017), along with confounding external factors that cannot be controlled in field conditions (Milton et al., 2009), as well as instrumental noise generated by the acquisition device itself (Geladi et al., 2004). Expertise is crucial especially for canopy imaging, due to the variation in leaf area index, leaf inclination, and the effects of internal shadowing and light scattering (Knipling, 1970; Blackburn, 2007; Govender et al., 2009; Römer et al., 2012; Behmann; Steinrücken, et al., 2014). A separate set of specialized skills is required to deal with the large volumes of the captured data (West; Bravo, et al., 2003; Deery et al., 2014).

Chapter 2

Problem area and literature review

2.1 Choice criteria of the topics covered by the research

Early crop stress diagnosis from HS data is a broad subject. Therefore, a cross-sectional approach has been adopted for this dissertation.

The studied problems were selected to extend beyond the agronomy context. The additional themes comprise field phenotyping, precision farming, and soil quality monitoring. Another criterion was to include spectral datasets acquired using more than one type of a device and platform. Proximal and remote sensing are covered as well as use of imaging and non-imaging equipment in conditions corresponding to different degrees of control. Abiotic and biotic stressors are featured; both plants and soil are subjected to the measurements.

On a finer level, the topic choice aimed to focus on parameters that have been difficult to estimate using spectral methods. These include fusarium head blight (FHB) infection severity due to localized character of the symptoms on a plant, or soil lability due to low concentrations of labile soil carbon (SC) pools in samples. For crop stress indicators that are easier to assess, such as leaf chlorophyll content or soil total carbon, the focus was shifted towards reducing the costs of employing the existing solutions. Here, a possibility to develop an affordable narrow-band imager and to reduce the number of samples collected for traditional laboratory analysis are explored, respectively. Although, on par with current trends in research, full-spectrum methods are given preference in the studies, one experiment proposes a novel look at traditional vegetation indexes.

2.2 Quantification of spectra variability as a potential source of crop stress information

This section is adapted from Želazny; Lukáš (2020), previously published by MDPI Remote Sensing.

As highlighted by Kruschke et al. (2017), "stressors [...] can increase the variance of a group because not everyone responds the same way to the stressor". In the context of close-range crop HS imaging, the "group" can refer to plant foliage or leaf tissue composed of individual leaves and cells, respectively, each responding to the change in the environment in a distinct way. Especially characteristic for stress-induced leaf senescence is the source–sink differentiation between the older and younger leaves (Munné-Bosch et al., 2004). The potential of imaging spectrometry to provide an insight into the spatial variation of stress symptoms across crop foliage was demonstrated for drought (Nansen, 2012; Römer et al., 2012; Behmann; Steinrücken, et al., 2014; Bruning et al., 2019), nitrogen deficiency (Jay et al., 2014), pest infestation (Nansen, 2012), and herbicide exposure (Kong et al., 2016). However, this effect was typically not quantified.

Quantitative studies on crop responses to stress conditions frequently employ traditional experimental designs, such as a randomized block design, coupled with linear modeling for statistical inference (Tesfamariam et al., 2010; Majidi et al., 2015; Peteinatos et al., 2016). The frequentist approach prevails in the fitting and evaluation of these models. Various authors noted shortcomings of the frequentist statistics, and have advocated Bayesian methods as an alternative (Zyphur et al., 2015; Kruschke et al., 2017; Zyl, 2018). Historically, first the lack of and then the high computational demands of suitable numerical methods posed obstacles towards a wider adoption of the Bayesian paradigm (Che et al., 2010). These hindrances have been largely removed by an increase in computer speeds (Gelfand, 2000; Che et al., 2010), followed by improved accessibility of parallel computing (Visser et al., 2015), and the availability of software with capabilities suited to the needs of the scientific community (Che et al., 2010; Salvatier et al., 2016; Carpenter et al., 2017; Bürkner, 2018).

One major appeal of Bayesian statistics is the ease with which interval estimates of model parameters can be derived, even for complex models. Notably, it is possible to obtain estimates with respect to not only the mean values but also standard deviations, shape factors, or hurdle values — again, also for complex models (Kruschke et al., 2017). In the context of stress detection with imaging spectroscopy, this capability can be readily exploited to quantify the influence of a stressor on

the spectral variation across the foliage of an affected plant.

Given the continuing increases in average temperatures (Lobell et al., 2011) and projections of more frequent and severe droughts in agricultural regions (Naumann et al., 2018; Arnell et al., 2019), water deficiency has been among the most extensively studied crop stress factors (Daryanto et al., 2017). In pot experiments, crop responses to drought can be investigated by varying the watering regime and comparing the obtained plant reactions across the treatments (Linke et al., 2008; Behmann; Steinrücken, et al., 2014; Sun et al., 2018; Asaari et al., 2019). An alternative approach is to exploit the variability of water management strategies exhibited by individual genotypes (Gilbert; Zwieniecki, et al., 2011; Buezo et al., 2019).

Several dehydration avoidance mechanisms have been described in crops (Blum, 2005; Raza et al., 2017). Plants can rapidly respond to water deficit by closing their stomata, which reduces the leaf transpiration. As a trade-off, this reduction leads to a simultaneous decrease in the photosynthesis rate, related to limited CO_2 assimilation (Flexas et al., 2004; Ashraf et al., 2013). It has been proposed to differentiate crop cultivars with respect to their stomatal conductance regulation. Plants that manage their water resources in a conservative way and maintain a steady CO_2 fixation rate that is affected by moisture availability only to a limited extent have been termed as water-savers. Water-spenders, on the other hand, maximize their CO_2 assimilation, depleting the available water resources at the onset of a drought due to delayed closure of the stomata (Nakhforoosh et al., 2016; Urban et al., 2017). Cultivars with high baseline stomatal conductance tend to not exhibit a mid-day depression in photosynthetic rates. They are capable of sustaining a high photosynthesis rate and can avoid heat stress due to the cooling action of transpiration, provided that water is available (Roche, 2015).

In addition to physiological parameters, a trace of a drought episode can be detectable in a spectral signature of the affected crop, as was demonstrated by Linke et al. (2008) for wheat and by Sun et al. (2018) for maize. The authors tested the changes of several vegetation indexes in plants exposed to repeated drought and recovery cycles. They observed a full recovery after the first cycle, but the second recovery was incomplete. As a possible cause, the authors suspected progressing cell deterioration due to oxygen radicals, which could not be neutralized in the absence of carotenoids, removed in the course of the preceding stress episode.

Studies devoted to drought effects on crop hyperspectra have been primarily focused on the species that dominate the global commodity market. Those include maize (Ge; Bai, et al., 2016; Asaari et al., 2019) and other staple cereals (Römer et al., 2012; Behmann; Steinrücken, et al., 2014; Bruning et al., 2019). Relatively

much attention has also been given to fruit crops (Kim et al., 2011; Zarco-Tejada et al., 2012; Zovko et al., 2019). On the other hand, numerous other species have so far been largely neglected by the studies, including those of regional importance.

Due to its nutritional (Din et al., 2011; Xia et al., 2018; Sabagh et al., 2019) and technical (Högy et al., 2010; Xia et al., 2018) value, oilseed rape (*Brassica napus* L.; hereafter, OSR) is an important crop in many parts of the world. It is widespread in North America (Zhang; Lu, et al., 2014; Bonjean et al., 2016), China (Bonjean et al., 2016), Europe (Zhang; Lu, et al., 2014), and India (Kumar et al., 2017). OSR is susceptible to drought (Din et al., 2011; Raza et al., 2017) and, along with other brassicas, the future cultivation of this species is endangered by dry spells (Zhang; Lu, et al., 2014; Majidi et al., 2015).

OSR has been the subject of various HS imaging studies. Based on field experiments, Piekarczyk et al. (2011) and Zhang; He (2013) attempted to predict its yield using vegetation indexes and partial least squares (PLS) modeling, respectively. Kumar et al. (2017) cite several publications devoted to OSR pests and diseases. Xia et al. (2018) analyzed imagery of water-logged plants. Effects of herbicide exposure were studied by Kong et al. (2016). In contrast to these stress factors, the possibilities of capturing the OSR response to drought using a HS camera remained unaddressed.

OSR is characterized by relatively large leaves, even in early developmental phases. HS imaging that captures leaf-level spectral variation may, therefore, prove to be a suitable approach for water deficiency detection in this crop (Bruning et al., 2019).

2.3 Discrete crop stress severity predictions using ordinal SVM trained to hyperspectra

This section is adapted from Żelazny; Chrpová, et al. (2021), previously published by Elsevier Biosystems Engineering.

Data associated with ordinal measurements scales require careful statistical treatment. The family of cumulative link models was developed for use in linear modeling with this class of data (McCullagh, 1980; Fernández-Navarro, 2017), and Liddell et al. (2018) demonstrated that false conclusions can be obtained when an inappropriate model family is employed. In the realm of predictive modeling, naive treatment of ordinal data can increase computational costs of model training (Behmann; Schmitter, et al., 2014), and ordinal extensions of machine learning algorithms have been proposed (Fernández-Navarro, 2017). The Behmann; Schmitter, et al. (2014) and Behmann; Steinrücken, et al. (2014) drought studies comprised early application

of ordinal machine learning to plant stress evaluation based on HS images.

Ordinal measurement scales are commonly employed in field phenotyping of cultivar disease resistance, where an expert visually rates disease severity in breeding lines (Bock et al., 2010). This traditional approach is a laborious and subjective process (Bauriegel; Herppich, 2014; Deery et al., 2014; Mahlein, 2016; Su et al., 2021), and it constrains progress in breeding programs (McMullen et al., 2012; Steiner; Buerstmayr, et al., 2017). Spectroscopic methods can be superior to visual rating by providing rapid and unbiased assessment in early infection stages (West; Canning, et al., 2017; Thomas; Kuska, et al., 2018).

FHB is a cereal fungal disease caused by *Fusarium* spp. (Jaillais et al., 2015; Khaledi et al., 2017; Saccon et al., 2017; Mielniczuk et al., 2020). The pathogen severely impairs yield, and the grain quality is greatly diminished by mycotoxin action (Bauriegel; Herppich, 2014; Cambaza et al., 2019; Mielniczuk et al., 2020). Infection occurrences have been increasing because of the high prevalence of maize in crop rotations (Dammer et al., 2011; Bauriegel; Herppich, 2014) and due to the adoption of reduced tillage systems by farmers (McMullen et al., 2012; Gilbert; Haber, 2013; Mielniczuk et al., 2020). Climate change may also aggravate this problem in the future (Gilbert; Haber, 2013; Vaughan et al., 2016).

FHB control is based on cultural preventive measures and fungicide applications (McMullen et al., 2012; Mielniczuk et al., 2020), the latter of which are scheduled according to the disease risk levels predicted from weather forecasts (McMullen et al., 2012; Xiao et al., 2020). There is also ongoing research into biological control agents (McMullen et al., 2012; Mielniczuk et al., 2020). However, the progress in FHB control does not remove a need for the development of new resistant cultivars (Gilbert; Haber, 2013; Buerstmayr et al., 2020; Mielniczuk et al., 2020), which need to be screened in phenotyping trials.

The primary focus of FHB spectroscopy studies has been post-harvest mycotoxin detection in extracted kernels under laboratory conditions. The research has been reviewed by Saccon et al. (2017) and Femenias et al. (2020), and findings on this topic continue to be published (Zhang; Chen; Zhang, et al., 2020; Zhang; Wang; Lin; Weng, et al., 2020; Shen et al., 2022). Less attention, however, has been given to whole spikes (Bauriegel; Giebel, et al., 2011; Alisaac et al., 2018; Whetton; Hassall, et al., 2018; Huang; Wu, et al., 2019; Mahlein; Alisaac, et al., 2019; Huang; Li, et al., 2020; Zhang; Wang; Lin; Yin, et al., 2020).

Field research has also been limited until recent years, as highlighted by Whetton; Waine, et al. (2018) and Huang; Wu, et al. (2019). Whetton; Waine, et al. (2018) employed a proximal push-broom HS imager coupled with an artificial light source

11

for FHB occurrence density estimation in winter wheat and barley. Söderström et al. (2013) predicted the deoxynivalenol Fusarium mycotoxin in oats from data captured by three different sensors, including a multispectral satellite imager. Further research included the Liu; Dong; Huang; Du; Ren, et al. (2020) proposal of a novel broad-band vegetation index to detect the disease from Sentinel-2 multispectral imagery. Xiao et al. (2020) augmented a time series of satellite observations with meteorological data for the same purpose, and Liu; Dong; Huang; Du; Ma (2020) investigated the feasibility of FHB monitoring at the field scale using UAV equipped with HS camera.

Whetton; Waine, et al. (2018) considered that the small number of FHB spectroscopic field studies may stem from the difficult detection of the disease in a crop stand. Unlike well-studied rust and powdery mildew (Franke et al., 2007; Huang; Lamb, et al., 2007; Kuckenberg et al., 2009; Cao et al., 2013; Huang; Guan, et al., 2014; Feng et al., 2016; Zhang; Wang; Yuan, et al., 2017), FHB symptoms affect only the spikes, and these constitute a small fraction of the total biomass. More recently, Liu; Dong; Huang; Du; Ma (2020) proposed the identification of infected field areas based on whole canopy characteristics, but the feasibility of this novel and potentially controversial approach requires further scrutiny.

There is a limited transfer ability of spectrometric disease detection methods developed for large-scale farming applications to field phenotyping. Simultaneous screening of multiple genotypes constrains the size of experimental plots to a level below the spatial resolution attainable with many remote or even proximal sensing systems (Rebetzke et al., 2014; Barmeier et al., 2016). Artificial inoculation may be limited to a small subset of plants in each plot, thus further limiting the choice of spectral data acquisition techniques. The plot size reduction can also modify its spectral properties (Barmeier et al., 2016), and this invalidates models which assume a continuous canopy. Moreover, disease severity must be estimated when searching for tolerant cultivars, rather than mere disease occurrence (Qiu et al., 2019). This task has only recently been attempted (Bock et al., 2010; Huang; Wu, et al., 2019; Qiu et al., 2019; Zhang; Wang; Gu, et al., 2019). Moreover, the predictions must consider masking the disease signal by genotype differences (Pinter et al., 1985). All these challenges create the need for research programs specifically aimed at FHB field phenotyping.

2.4 Specialized narrow-band imagers as an alternative for HS cameras in crop stress detection

This section is adapted from Żelazny (2020), previously published by Estonian Agricultural University, Faculty of Agronomy Agronomy Research.

Development of simplified narrow-band imagers has been proposed as a way to overcome the problem of high HS camera prices. These products would resemble multispectral devices by their limited number of bands, but would be suited to specialized applications, including stress mapping (Govender et al., 2009; Deery et al., 2014; Mahlein, 2016; Lowe et al., 2017; Alckmin et al., 2020). In addition to the lower price (West; Bravo, et al., 2003; Amigo et al., 2015), they would offer a high speed of operation (Rapaport et al., 2015).

The bands for use in the envisioned devices can preselected using feature selection algorithms (Govender et al., 2009; Mishra et al., 2017). These methods remove wavelengths associated with redundant information (Mewes et al., 2011; Behmann; Steinrücken, et al., 2014). Therefore, the predictive performance of the device would remain at a satisfactory level. As demonstrated by the Zhang; He (2013) OSR yield study, substantial reduction of data volume can be attained without impairing model performance. Discarding of 98 % of HS bands had a minimal effect on the quality of nitrogen content prediction in pepper plants, while significantly simplifying the obtained model (Yu et al., 2014). Feature selection was even shown to improve the prediction accuracy in some applications (Ding et al., 2005; Mehmood et al., 2012). Behmann; Steinrücken, et al. (2014) proposed a support vector machine (SVM) model for detecting water stress in barley. The model inputs comprised vegetation indexes, the combinations of which were determined using wrapper feature selection. Increased detection sensitivity was obtained, allowing for earlier drought detection relative to the raw indexes. Owing to the reduced number of inputs, fewer computations are required to process data subjected to feature selection, and model interpretation is facilitated (Ding et al., 2005).

Mehmood et al. (2012) reviewed current feature selection methods suitable for PLS modeling. Under the filtering approach, variables are evaluated independently of model fitting, according to a measure the value of which determines which of them will be discarded. In the minimum redundancy, maximum relevance (MRMR) method, this measure is the mutual information shared by the candidate feature and the predicted variable, reduced by the average mutual information shared by the candidate feature and the features already accepted for inclusion into the model. This mutual information is a function of the

correlation coefficient (De Jay et al., 2013). With wrapping, models are fitted to multiple preselected feature subsets, and the fit quality itself serves as the selection performance criterion, making it a computationally more demanding approach (Mehmood et al., 2012). The wrapper forward selection method is analogous to the forward selection in the stepwise regression: candidate features are picked one by one from the feature pool, and their influence on the performance of the refitted model is assessed. The variable associated with the highest performance increase is kept in the model, and the process continues iteratively until there is no further improvement.

The indispensability of chlorophyll for plant photosynthesis (Sims et al., 2002; Main et al., 2011) and its contribution to crop optical properties (Ollinger, 2011) make the estimation of leaf chlorophyll concentration an important remote sensing application. In large-scale assessments, leaf chlorophyll remote sensing is useful for yield prediction (Moharana et al., 2016). At finer spatial scales, it can be used for the delineation of management zones for precision agriculture (Miao et al., 2009). As chlorophyll breaks down under stress, its monitoring provides information about the crop status, and enables a timely intervention to prevent the yield loss (Peñuelas; Baret, et al., 1995; Sims et al., 2002).

Numerous studies have been devoted to predicting chlorophyll content from HS images. Partial least squares regression (PLSR) was employed to analyze winter wheat leaf laboratory samples (Zhang; Wang; Ma, et al., 2012). Scanning of single leaves under controlled illumination allowed the authors to evade the challenges inherent to canopy-level imaging in outdoor conditions, and without doubt contributed to extremely accurate (R statistics up to 0.99) predictions. Unfortunately, unclear study design description undermines the trustworthiness of the findings. Meij et al. (2017) employed PLS to predict chlorophyll content in oats from UAV campaign data. Robust leaf chlorophyll content predictions for multiple crops were obtained with kernel ridge and Gaussian process regression. On the other hand, artificial neural networks, an approach with a comparable level of sophistication, failed to provide consistently reliable estimates (Caicedo et al., 2014). By applying SVM to maize hyperspectra, Karimi et al. (2008) obtained very good validation estimates for the tasseling stage. The prediction quality was worse, but still satisfactory, for the early growth stage, which the authors attributed to the soil showing through the crop canopy.

Possibilities of extending these workflows with feature selection to simplify the imagers, obtain parsimonious models, and potentially improve their accuracy remain unexplored. Reliable non-imaging solutions available on the market, such as

chlorophyll meters (Govender et al., 2009; Miao et al., 2009) working with reduced numbers of spectra, justify undertaking research in that direction.

2.5 Soil lability link to crop stress and its estimation using spectroscopic methods

This section is adapted from Żelazny; Šimon (2022), previously published by MDPI Agriculture.

SC is a primary indicator of soil quality (Reeves, 1997; Bünemann et al., 2018), and in recent years estimation of atmospheric CO_2 sequestration has boosted interest in SC monitoring (Madari et al., 2005; Stenberg et al., 2010; Batjes et al., 2015; Paustian et al., 2019; Smith; Soussana, et al., 2019). In addition to SC quantity, its fractional composition can be of interest in evaluating soil status. Research has been devoted to the labile fraction, which can give insight into SC turnover processes (Kan et al., 2021). Labile C determines the rate of nitrogen release from soil organic matter, a factor to be accounted for while fertilizing the soil (Körschens et al., 1990; Thomas; Whalen, et al., 2016), and it can also inform about long-term stability of sequestered carbon (Page et al., 2013). Traditional assessment of SC lability is costly and time-consuming due to the laboriousness of laboratory SC fractionation (Zimmermann et al., 2007; Yang; Xie, et al., 2012; Jaconi et al., 2019). Environmental concerns have also been raised (Janik; Skjemstad, et al., 2007; Gredilla et al., 2016).

Higher throughput and economical viability can be attained with soil spectroscopy (Viscarra Rossel et al., 2006; Ge; Thomasson, et al., 2011; Paustian et al., 2019; Barra et al., 2021). Here, mid-infrared diffuse reflectance infrared Fourier transform spectroscopy (MIR-DRIFTS) is one of the methods considered suitable for chemical soil analysis (Viscarra Rossel et al., 2006; Janik; Skjemstad, et al., 2007; Soriano-Disla et al., 2014) owing to fundamental vibrations of soil molecules arising in the MIR spectral region (McCarty et al., 2002; Du et al., 2009; Kuang et al., 2012; Soriano-Disla et al., 2014). In particular, it can give accurate estimates of SC content (Viscarra Rossel et al., 2006; Reeves III, 2010; Bellon-Maurel et al., 2011; Kuang et al., 2012; Soriano-Disla et al., 2014; Barra et al., 2021), and according to Reeves III (2010), this high performance may extend to SC fraction assessments. However, the modest number of publications devoted to SC lability (Zhang; Yang, et al., 2018) is in contrast with the extensive literature on total carbon (TC) or the large organic carbon (OC) pool estimation with MIR-DRIFTS.

2.6 Calibration spiking as a possibility to reduce the number of laboratory samples in spectroscopic crop stress evaluations

This section is adapted from Żelazny; Šimon (2022), previously published by MDPI Agriculture.

Spectroscopic methods do not completely remove the need for traditional measurements, as predictive model training, tuning, and validation depend on their availability. An approach termed "calibration spiking" was proposed as a way to maximize reuse of reference data collected in the past while avoiding calibration domain mismatch when a model is applied in a new setting (Stenberg et al., 2010; Cezar et al., 2019). It consists of extending the reference library with a limited number of samples collected at the target site, and recalibrating the predictive model to the resulting dataset (Capron et al., 2005). To minimize their number, samples for calibration spiking can be picked according to leverage selection (Nocita et al., 2015). Under leverage selection, an extended number of samples is subjected to spectral analysis, after which a subsample is picked to be analyzed in the traditional way, based on their spectral dissimilarity. In this way, the resulting library size can be reduced while preserving the representativeness of the samples therein with respect to the spectral variation of the soils occurring in the target region (Capron et al., 2005).

However, even with a modestly-sized reference dataset, an issue of the disproportion between the number of library and spiking samples arises. One way of addressing this problem is to use a subset of the library samples (Guerrero; Zornoza, et al., 2010). As an alternative, which does not involve information loss, local samples can be given bigger weight relative to the samples in the library (Cezar et al., 2019). The weighing is typically performed by multiplying the local sample occurrences in the model training dataset (Guerrero; Stenberg, et al., 2014; Cezar et al., 2019). As an alternative approach, a model allowing for observation weights to be included among its inputs is employed, instead (e.g., Sankey et al., 2008). Indahl et al. (2009) proposed combining PLS with canonical correlation analysis, and termed the new class of models canonical PLS. Among the features of canonical partial least squares regression (CPLSR) that are absent in PLSR is the possibility to weight the individual observations. However, the suitability of CPLSR to calibration spiking has not received research attention.

2.7 Computational experiments in crop stress detection from hyperspectra

The multivariate character of HS data with autocorrelated bands (Blackburn, 2007) frequently demands an application of pre-processing methods to strengthen the stressor signal while suppressing confounding factors, instrumental noise, and other artifacts (Gholizadeh; Borůvka, et al., 2013). Rinnan et al. (2009) provide a comprehensive review of available transformations, which they classify into two major groups. An even bigger array of possibilities pertains to the choice of a predictive model family that can be trained to the spectral data. These range from traditional chemometrics models (Frank et al., 1993), through machine learning algorithms (Kuhn et al., 2013), to the latest deep learning approaches, which have been proliferating as a byproduct of artificial intelligence research (Gao; Luo, et al., 2020). Availability of leverage sampling (Gani et al., 2016), feature selection, and feature engineering (Guyon et al., 2003) methods further compounds the number of data treatment possibilities.

This "garden of forking paths" problem creates a need for computational experiments that compare the performance of data processing workflows for individual applications based on data acquired in various settings. Only then a choice from the multitude of available options can be made to develop a commercial product that will be reliable in a broad range of circumstances and can be offered to a mainstream consumer. Workers specializing in crop stress sensing have recognized this need, and such assessments make a significant share of current research (e.g., Buddenbaum et al., 2012; Behmann; Schmitter, et al., 2014; Cezar et al., 2019; Huang; Wu, et al., 2019; Liu; Dong; Huang; Du; Ma, 2020; Zhang; Chen; Yin, et al., 2020). Naive comparisons of single models prevail, and only a minority of authors (Verrelst et al., 2012; Ng et al., 2018; Baumann et al., 2021) evaluate model ensembles developed with multiple versions of a training dataset. Application of formal methods to quantify performance differences between individual approaches is also unexplored.

17

Chapter 3

Hypotheses and study strategy

In a proximal-sensing HS imaging study with potted OSR plants, it was hypothesized that the variability of responses to drought would provide more reliable separation between the experimental treatments than the mean values of spectral indicators. In a field phenotyping trial involving both point and imaging spectrometry, prediction quality sufficient for genotype screening was hypothesized for models trained to FHB-infected spike spectra acquired at a late stage of the disease. Prediction performance deterioration was expected in the case of the spectral data obtained at an earlier date. An implicit hypothesis drove a study on leaf chlorophyll content estimation from remote HS imagery. Two feature selection methods representing distinctive algorithm families were expected to reduce the number of bands to a number typical for multispectral devices without impairing the prediction accuracy.

One study was devoted to mid-infrared (MIR) soil spectroscopy. Improved predictions of SC and its lability were expected with calibration spiking of MIR-DRIFTS spectra, as described in the literature. Furthermore, superiority of embedded weighing of PLSR spiking observations to the traditional weighing scheme, under which multiple copies thereof are included in a training dataset, was hypothesized. Simultaneously, it was assumed that the number of samples subjected to traditional laboratory analysis could be reduced by proper pre-processing of the spectra or by employing leverage sampling.

A balance between examination of the questions that are relevant to the current state of scientific inquiry and novelty is sought in the dissertation thesis. The novel elements include the following: Not only global commodity crops, but also minor and less researched species with regional significance, are featured. In two of the three crop sensing studies, more than one cultivar is employed to generalize the validity of the findings and examine their robustness. The conventional "drought" and "control" watering regimes are extended with a "rewatered" regime in the OSR study to find out whether a trace of the drought episode will remain in the spectral patterns. Field phenotyping and a more challenging disease are the focus of the FHB study

instead of leaf diseases studied in the context of precision agriculture, which have been prevalent in crop disease spectroscopy. The studies exploit an opportunity to compare multiple spectral data pre-processing or processing schemes. Less known predictive model families are tested, such as ordinal SVM model ensembles or PLSR extensions that generalize the dependent variable to a vector of variables. Multiple data partitioning scenarios are generated for more robust assessment of the workflow performances, and in one study, these are compared in a formal fashion using linear models. To this end, Bayesian modeling is attempted instead of the dominant but criticized frequentist statistics. Not only SC, but also its lability is investigated in the soil spectroscopy study, and the data are analyzed in the compositional data analysis framework.

The collected data and scripts were made available in public repositories to enable the reproduction of the results and adaptation to future research (Sandve et al., 2013; Piccolo et al., 2016; Szucs et al., 2017). One of the studies was preregistered (Nosek et al., 2018) to address the "researcher degrees of freedom" (Bakker et al., 2020) and "file drawer" (Rosenthal, 1979) problems.

In addition to the published material, multiple further HS data acquisition campaigns were performed, but the obtained spectra were of insufficient quality to proceed with their analysis. A number of the encountered problems pertained to the radiometric dimension of the data, for example, improper artificial lighting introduced artifacts into proximal sensing imagery of juvenile plants exposed to low temperatures. Two attempts were made to capture airborne imagery of a commercial field with variable sowing density from a UAV, one unsuccessful due to technical failure resulting in an emergency landing, and the other due unfavorable illumination conditions and image pixel saturation. Some of the issues were more related to the geometric dimension of a dataset, as described for the HS camera data in the field phenotyping study, for instance. The same experiment was captured from a UAV, but since the infection did not spread to the untreated neighboring spikes, the affected areas could not be discerned in the data product.

Chapter 4

Dissertation objectives

The presented dissertation is aimed at filling some of the gaps with regard to early plant stress diagnosis. It investigates the underexplored relationships between physiological status, the degree of tissue damage, and field crop HS characteristics. Apart from plant material, studies on soil carbon are included in recognition of the contribution of soil properties to crop stress. The secondary goal is to introduce and popularize in plant sciences statistical methods, data analysis tools and workflows that to a large degree have been absent in this field — in crop spectroscopy, in particular. Advantages of the most promising of these methods are demonstrated. Finally, the dissertation responds to the reproducibility crisis and it contributes to the Open Science movement.

The influence of drought on narrow-band vegetation indexes and principal component analysis (PCA) scores is studied in juvenile OSR, based on proximal imagery acquired in semi-controlled conditions (Żelazny; Lukáš, 2020). The objective is to determine the spectral response of plants representing two water management strategies to three types of watering regimes. In addition, the effect of spectral pre-processing is evaluated in the PCA part of the study.

The spectral response of a collection of winter wheat cultivars, with varied FHB resistance, is investigated after the exposure of the plants to *Fusarium culmorum* (Żelazny; Chrpová, et al., 2021). The aims of the study are: (1) to estimate the infection severity prediction accuracy with respect to visual rating by an expert, and (2) to examine the accuracies with respect to the timing of the inoculation and spectral data acquisition. Two approaches to spectra aggregation are also compared and, like in the preceding study, several spectral pre-processing schemes are tested.

A possibility of choosing a small set of hyperspectra for the purpose of developing a chlorophyll meter with imaging capability is explored based on an analysis of publicly available aggregated HS data (Żelazny, 2020). The aim is to first reproduce selected results from the original study that was based on that dataset, and then investigate the effect of two feature selection approaches on the prediction of leaf

chlorophyll concentration in oats.

The feasibility of MIR-DRIFTS for quick diagnosis of SC properties is investigated (Żelazny; Šimon, 2022). The study compares SC content and SC lability estimates obtained using MIR-DRIFTS spectroscopy followed by three levels of spectral pre-processing to those derived with reference laboratory methods. The second study objective is to determine the influence of calibration spiking and its intensity on the performance of the soil carbon predictions. Internally and externally weighted CPLSR models trained to MIR-DRIFTS spectra are also compared, and the performance of two leverage sampling algorithms is evaluated. Long-term experiments serve as the reference library, and two commercial sites — as the prediction targets and the sources of the spiking samples.

Chapter 5

Research methods

5.1 Drought stress detection in juvenile oilseed rape using HS imaging with a focus on spectra variability

This section is adapted from Żelazny; Lukáš (2020), previously published by MDPI Remote Sensing.

The seed of 'Cadeli' and 'Viking' OSR cultivars was obtained from OSEVA PRO s.r.o. (Opava, Czech Republic) and started by placing it for two days in a thermostat (BT-120, Laboratorní přístroje Praha, Praha, Czechoslovakia), set to 20 °C. The seedlings were transplanted to pots, at the number of five seedlings per pot, and grown in a growth chamber, model Tyler T-16/4 (Budapest, Hungary), in 18–20 °C, under 16-hour photo period, exposed to $400\,\mu\mathrm{mol\,m^{-2}\,s^{-1}}$ irradiance.

The watering regime followed one of the three treatments: Control pots were watered daily to 70 % of the substrate water capacity. For pots in the dry treatment, after 14 days since sowing, the watering was reduced to 45 % of the water capacity, and after 10 more days the watering stopped completely. The pots in the rewatered treatment were treated according to the same plan, but one day before the experiment termination they were watered to 100 % of the water capacity.

Images of the plants in the 3–4 leaves growth phase were captured using a 2D frame HS camera (Rikola, Senop, Oulu, Finland) in natural light conditions. Four pots — two pots per cultivar, placed in a photo tent against a dark background were captured in each image. The HS data cubes comprised 41 evenly spaced bands from 503 to 903 nm. The spatial resolution was 1010 px × 1010 px, and the integration time was set to 30 ms. The pot rims were approximately 0.70 m away from the camera lens, and an irradiance sensor was also placed inside the tent.

The images of the plants were interleaved with images of spectralon. For every band, a mixed-effect empirical line model (Smith; Milton, 1999) was fitted, and reflectance derived from radiance values. The images were then subjected to band

registration, and square regions of interest were inscribed inside the pots.

A sample of 200 px was drawn from all regions of interest. Each was hand-classified as either background, fresh-leaf, dry-leaf, or edge pixel. The spectra of the pixels were partitioned into training and test dataset, in 3:1 proportion. A SVM model was fitted, tuned, and applied. The effects of the experimental treatments on the dry-leaf pixel proportion were assessed using a Bayesian linear mixed-effect model.

Twenty vegetation indexes (SR, GI, RGI, DVI, NDVI, RDVI, PSRI, $PSSR_a$, $PSND_a$, RNDVI, PRI_{570}, PRI_{512}, PRI_{norm}, MTCI, MCARI, TCARI, OSAVI, TCARI/OSAVI, CI_{green}, and CI_{re}) were calculated for each leaf pixel except for background and edge pixels. From each pot image, 36 pixels were sampled on a regular grid. To assess the influence of the experimental treatments on the index values, an ensemble of Bayesian linear mixed-effect models were fitted.

An attempt to exploit the full-spectrum information was made. In addition to analyzing raw spectra, Savitzky–Golay filter, multiple scatter correction, derivation, and second derivation pre-processing was tested. The spectra were subjected to principal component analysis, and pixel subsets were grid-sampled analogously to vegetation indexes. The influence of the experimental treatments on PCA scores was then assessed using multivariate linear modeling, with the model outcomes consisting of the first four score values.

5.2 Fusarium head blight detection from spectral measurements in a field phenotyping setting

This section is adapted from Żelazny; Chrpová, et al. (2021), previously published by Elsevier Biosystems Engineering.

Winter wheat was sown on 1–3 October 2019 in 12 parallel blocks, with two rows of aligned hill plots per block. Each plot pair represented one cultivar, arranged according to the earliness in the first three blocks and the remainder in alphabetic order. These cultivar arrangements were replicated three times. In each block, one of the rows was inoculated with the isolate B of *Fusarium culmorum* in the winter wheat flowering stage. Suspension of 0.8×10^7 spores ml^{-1} was applied from all sides to 10 spikes of each hill plot, tied together to form a cluster. Those were covered for 24 hours with 40 cm × 60 cm polyethylene bags. Simultaneously, 10 spikes were tied and covered with a bag also in the facing hill plots. The plots were maintained under irrigation except for rainy and subsequent days.

Two spectral data acquisition campaigns were planned: (1) soon after the ini-

tial infection symptoms become visible to a naked eye and (2) once the symptoms become severe. These dates coincided with the milk- and dough-ripening developmental phases of the plants. The data were collected using the ASD Fieldspec 4 Hi-Res (Malvern Panalytical, Malvern, UK) spectroradiometer equipped with a contact probe. Each spike cluster was subjected to a sequence of five measurements, with three spikes were positioned between the probe and a black non-woven textile. The device was calibrated at the beginning of each row and then roughly every ten minutes using white Spectralon. The resulting dataset comprised complete spectral data for 67 hill-plot pairs. Collection of imaging data was also attempted with the same HS camera as in the preceding study, mounted on a tripod. The acquisition process was slow due to difficult aiming and waiting for proper light conditions, and the imagery had to be discarded.

On the same days, visual symptom scores (VSSs) were assigned to each hill plot of the infected rows according to a 9-point scale denoting the percentage of the infected spikelets in the spike cluster (9 points: $<5\%$, 8: 5–17 %, 7: 18–30 %, 6: 31–43 %, 5: 44–56 %, 4: 57–69 %, 3: 70–82 %, 2: 83–95 %, 1: $>95\%$). In the case of the dough-ripening phase, the facing hill plots were also rated one day later, and the hill plot pairs for which that rating was below 9 were excluded from further analysis. The final number of retained hill plot pairs was 48.

The hill plots were 10 times randomly split into calibration and validation data partitions, in the 3:1 proportions. For each combination of partitioning, preprocessing, hill plot, and spectral band, the median and mean reflectances across multiple measurements or pixels were derived. Seven scenarios of aggregated spectra pre-processing were tested: (1) no pre-processing, (2) subtraction of facing hill plot spectrum, (3) division by facing hill plot spectrum, (4) standard normal variate (SNV), (5) subtraction followed by SNV, (6) division followed by SNV, and (7) maximum normalization followed by Savitzky–Golay derivative and smoothing to approximate the approach of Whetton; Waine, et al. (2018).

Ordinal SVMs with the radial basis function kernel (Behmann; Steinrücken, et al., 2014) were trained for each combination of spectra acquisition campaign, spectra aggregation function, plot partitioning, spectra pre-processing, and VSS dataset — 560 model ensembles in total. The tuning hyperparameters (C and gamma) were determined with an aid of Bayesian optimization according to the leave-one-out cross-validation maximum accuracy criterion. Validation accuracies of each model were estimated; allowing for misclassification of zero, one, and two class differences. The joint effects of spectra acquisition campaign and ground-truth data collection timings, spectra pre-processing scenarios, and magnitudes of error tolerances

on the probability of a correct VSS rating were modeled using mixed-effect Bayesian generalized linear models (Bürkner, 2018).

5.3 Application of feature selection for predicting leaf chlorophyll content in oats from HS imagery

This section is adapted from Żelazny (2020), previously published by Estonian Agricultural University, Faculty of Agronomy Agronomy Research.

The Meij et al. (2018) HS dataset was obtained, and reproduction of the Meij et al. (2017) PLS modeling result was prepared. Soil-plant analysis development (SPAD) leaf chlorophyll concentration was modeled as the dependent variable, and the reflectance values as the independent variables. The number of latent variables was tuned using leave-one-out cross-validation by calculating the cross-validation root mean square error (RMSE) for each value from between 1 and 20.

Next, two approaches to feature selection were tested: a filter method based on the MRMR criterion, and a forward selection wrapper method. For each, the study aimed to obtain a series of models with the input feature number ranging from 2 to all 94 bands (i.e., no selection). In this way, the influence of feature selection intensity on the prediction quality could be investigated. Validation RMSE, normalized root mean square error (NRMSE), and R^2 statistics were derived.

5.4 Calibration spiking of MIR-DRIFTS soil spectra for carbon predictions using CPLSR and log-ratio transformations

This section is adapted from Żelazny; Šimon (2022), previously published by MDPI Agriculture.

Soil samples were collected at the territory of the Czech Republic: (1) time series of archived samples obtained from long-term agricultural trials, and (2) topsoil samples from two commercial sites, a field in Janovice (45 samples) and Ústí nad Orlicí district (335 samples). The samples were dried, sieved through 2-mm mesh, and milled. MIR-DRIFTS spectra were measured using a Thermo Nicolet Avatar 320 FTIR spectrometer with a Ge beam splitter and a TGS detector, equipped with a Smart Diffuse Reflectance accessory (Nicolet, Madison, USA) in a 1:3 mixture of soil and KBr (Sigma-Aldrich, Darmstadt, Germany) prepared by hand in an agate mortar. TC content was determined using Vario/CNS analyzer (Elementar Analysensys-

teme GmbH, Langenselbold, Germany), and hot-water extractable carbon (HWC) content was determined according to Körschens et al. (1990).

For each commercial site, 10 independent sets of 12 samples were picked randomly for testing of predictive model quality. The samples in the library part of the dataset were the calibration samples in the baseline modeling scenario. The target-site spectra not included in a testing partition made a pool from which samples were picked for model training in other scenarios.

Noisy bands up to $600\,\text{cm}^{-1}$ and the 2268–$2389\,\text{cm}^{-1}$ wavenumber range affected by CO_2 were discarded. The spectra were further processed using a moving-average filter with an 11-band window. In addition to analyzing the obtained "raw" spectra, five further pre-processing schemes (Rinnan et al., 2009) were tested. First, the moving-average smoothing of the spectra was either followed with multiplicative scatter correction (MSC) or left unchanged. In the second phase, SNV, derivative transformation using the Savitzky–Golay filter with additional third-order polynomial smoothing applied over a moving window of 11 bands, or no transformation were applied to the result.

Calibration spiking was introduced, based on increasing spiking sample counts to the level of 16 samples with a step of 4 samples. In addition to the random scheme, two leverage sampling approaches were assessed for spiking sample selection: the Kennard–Stone algorithm preceded by PCA (Ng et al., 2018) and sampling based on conditioned Latin Hypercube (Minasny et al., 2006). Scenarios mirroring the calibration spiking scenarios, but with no samples from the long-term experiments, were also included.

Three components summing up to the whole soil sample were derived from the TC and HWC measurements: (1) HWC, (2) the part of TC resistant to hot-water extraction (nHWC), (3) and the non-TC part of a sample $(1 - \text{TC})$. In the next step, the component values were transformed into two isometric log-ratio (ilr) coordinates (Kynčlová et al., 2015): (1) ilr_{TC}, balancing TC and remainder of a sample; and (2) ilr_{HWC} balancing HWC and nHWC.

A bivariate PLSR model was trained for predicting the ilr coordinate values from MIR-DRIFTS spectra. For data partitionings that included both reference-library and target area samples, the influence of spiking sample weighing was examined by introducing models with 5-fold and 25-fold weighted local observations, in addition to unweighted models. The weighing was performed either by data row multiplication (standard approach), in which case a partial least squares regression 2 (PLSR2) model was used, or by exploiting the internal weighing capability of the CPLSR model family (proposed approach).

The numbers of PLSR components were tuned using cross-validation according to one-sigma heuristics (Dangal et al., 2019; Mevik et al., 2019) with values between 1 and 12 considered. The performance of each model was evaluated using test data partitions in terms of R^2 and root mean square error of prediction (RMSEP) statistics, prediction bias, followed by ratio of prediction to deviation (RPD) and ratio of prediction to interquartile range (RPIQ) statistics (Baldock et al., 2013; Clairotte et al., 2016; Cezar et al., 2019).

Chapter 6

Results and discussion

6.1 Drought stress detection in juvenile oilseed rape using HS imaging with a focus on spectra variability

This section is adapted from Żelazny; Lukáš (2020), previously published by MDPI Remote Sensing.

Accurate classification of the reflectance spectra was obtained with SVM, with all but 2 of the 47 pixels in the test set correctly assigned. The high pixel classification accuracy and plausible spectral patterns that can be discerned in the obtained classes highlight the potential of SVMs to segment OSR images.

When proportions of pixels identified by the SVM model as dry in a HS image were compared using linear model estimates, the narrowest posterior distribution was obtained for the cultivar contrast under the dry regime. However, since the distribution was centered close to the value of 1, it fails to provide information on the sign of the difference. Wide posterior distributions were obtained for the two remaining comparisons in this group. The contrasts involving the watering regime suggested an effect of restricted water supply on the dry pixel occurrence, albeit with high uncertainty. As expected, all but two comparisons indicated lower dry leaf surface area with improved water availability, especially for 'Viking'. Here, the multiplicative effect size (ES) along with the 95% credibility interval was $ES_{watered/dry} = 0.009\,[0.00005, 1.6]$, in agreement with the high drought sensitivity of this cultivar.

'Viking' and 'Cadeli' maintained under the watered treatment clearly differed according to the MCARI and MTCI mean index values. MCARI was higher for 'Viking' by $ES_{Viking-Cadeli} = 0.06\,[0.01, 0.11]$ units. As green reflectance is one of the components of MCARI, this effect can be explained by increased photosynthetic activity fostered by the favorable hydric conditions. The tendency towards minimizing the periods of stomatal closure allows the 'Viking' water-spender to thrive in the control watering regime. A link between low water stress and high MCARI values was demon-

strated by Peteinatos et al. (2016) in a spring wheat experiment. This interpretation can be questioned in the light of an inverse relationship obtained by Haboudane; Tremblay, et al. (2008) for corn. However, both these authors and Gitelson (2016) had observed a discrepancy pattern between MCARI and MTCI, which was also obtained in the present study. To a limited extent, the cultivars in the control treatment differed in terms of TCARI ($ES_{Viking-Cadeli} = 0.04 [-0.02, 0.10]$) and TCARI/OSAVI ($ES_{Viking-Cadeli} = 0.10 [-0.02, 0.21]$). The suitability of these indexes to crop water status diagnosis can be linked to pigmentation changes in drought-affected tissues (Haboudane; Miller, et al., 2002). Just as for MCARI, their increased values associated with the 'Viking' cultivar can be linked to its water-spender management strategy. In addition to the vegetation index mean values, their standard deviations differed across the two cultivars. An explanation linking those effects to the differing water management strategies seems dubious. More plausibly, they were caused by additional cultivar properties, in particular those related to the leaf surface and structure of the forming canopy (Mishra et al., 2017).

The influence of the watering regimes on the leaf spectra was captured by the mean values of several vegetation indexes. Unsurprisingly, particularly large differences were obtained for the watered:dry contrast. The RGI index exhibited high sensitivity in 'Cadeli', with its values lower in the control plants ($ES_{watered-dry} = -0.96 [-2.21, 0.21]$). The ESs obtained for both cultivars were in agreement with RGI increase in maize exposed to drought (Sun et al., 2018). The potential usefulness of this index is further illustrated by its strong negative correlation to leaf water status indicators investigated by Rodríguez-Pérez et al. (2007) in a commercial vineyard. Water availability had a positive influence on the MTCI, RNDVI, and GI indexes in the 'Cadeli' cultivar, with the effect not as strong as for RGI, but more precisely estimated. The RNDVI difference ($ES_{watered-dry} = 0.19 [-0.01, 0.41]$) was similar in magnitude to the spring wheat cultivars responses reported by Gutierrez et al. (2010). Although RNDVI was originally developed for woody species (Gitelson; Merzlyak, 1994), in light of the above findings, it seems to be also suited to OSR cultivation.

The variation in RGI and PSRI vegetation indexes exhibited sensitivity to the difference between the dry and control leaf spectra in 'Cadeli'. The estimated multiplicative treatment effects were $ES_{watered/dry,\sigma} = 0.10 [0.07, 0.16]$ and $0.10 [0.06, 0.17]$, respectively. They were weaker in 'Viking' plants. The variation in the majority of the remaining indexes were affected by the discussed treatment contrast for at least one of the cultivars. Several indexes revealed the difference between the rewatered and dry treatment, particularly PSRI

($ES_{\text{rewatered/dry},\sigma} = 0.33\,[0.16, 0.68]$ for 'Cadeli'). PSRI was proposed by Merzlyak et al. (1999) as an indicator index of leaf senescence, and was useful in discerning between barley drought senescence classes in the Behmann; Steinrücken, et al. (2014) study. The obtained PSRI standard deviation sensitivity to the contrasting watering regimes in 'Cadeli' can be linked to the source–sink character of the leaf senescence process (Munné-Bosch et al., 2004).

TCARI/OSAVI exhibited a strong response to the watered:rewatered contrast in both cultivars ($ES_{\text{watered/rewatered},\sigma} = 0.40\,[0.28, 0.55]$ for 'Cadeli'). Its formulation allows to disentangle the effect of chlorophyll and LAI (Haboudane; Miller, et al., 2002), as demonstrated by Haboudane; Tremblay, et al. (2008) and Perry; Roberts (2008). This is a valuable property, considering that chlorophyll content can increase in leaves as they shrink in drought conditions (Peñuelas; Gamon, et al., 1994; Linke et al., 2008). As the most striking finding, all of the affected indexes exhibited variation decrease with an improving water availability. This remarkable consistency corroborates the relationship between the stress level and symptom variability mentioned by Kruschke et al. (2017). In the light of this finding, vegetation index standard deviations appear to be sensitive stress indicators in the context of drought diagnosis using proximal HS imaging, perhaps more so than index means.

Regardless of the spectra pre-processing, no separation of the cultivars was obtained using the means of the principal component scores. However, the comparison of the first principal component (PC1) score standard deviations revealed less varied values for the watered 'Viking' plants relative to the 'Cadeli' cultivar with MSC pre-processing ($ES_{\text{Viking/Cadeli},\sigma} = 0.53\,[0.32, 0.96]$). This pattern can be explained by referring to the higher stress level of the latter genotype (Kosová et al., 2018). MSC is known to remove some scatter and baseline shift artifacts (Rinnan et al., 2009). In the present study, it might have mitigated the influence of variable illumination conditions on the captured HS data cubes.

A consistent pattern of watering-treatment effect was apparent for the remaining pre-processing approaches. The variability of the PC1 scores was found to be higher in the watered regime than in both dry ('Cadeli') and rewatered plants (both cultivars). When the latter two treatments were compared with each other, the dry spectra appeared to be more variable. Similarly to the TCARI/OSAVI standard deviation, the observed pattern may reveal a trace of a past severe drought episode in a seemingly healthy and well-hydrated crop.

6.2 Fusarium head blight detection from spectral measurements in a field phenotyping setting

This section is adapted from Żelazny; Chrpová, et al. (2021), previously published by Elsevier Biosystems Engineering.

Lower FHB VSSs were associated with increased red reflectance and the lowered NIR shoulder next to the red edge region. There was a successive blue shift as the ratings decreased towards the value of 4, and a qualitative change in the signature shape for the three lowest VSSs. Here, the curves assumed an almost linear shape in the VIS and red-edge regions and beyond. The wax-ripening later phenological phase was characterized by increased green and red reflectance relative to the milk-ripening phase.

The increase in VIS-region indicates pigment breakdown in the presence of fungi (West; Bravo, et al., 2003; Morin et al., 2017; Saccon et al., 2017). This agrees with the FHB "bleaching" symptoms (McMullen et al., 2012; Bauriegel; Herppich, 2014; Ma et al., 2020). The observed red-edge displacement towards the shorter wavelengths is also well-described (Martinelli et al., 2015; West; Canning, et al., 2017), as is the NIR reflectance decrease (Hamid Muhammed, 2005; Alisaac et al., 2018; Mahlein; Alisaac, et al., 2019; Zhang; Wang; Lin; Yin, et al., 2020; Zhang; Chen; Yin, et al., 2020). The qualitative change in spectral signature shape in the greatest disease severities suggests the onset of a new biological process, in particular the spikes becoming dry (Hamid Muhammed, 2005).

Most pre-processing schemes retained the possibility of discerning at least the curves corresponding to the lowest or highest infection levels. Subtraction and division resulted in the top-rated spectra following the horizontal line centered at 0 and 1 value respectively. The SNV yielded three bundles of curves, while prepending it with subtraction amplified the curve separation in the red-edge region, while suppressing the SWIR signal and introducing noise. Prepending the SNV with division also gave "noisy" signatures, and transformation adapted from Whetton; Waine, et al. (2018) resulted in a tight gradient.

While 100% correct VSS ratings within two-point error tolerance were predicted for large fractions of the test data partitions, predictions were poor in all zero-tolerance scenarios. However, relatively high performance was associated with the early-date predictions regardless of spectral data timing provided one-point VSS error was permitted. One VSS point difference in this study indicates disease symptom difference of approximately ten percentage points. Although this error allowance is much greater than errors reported in other FHB proximal sensing

experiments, some authors (Hamid Muhammed, 2005; Dammer et al., 2011; Jin et al., 2018) made no attempt to discriminate between disease severity levels or transformed the disease severities to binary scale prior to model training (Liu; Dong; Huang; Du; Ma, 2020; Ma et al., 2020).

A field phenotyping study by Qiu et al. (2019) estimated individual spike surface area percentages with visible FHB symptoms from RGB imagery. That process closely matched the expert rating procedure. The authors mentioned the "optimum performance" of their model but provided no numerical measures to enable comparison with other studies. This was followed by Su et al. (2021) proposal of a refined approach for application to 55 wheat genotypes. Although the authors reported negative prediction bias when disease severity was quantified on a 14-point scale, some performance measures exceeded 70 %. This suggests that RGB imagery combined with advanced processing is sufficient to phenotype FHB resistance in the field; at least when the disease is visible to the naked eye. However, it is important to note that the ground truth in this experiment was obtained by manually labeling the images, rather than based on field-rating by an expert, who would examine all sides of the spikes. In a deep-learning study with six-point ground-truth scale, Zhang; Wang; Gu, et al. (2019) reported accuracies high enough to justify replicating their experiment and testing the proposed approach in a phenotyping setting.

The SVM accuracy approached 100 % when two VSS-levels difference was permitted, and this was sufficient to reliably distinguish between low, medium, and heavily-infected cultivars. This result was similar to the four-level classification tested by Huang; Wu, et al. (2019), where the authors had reported greatest accuracies in the 85–90 % range. Despite its coarseness, this differentiation is sufficient for certain field phenotyping tasks; including the preliminary screening of disease-intolerant breeding lines (Bock et al., 2010).

An attempt was made to resolve the hypothesis of deteriorated model quality when late-date spectral signatures are replaced with early-date spectra. To that end, the prediction success probabilities corresponding to the different scanning dates were compared with the same VSS rating dates maintained. There was a positive effect of an earlier spectral campaign when subtraction or division pre-processing was applied after median aggregation to determine infection severity with single-point tolerance. The classification success odds increase was below 30 % when a predictive model was trained to the VSSs collected on the early date. On the other hand, a possible greater than 50 % increase was obtained for the late-date ratings. The early date corresponded to the milk-ripening stage, which Whetton; Hassall, et al. (2018) identified as being optimal for Fusarium infection severity prediction using PLSR

in laboratory. However, the diseased spikes have similar color to spikes unaffected by FHB as the crop matures, as reflected in temporal difference between spectral curves obtained in the present study. This ripening adversely affects spike detection in images (Fernandez-Gallego et al., 2018), and is detrimental to disease severity estimation accuracy (Bauriegel; Giebel, et al., 2011; Dammer et al., 2011; Qiu et al., 2019).

The effects of pre-processing schemes on the SVM prediction quality were examined, with the raw-spectra scenarios forming the baseline for comparison. The prediction performance remained almost the same if an error of two VSS points was allowed, and the response was weak for models trained to the early-date infection ratings when the tolerance was reduced to one point. Slight, but consistent performance decreases were associated with median aggregation followed by subtraction and division combined with SNV, and also with the pre-processing adapted from Whetton; Waine, et al. (2018). The detrimental effects of these schemes were confirmed by the zero-error tolerance scenario, where the probabilities of a correct VSS assignment were reduced by 63 % [32, 80], 56 % [19, 76], and 58 % [22, 78] when the milk-ripening phenological phase was involved. It was likely due to the "noisy" spectra resulting from these combinations.

Mean spectra aggregation preserved the negative effects of subtraction followed by SNV and the scheme adapted from Whetton; Waine, et al. (2018), which yielded a tight spectral VSS gradient. Although the ratings were consistently ordered, the signatures crossed and overlapped. This effect and the gradient steepness may have weakened the disease severity signal. Some evidence of accuracy increase with SNV was obtained. This transformation can reduce noise resulting from light beam scattering, and it was employed in some studies on Fusarium-affected kernels (Femenias et al., 2020). It can be recommended if both spectra acquisition and VSSs ratings are performed at the milk-ripening stage,

6.3 Application of feature selection for predicting leaf chlorophyll content in oats from HS imagery

This section is adapted from Żelazny (2020), previously published by Estonian Agricultural University, Faculty of Agronomy Agronomy Research.

When the experimental plot spectra were matched to the ground-truth data, high leaf chlorophyll concentration could be associated with increased NIR reflectance and a steep red edge. Both these regions have been repeatedly deemed as important for chlorophyll prediction by earlier studies (Govender et al., 2009; Main et al.,

2011). On the other hand, contrary to expectation, no apparent red-edge shift could be discerned.

Despite the variety of existing PLS flavors and implementations, the attempt to reproduce the validation results of the Meij et al. (2017) paper was successful. Only NRMSE exhibited a slight deviation from the original value. However, the high number of bands contributing to the model made the "full-spectrum" approach infeasible for practical application — at least until HS imagers would become affordable (Aasen et al., 2018). In addition, model tuning set the number of the latent variables to five, making an insight into its workings challenging.

The cross-validation of models that employed filter feature selection resulted in two local error minima. The absolute minimum corresponded to 19 input bands, a much lower number than for the reference model, but still too high for developing an affordable specialized device. What is more, the model exhibited higher validation error and involved even more latent variables (seven) than the reproduction model.

The second minimum reduced this set to three wavelengths, a seemingly good middle-ground between technical feasibility, model interpretability, and its expected estimation error. The validation statistics improved, corroborating the positive influence of feature selection on prediction accuracy (Mehmood et al., 2012). The first PLS component loading of this tuned model corresponded to the NIR reflectance, in accordance with the visual assessment, above. The second component could be interpreted as canopy darkness, and linked to the absorbance in the photosynthetically-active spectral region. The third component value was linked to canopy blueness.

Wavelength combinations similar to that picked by the filtering algorithm seldom occur in vegetation index formulations. They can be found in the Enhanced Vegetation Index (Gao; Huete, et al., 2000), the Structure Insensitive Pigment Index (Peñuelas; Baret, et al., 1995), the Modified Simple Ratio, and the Modified Normalized Difference (mND_{705}) (Sims et al., 2002). In the study by Main et al. (2011), the first three indexes fared poorly when used for predicting chlorophyll content in maize leaves at various developmental stages.

Regarding mND_{705}, it was among the best-performing indexes in Main et al. (2011) and in Miao et al. (2009) — also a maize study. On the other hand, it was reported as a poor predictor of chlorophyll content in rice (Moharana et al., 2016). The mND_{705} index formula includes a blue band, which accounts for specular reflectance (Sims et al., 2002). The third latent variable of the discussed PLS model might have played the same role, or it might have adjusted for Rayleigh scattering.

The forward selection within the wrapper approach stopped after picking one band (775 nm), thus reducing the PLS model to a classical regression model

with a single independent variable. The selected wavelength lied in the NIR spectral region, in agreement with the observation from the visual assessment. According to the validation statistics, despite extreme simplicity, the model performed surprisingly well in terms of RMSE. However, the low R^2 value put in question its practical utility.

6.4 Calibration spiking of MIR-DRIFTS soil spectra for carbon predictions using CPLSR and log-ratio transformations

This section is adapted from Żelazny; Šimon (2022), previously published by MDPI Agriculture.

Among the two target sites, Ústí nad Orlicí spectral signatures were highly varied. In certain regions, the signatures extended beyond the envelope of the library samples regardless of pre-processing, whereas the Janovice spectra were enveloped by the library spectra. Also the C measurement variation was high in Ústí nad Orlicí, not much smaller than that of the library samples despite the different geographical scales. This pattern indicates high spatial heterogeneity of Ústí nad Orlicí district soils. Consequently, it corroborates the need for dense soil sampling to map and monitor SC in the conditions of Czech Republic and, arguably, beyond (Paustian et al., 2019; Smith; Soussana, et al., 2019). The need to develop cost-effective assessment methods follows this consideration (Paustian et al., 2019). Janovice C measurements exhibited an apparent mismatch relative to the reference library measurements. Both TC and HWC were high here, and the only library samples with similar characteristics were a group of experimental plots assigned to compost fertilization treatments.

Regardless of the data subset, the raw measurements were skewed towards lower values. The skew, and to a degree high kurtosis, were reduced after the ilr transformations. As high performance of a PLSR model can be attained when the predicted variable has a Gaussian distribution, the obtained effect provides evidence of compatibility of log-ratios with PLSR predictive modeling.

The predictive performance of the PLSR models varied substantially. Although negative values were obtained for the worst models, models corresponding to R^2 in excess of 0.80 could be found for each ilr coordinate and target site combination. According to Janik; Skjemstad, et al. (2007), this is a high quality result. However, after aggregating the values across all data partitionings, R^2 exceeded 0.50, still an unsatisfactory value, only for Janovice while predicting ilr_{TC}, whereas both

ilr_{HWC} and Ústí nad Orlicí scenarios gave poor results. The worst negative biases and RMSEP values were comparable, amounting to 0.4–0.5 for ilr_{TC} and 0.2–0.3 for ilr_{HWC}. In terms of raw component values, these correspond to approximately 1.30 TC percentage points and 0.09–2.79 mg g^{-1} HWC, depending on the baseline HWC value. The best models had RMSEP of only 0.04 for ilr_{TC} (approximately 0.12 pp TC) and 0.03 for ilr_{HWC} (0.34 mg g^{-1} HWC for high value range and less for low value range). More conservative estimates, based on partitioning medians, suggested a possibility of predicting ilr_{TC} with an error of 0.13 (0.38 pp TC) and 0.08 (0.23 pp TC) in Ústí nad Orlicí and Janovice, respectively. For ilr_{HWC}, the corresponding values were 0.11 and 0.04 (0.04–1.23 and 0.01–0.45 mg g^{-1} HWC). Models with RPD or RPIQ above 2.5 or even 3.0 were obtained in some scenarios and test data partitions, described in literature as good and excellent predictions (Saeys et al., 2005). However, typically one could not expect the performance higher than 1.7, that is, barely sufficient to estimate the values even as high or low.

These evaluations do not corroborate the purported potential of MIR-DRIFTS to become a cost-effective yet reliable laboratory method for SC assessment (Viscarra Rossel et al., 2006; Reeves III, 2010; Kuang et al., 2012; Soriano-Disla et al., 2014; Seybold et al., 2019). Bellon-Maurel et al. (2011) and Barra et al. (2021) summarized model quality estimates for predicting TC and OC from MIR spectra. Although highly-performing models prevail in reported research, a number of SC studies suffer from methodological issues that arguably bias the results towards higher accuracy. In addition to that, not all models have been reported to perform well. The Bellon-Maurel et al. (2011) review includes formulations that resulted in modest RPD values, similar to those obtained in the present study. In the more recent Page et al. (2013) work, MIR-DRIFTS substantially underestimated OC loss over time in a long-term experiment. Moreover, the estimated effect of evaluated management treatments contradicted that inferred using traditional OC determination. Calderón et al. (2017) predicted OC in several crop experiments using PLSR, and obtained RMSEP of 0.67–0.80 pp, that is beyond the upper RMSEP conservative bracket for TC in the present study.

PLSR models trained to the spectral library, that is, with zero target-site samples, performed poorly, especially for Janovice. Training of the models to a selection of target-site samples, only, while excluding the spectral library, had a clearly positive effect on all measures even with only 4 training samples. This effect can be linked to the absence of samples with compost fertilization in the training dataset, analogously to the effect observed by Calderón et al. (2017) after training a model without an atypical site found in their data.

In Janovice scenarios with PLSR2 models, augmenting the library samples with 25-fold weighed spike samples yielded results competitive with the local approach. R^2 up to 0.71 could be attained with only 4 spiking samples for ilr_{TC} — in contrast to R^2 of corresponding local-only models, which was always negative. This effect is similar to the OC prediction pattern with NIR spectroscopy obtained by Guerrero; Stenberg, et al. (2014) while increasing the number of target samples from 8 to 16 and 32. It is also in line with the hypothesis on calibration spiking potential to reduce the number of necessary reference samples.

A pattern of Kennard–Stone leverage sampling inferiority could be discerned for Ústí nad Orlicí. This leverage sampling scheme tends to pick distant observations, located at the edges of a hyperspace. It also operates incrementally, as opposed to conditioned Latin hypercube, in the case of which the spectra are picked at once and can be more representative of a dataset (Ng et al., 2018). When applied to the heterogeneous Ústí nad Orlicí dataset, Kennard–Stone might have picked outlier spiking samples, perhaps corresponding to soils with atypical textures (Stumpe et al., 2011) or mineralogy (Reeves et al., 2001; Calderón et al., 2017).

Contrary to the hypothesis, the application of the CPLSR method was clearly detrimental for the prediction quality of both ilr_{TC} and ilr_{HWC} in Janovice samples compared to the standard approach. Sankey et al. (2008) attempted to predict SC from VisNIR spectral data using boosted regression trees for different levels of local sample weights relative to the weights of the samples in the reference library. The authors expressed skepticism with respect to their results, in which the model performance decreased substantially for one target site. Although a positive relationship was observed for another, the obtained improvement was modest. Still, given the limited number of studies devoted to the topic so far, it seems worthwhile to further explore effects of embedded weighing with other data and other classes of predictive models (Janik; Forrester, et al., 2009; Deiss et al., 2020).

Chapter 7

Conclusions and recommendations

This chapter is adapted from Żelazny; Lukáš (2020), previously published by MDPI Remote Sensing; Żelazny; Chrpová, et al. (2021), previously published by Elsevier Biosystems Engineering; Żelazny (2020), previously published by Estonian Agricultural University, Faculty of Agronomy Agronomy Research; and Żelazny; Šimon (2022), previously published by MDPI Agriculture.

According to the latest literature, the superiority of hyperspectral measurements relative to multispectral measurements is unequivocal with respect to crop stress detection. However, HS data carry with them the "curse of dimensionality" problem, can be more affected by noise, and there is an increased risk of predictive model overfitting. Therefore, it is imperative that hyperspectra processing workflows continue to be developed, while those already published need to be validated by multiple teams in a range of conditions to ensure their robustness.

In imaging solutions, large volumes of HS data cubes pose a computational challenge, and the acquisition equipment is expensive. Before these issues become addressed, a focus on small sets of pre-selected narrow bands captured by more affordable specialized devices can be a feasible middle path. In the study with oats, it was possible to train a PLSR model that predicted leaf chlorophyll from only three spectral inputs. The pigment concentration could be estimated from NIR reflectance, canopy darkness, and its blueness. Chlorophyll content estimation is of practical use for precision agriculture, as crop stress occurrence triggers pigment deterioration. A relatively simple sensor for airborne field mapping or for on-line mapping during a tractor operation can, therefore, be envisioned exploiting the model.

The robustness of the solution remains to be investigated using an ensemble of dataset partitionings and ground truth obtained from laboratory analyzes instead of a SPAD chlorophyll meter. HS patterns of chlorophyll-deficient oats obtained from airborne imagery mostly matched those described in literature for vegetation in general. This increases the odds that a model could be subsequently transferred

also to other crop species.

In the study where a HS camera was employed to acquire data cubes of water-stressed OSR plants, a SVM accurately distinguished between healthy leaf zones, those bearing signs of drought discoloration, and the background. At the same time, the process required time-consuming pixel labelling to train the classification model, and possibilities of reducing this effort still need to be tested. There was an increase in the discoloured leaf surface area in the 'Viking' cultivar when subjected to drought stress, a pattern that would corroborate the higher susceptibility of this genotype. Future studies should include additional cultivars representing each water management strategy to control for nuisance effects of traits unrelated to OSR drought tolerance.

Several vegetation index means responded to the difference between the control and water-deprived plants, especially RGI, MTCI, RNDVI, and GI; but the most striking finding was a consistent increase in the multiple index standard deviations to worsening of the hydric regime. This result suggests higher sensitivity of the vegetation index variability measures relative to the means for OSR drought stress diagnosis. Vegetation indexes are frequently derived from only two spectral bands. The proposed approach could then, similarly to feature selection, contribute to the development of affordable crop stress sensors. However, it needs to be noted that imagery with high spatial resolution is required to capture within-leaf reflectance variability. For this reason, the imager would be primarily suited to proximal applications — for example, mounted on agricultural machinery or a field phenotyping robot.

Additional factors can be introduced in follow-up studies to verify the robustness of the findings and their application to earlier drought stress detection. A single campaign could be replaced by a time series to capture the temporal development of the drought stress and of the spectral responses. Another modification would be to restrict the watering of the plants at an earlier developmental phase. More insights could be obtained by augmenting the new dataset with biochemical and physiological measurements. Regardless of the study extensions, the observed patterns need to be replicated in an independent experiment with a larger sample.

In the winter wheat field phenotyping study, the patterns in the acquired spectral signatures of FHB-infected spikes could be explained by plant stress biology and linked to the disease visual symptoms. These agreed with the spectral patterns obtained in other FHB studies, and this supports utility of a spectroradiometer with a contact probe for FHB phenotyping applications. Ordinal SVM models applied to the raw spectra yielded predictions with confident distinction of low, moderate, and high disease severity at the milk-ripening developmental phase. The proposed

approach enables screening of unpromising breeding lines while developing new crop cultivars. Other predictive models require testing for applications that demand finer FHB severity ratings. In the context of precision agriculture, the solution could be adapted to facilitate separate harvest of healthy and FHB-infected grains.

Spike-drying discolor masked the disease signal at the wax-ripening developmental stage. Predictions at this phase can be improved by subtracting healthy-spike reflectance measurements from the analyzed signatures or by division. However, gains from these pre-processing schemes must be compared to the costs of additional scans. While deep learning may be a more suitable alternative for FHB severity estimation, the current published results require replication with multiple cultivar datasets and avoid over-simplification of ground-truth data. Further investigation is also required to determine if a single trained model can be applied year-to-year to reduce the reference data collection effort.

In the soil spectroscopy study, conservative estimates of PLSR model performances were lower than the values typically reported for MIR-DRIFTS SC predictions. Introduction of spiking samples combined with heavy weighing and PLSR2 modeling was associated with an increase of model quality. It, therefore, appears to be a promising cost-effective and environmentally-friendly SC monitoring solution. On the other hand, application of Kennard–Stone leverage sampling is not recommended, at least for heterogeneous datasets. The evaluated application of MIR-DRIFTS spectroscopy is primarily relevant for large-scale SC inventories. However, C lability mapping can also be of interest to precision agriculture practitioners, as a factor to be accounted for in variable-rate fertilization.

Not enough representative training data were available to satisfactorily predict SC properties in the more geographically extensive district-scale dataset. The large training-dataset requirement calls for international cooperation to standardize data collection and treatment workflows, harmonize spectral libraries, and facilitate their use. Log-ratio transformation of SC laboratory reference measurements improves data distributional properties and, therefore, is compatible with PLSR modelling. Accounting for carbon saturation limits, and treating spectral measurements as compositional are potential further refinements of this approach.

Early crop stress detection based on HS reflectance is a broad research subject, and as such it could benefit immensely from Open Science practices to avoid work duplication, facilitate predictive workflow validation, synthesis of the findings, and their transfer to the industry. For this reason, the studies included in this dissertation aimed not only to address pertinent research questions and explore novel solutions, but also to share the collected data, analyze them in a reproducible fashion,

and avoid research bias. In particular, the pre-registered FHB study demonstrates that with proper tooling, some of the recommended best Open Science practices can be followed without adding excessive burden on researchers. The material accompanying the publications included in this dissertation thesis can be consulted by teams interested in making their work more transparent and accessible to the research community.

Bibliography

AASEN, H.; HONKAVAARA, E.; LUCIEER, A.; ZARCO-TEJADA, P., 2018. Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures, and Data Correction Workflows. *Remote Sensing*. Vol. 10, no. 7, p. 1091. Available from DOI: 10.3390/rs10071091.

ADÃO, T.; HRUŠKA, J.; PÁDUA, L.; BESSA, J.; PERES, E.; MORAIS, R.; SOUSA, J. J., 2017. Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sensing*. Vol. 9, no. 11, p. 1110. Available from DOI: 10.3390/rs9111110.

AGUATE, F. M.; TRACHSEL, S.; PÉREZ, L. G.; BURGUEÑO, J.; CROSSA, J.; BALZARINI, M.; GOUACHE, D.; BOGARD, M.; CAMPOS, G. de los, 2017. Use of Hyperspectral Image Data Outperforms Vegetation Indices in Prediction of Maize Yield. *Crop Science*. Vol. 57, no. 5, pp. 2517–2524. Available from DOI: 10.2135/cropsci2017.01.0007.

ALCKMIN, G. T. de; LUCIEER, A.; ROERINK, G.; RAWNSLEY, R.; HOVING, I.; KOOISTRA, L., 2020. Retrieval of Crude Protein in Perennial Ryegrass Using Spectral Data at the Canopy Level. *Remote Sensing*. Vol. 12, no. 18, p. 2958. Available from DOI: 10.3390/rs12182958.

ALISAAC, E.; BEHMANN, J.; KUSKA, M.; DEHNE, H.-W.; MAHLEIN, A.-K., 2018. Hyperspectral quantification of wheat resistance to *Fusarium* head blight: comparison of two *Fusarium* species. *European Journal of Plant Pathology*. Vol. 152, pp. 869–884. Available from DOI: 10.1007/s10658-018-1505-9.

AMIGO, J. M.; BABAMORADI, H.; ELCOROARISTIZABAL, S., 2015. Hyperspectral image analysis. A tutorial. *Analytica Chimica Acta*. Vol. 896, pp. 34–51. Available from DOI: 10.1016/j.aca.2015.09.030.

ARNELL, N.; LOWE, J.; CHALLINOR, A.; OSBORN, T., 2019. Global and regional impacts of climate change at different levels of global temperature increase. *Climatic Change*. Vol. 155, pp. 377–391. Available from DOI: 10.1007/s10584-019-02464-z.

ASAARI, M. S. M.; MERTENS, S.; DHONDT, S.; INZÉ, D.; WUYTS, N.; SCHEUNDERS, P., 2019. Analysis of hyperspectral images for detection of drought stress and recovery in maize plants in a high-throughput phenotyping platform. *Computers and Electronics in Agriculture*. Vol. 162, pp. 749–758. Available from DOI: 10.1016/j.compag.2019.05.018.

ASHRAF, M.; HARRIS, P., 2013. Photosynthesis under stressful environments: An overview. *Photosynthetica*. Vol. 51, pp. 163–190. Available from DOI: 10.1007/s11099-013-0021-6.

BAKKER, M.; VELDKAMP, C. L.; ASSEN, M. A. van; CROMPVOETS, E. A.; ONG, H. H.; NOSEK, B. A.; SODERBERG, C. K.; MELLOR, D.; WICHERTS, J. M., 2020. Ensuring the quality and specificity of preregistrations. *PLoS biology*. Vol. 18, no. 12, e3000937. Available from DOI: 10.1371/journal.pbio.3000937.

BALDOCK, J.; HAWKE, B.; SANDERMAN, J.; MACDONALD, L., 2013. Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. *Soil Research*. Vol. 51, no. 8, pp. 577–595. Available from DOI: 10.1071/SR13077.

BARMEIER, G.; SCHMIDHALTER, U., 2016. High-Throughput Phenotyping of Wheat and Barley Plants Grown in Single or Few Rows in Small Plots Using Active and Passive Spectral Proximal Sensing. *Sensors*. Vol. 16, no. 11, p. 1860. Available from DOI: 10.3390/s16111860.

BARRA, I.; HAEFELE, S. M.; SAKRABANI, R.; KEBEDE, F., 2021. Soil spectroscopy with the use of chemometrics, machine learning and pre-processing techniques in soil diagnosis: Recent advances–A review. *TrAC Trends in Analytical Chemistry*. Vol. 135, p. 116166. Available from DOI: 10.1016/j.trac.2020.116166.

BARTON, C. V., 2001. A theoretical analysis of the influence of heterogeneity in chlorophyll distribution on leaf reflectance. *Tree Physiology*. Vol. 21, no. 12-13, pp. 789–795. Available from DOI: 10.1093/treephys/21.12-13.789.

BATJES, N. H.; WESEMAEL, B. van, 2015. Measuring and Monitoring Soil Carbon. In: BANWART, S. A.; NOELLEMEYER, E.; MILNE, E. (eds.). *Soil carbon: science, management and policy for multiple benefits*. Wallingford, United Kingdom: CAB International. Vol. 71, pp. 188–201. SCOPE.

BAUMANN, P.; HELFENSTEIN, A.; GUBLER, A.; KELLER, A.; MEULI, R. G.; WACHTER, D.; LEE, J.; VISCARRA ROSSEL, R. A.; SIX, J., 2021. Developing the Swiss mid-infrared soil spectral library for local estimation and monitoring. *Soil*. Vol. 7, pp. 525–546. Available from DOI: 10.5194/soil-7-525-2021.

BAURIEGEL, E.; GIEBEL, A.; GEYER, M.; SCHMIDT, U.; HERPPICH, W., 2011. Early detection of *Fusarium* infection in wheat using hyper-spectral imaging. *Computers and Electronics in Agriculture*. Vol. 75, no. 2, pp. 304–312. Available from DOI: 10.1016/j.compag.2010.12.006.

BAURIEGEL, E.; HERPPICH, W. B., 2014. Hyperspectral and Chlorophyll Fluorescence Imaging for Early Detection of Plant Diseases, with Special Reference to *Fusarium* spec. Infections on Wheat. *Agriculture*. Vol. 4, no. 1, pp. 32–57. Available from DOI: 10.3390/agriculture4010032.

BEHMANN, J.; SCHMITTER, P.; STEINRÜCKEN, J.; PLÜMER, L., 2014. Ordinal classification for efficient plant stress prediction in hyperspectral data. *ISPRS - International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XL, no. 7, pp. 29–36. Available from DOI: 10.5194/isprsarchives-XL-7-29-2014.

BEHMANN, J.; STEINRÜCKEN, J.; PLÜMER, L., 2014. Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*. Vol. 93, pp. 98–111. Available from DOI: 10.1016/j.isprsjprs.2014.03.016.

BELLON-MAUREL, V.; MCBRATNEY, A., 2011. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils – Critical review and research perspectives. *Soil Biology and Biochemistry*. Vol. 43, no. 7, pp. 1398–1410. Available from DOI: 10.1016/j.soilbio.2011.02.019.

BLACKBURN, G. A., 2007. Hyperspectral remote sensing of plant pigments. *Journal of Experimental Botany*. Vol. 58, no. 4, pp. 855–867. Available from DOI: 10.1093/jxb/erl123.

BLUM, A., 2005. Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive? *Australian Journal of Agricultural Research*. Vol. 56, no. 11, pp. 1159–1168. Available from DOI: 10.1071/AR05069.

BOCK, C.; POOLE, G.; PARKER, P.; GOTTWALD, T., 2010. Plant Disease Severity Estimated Visually, by Digital Photography and Image Analysis, and by Hy-

perspectral Imaging. *Critical Reviews in Plant Sciences*. Vol. 29, no. 2, pp. 59–107. Available from DOI: 10.1080/07352681003617285.

BONJEAN, A. P.; DEQUIDT, C.; SANG, T., 2016. Rapeseed in China. *OCL*. Vol. 23, no. 6, p. D605. Available from DOI: 10.1051/ocl/2016045.

BRUNING, B.; LIU, H.; BRIEN, C.; BERGER, B.; LEWIS, M.; GARNETT, T., 2019. The Development of Hyperspectral Distribution Maps to Predict the Content and Distribution of Nitrogen and Water in Wheat (*Triticum aestivum*). *Frontiers in Plant Science*. Vol. 10. Available from DOI: 10.3389/fpls.2019.01380.

BUDDENBAUM, H.; STEFFENS, M., 2012. The Effects of Spectral Pretreatments on Chemometric Analyses of Soil Profiles Using Laboratory Imaging Spectroscopy. *Applied and Environmental Soil Science*. Vol. 2012, p. 274903. Available from DOI: 10.1155/2012/274903.

BUERSTMAYR, M.; STEINER, B.; BUERSTMAYR, H., 2020. Breeding for Fusarium head blight resistance in wheat—Progress and challenges. *Plant Breeding*. Vol. 139, no. 3, pp. 429–454. Available from DOI: 10.1111/pbr.12797.

BUEZO, J.; SANZ-SAEZ, Á.; MORAN, J. F.; SOBA, D.; ARANJUELO, I.; ESTEBAN, R., 2019. Drought tolerance response of high-yielding soybean varieties to mild drought: physiological and photochemical adjustments. *Physiologia Plantarum*. Vol. 166, no. 1, pp. 88–104. Available from DOI: 10.1111/ppl.12864.

BULLOCK, D. S.; BULLOCK, D. G., 2000. From Agronomic Research to Farm Management Guidelines: A Primer on the Economics of Information and Precision Technology. *Precision Agriculture*. Vol. 2, pp. 71–101. Available from DOI: 10.1023/A:1009988617622.

BÜNEMANN, E. K.; BONGIORNO, G.; BAI, Z.; CREAMER, R. E.; DE DEYN, G.; GOEDE, R. de; FLESKENS, L.; GEISSEN, V.; KUYPER, T. W.; MÄDER, P.; PULLEMAN, M.; SUKKEL, W.; GROENINGEN, J. W. van; BRUSSAARD, L., 2018. Soil quality – A critical review. *Soil Biology and Biochemistry*. Vol. 120, pp. 105–125. Available from DOI: 10.1016/j.soilbio.2018.01.030.

BÜRKNER, P.-C., 2018. Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*. Vol. 10, no. 1, pp. 395–411. Available from DOI: 10.32614/RJ-2018-017.

CAICEDO, J. P. R.; VERRELST, J.; MUÑOZ-MARÍ, J.; MORENO, J.; CAMPS-VALLS, G., 2014. Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. *IEEE Journal of Selected Topics in Applied Earth Obser-*

vations and Remote Sensing. Vol. 7, no. 4, pp. 1249–1259. Available from DOI: 10.1109/JSTARS.2014.2298752.

CALDERÓN, F. J.; CULMAN, S.; SIX, J.; FRANZLUEBBERS, A. J.; SCHIPANSKI, M.; BENISTON, J.; GRANDY, S.; KONG, A. Y., 2017. Quantification of Soil Permanganate Oxidizable C (POXC) Using Infrared Spectroscopy. *Soil Science Society of America Journal.* Vol. 81, no. 2. Available from DOI: 10.2136/sssaj2016.07.0216.

CAMBAZA, E.; KOSEKI, S.; KAWAMURA, S., 2019. Why RGB Imaging Should Be Used to Analyze *Fusarium Graminearum* Growth and Estimate Deoxynivalenol Contamination. *Methods and Protocols.* Vol. 2, no. 1, p. 25. Available from DOI: 10.3390/mps2010025.

CAO, X.; LUO, Y.; ZHOU, Y.; DUAN, X.; CHENG, D., 2013. Detection of powdery mildew in two winter wheat cultivars using canopy hyperspectral reflectance. *Crop Protection.* Vol. 45, pp. 124–131. Available from DOI: 10.1016/j.cropro.2012.12.002.

CAPOLUPO, A.; KOOISTRA, L.; BERENDONK, C.; BOCCIA, L.; SUOMALAINEN, J., 2015. Estimating Plant Traits of Grasslands from UAV-Acquired Hyperspectral Images: A Comparison of Statistical Approaches. *ISPRS International Journal of Geo-Information.* Vol. 4, no. 4, pp. 2792–2820. Available from DOI: 10.3390/ijgi4042792.

CAPRON, X.; WALCZAK, B.; DE NOORD, O.; MASSART, D., 2005. Selection and weighting of samples in multivariate regression model updating. *Chemometrics and Intelligent Laboratory Systems.* Vol. 76, no. 2, pp. 205–214. Available from DOI: 10.1016/j.chemolab.2004.11.003.

CARPENTER, B.; GELMAN, A.; HOFFMAN, M. D.; LEE, D.; GOODRICH, B.; BETANCOURT, M.; BRUBAKER, M.; GUO, J.; LI, P.; RIDDELL, A., 2017. Stan: A Probabilistic Programming Language. *Journal of Statistical Software.* Vol. 76, no. 1, pp. 1–32. Available from DOI: 10.18637/jss.v076.i01.

CARTER, G. A., 1993. Responses of leaf spectral reflectance to plant stress. *American Journal of Botany*, pp. 239–243.

CARTER, G. A.; KNAPP, A. K., 2001. Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. *American Journal of Botany.* Vol. 88, no. 4, pp. 677–684. Available from DOI: 10.2307/2657068.

CEZAR, E.; NANNI, M. R.; GUERRERO, C.; SILVA JUNIOR, C. A. da; CRUCIOL, L. G. T.; CHICATI, M. L.; SILVA, G. F. C., 2019. Organic matter and sand estimates by spectroradiometry: Strategies for the development of models with applicability at a local scale. *Geoderma.* Vol. 340, pp. 224–233. Available from DOI: 10.1016/j.geoderma.2019.01.021.

CHE, X.; XU, S., 2010. Bayesian data analysis for agricultural experiments. *Canadian Journal of Plant Science.* Vol. 90, no. 5, pp. 575–603. Available from DOI: 10.4141/cjps10004.

CLAIROTTE, M.; GRINAND, C.; KOUAKOUA, E.; THÉBAULT, A.; SABY, N. P.; BERNOUX, M.; BARTHÈS, B. G., 2016. National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. *Geoderma.* Vol. 276, pp. 41–52. Available from DOI: 10.1016/j.geoderma.2016.04.021.

DAMMER, K.-H.; MÖLLER, B.; RODEMANN, B.; HEPPNER, D., 2011. Detection of head blight (*Fusarium* ssp.) in winter wheat by color and multispectral image analyses. *Crop Protection.* Vol. 30, no. 4, pp. 420–428. Available from DOI: 10.1016/j.cropro.2010.12.015.

DANGAL, S. R.; SANDERMAN, J.; WILLS, S.; RAMIREZ-LOPEZ, L., 2019. Accurate and Precise Prediction of Soil Properties from a Large Mid-Infrared Spectral Library. *Soil Systems.* Vol. 3, no. 1, p. 11. Available from DOI: 10.3390/soilsystems3010011.

DARYANTO, S.; WANG, L.; JACINTHE, P.-A., 2017. Global synthesis of drought effects on cereal, legume, tuber and root crops production: A review. *Agricultural Water Management.* Vol. 179, pp. 18–33. Available from DOI: 10.1016/j.agwat.2016.04.022.

DE JAY, N.; PAPILLON-CAVANAGH, S.; OLSEN, C.; EL-HACHEM, N.; BONTEMPI, G.; HAIBE-KAINS, B., 2013. mRMRe: an R package for parallelized mRMR ensemble feature selection. *Bioinformatics.* Vol. 29, no. 18, pp. 2365–2368. Available from DOI: 10.1093/bioinformatics/btt383.

DEERY, D.; JIMENEZ-BERNI, J.; JONES, H.; SIRAULT, X.; FURBANK, R., 2014. Proximal Remote Sensing Buggies and Potential Applications for Field-Based Phenotyping. *Agronomy.* Vol. 4, no. 3, pp. 349–379. Available from DOI: 10.3390/agronomy4030349.

DEISS, L.; MARGENOT, A. J.; CULMAN, S. W.; DEMYAN, M. S., 2020. Tuning support vector machines regression models improves prediction accuracy of soil

properties in MIR spectroscopy. *Geoderma*. Vol. 365, p. 114227. Available from DOI: 10.1016/j.geoderma.2020.114227.

DIN, J.; KHAN, S.; ALI, I.; GURMANI, A., 2011. Physiological and agronomic response of canola varieties to drought stress. *Journal of Animal & Plant Sciences*. Vol. 21, no. 1, pp. 78–82.

DING, C.; PENG, H., 2005. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology*. Vol. 3, no. 2, pp. 185–205. Available from DOI: 10.1142/S0219720005001004.

DU, C.; ZHOU, J., 2009. Evaluation of Soil Fertility Using Infrared Spectroscopy – A Review. In: *Climate Change, Intercropping, Pest Control and Beneficial Microorganisms*. Springer, pp. 453–483. Available from DOI: 10.1007/978-90-481-2716-0_16.

FEMENIAS, A.; GATIUS, F.; RAMOS, A. J.; SANCHIS, V.; MARÍN, S., 2020. Use of hyperspectral imaging as a tool for *Fusarium* and deoxynivalenol risk management in cereals: A review. *Food Control*. Vol. 108, p. 106819. Available from DOI: 10.1016/j.foodcont.2019.106819.

FENG, W.; SHEN, W.; HE, L.; DUAN, J.; GUO, B.; LI, Y.; WANG, C.; GUO, T., 2016. Improved remote sensing detection of wheat powdery mildew using dual-green vegetation indices. *Precision Agriculture*. Vol. 17, no. 5, pp. 608–627. Available from DOI: 10.1007/s11119-016-9440-2.

FERNANDEZ-GALLEGO, J. A.; KEFAUVER, S. C.; GUTIÉRREZ, N. A.; NIETO-TALADRIZ, M. T.; ARAUS, J. L., 2018. Wheat ear counting in-field conditions: high throughput and low-cost approach using RGB images. *Plant Methods*. Vol. 14, p. 22. Available from DOI: 10.1186/s13007-018-0289-4.

FERNÁNDEZ-NAVARRO, F., 2017. A Generalized Logistic Link Function for Cumulative Link Models in Ordinal Regression. *Neural Processing Letters*. Vol. 46, pp. 251–269. Available from DOI: 10.1007/s11063-017-9589-3.

FLEXAS, J.; BOTA, J.; LORETO, F.; CORNIC, G.; SHARKEY, T., 2004. Diffusive and metabolic limitations to photosynthesis under drought and salinity in C_3 plants. *Plant Biology*. Vol. 6, no. 3, pp. 269–279. Available from DOI: 10.1055/s-2004-820867.

FRANK, I. E.; FRIEDMAN, J. H., 1993. A Statistical View of Some Chemometrics Regression Tools. *Technometrics*. Vol. 35, no. 2, pp. 109–135.

FRANKE, J.; MENZ, G., 2007. Multi-temporal wheat disease detection by multi-spectral remote sensing. *Precision Agriculture*. Vol. 8, pp. 161–172. Available from DOI: 10.1007/s11119-007-9036-y.

GAGO, J.; DOUTHE, C.; COOPMAN, R. E.; GALLEGO, P. P.; RIBAS-CARBO, M.; FLEXAS, J.; ESCALONA, J.; MEDRANO, H., 2015. UAVs challenge to assess water stress for sustainable agriculture. *Agricultural Water Management*. Vol. 153, pp. 9–19. Available from DOI: 10.1016/j.agwat.2015.01.020.

GANI, W.; LIMAM, M., 2016. A kernel distance-based representative subset selection method. *Journal of Statistical Computation and Simulation*. Vol. 86, no. 1, pp. 135–148. Available from DOI: 10.1080/00949655.2014.996758.

GAO, X.; HUETE, A. R.; NI, W.; MIURA, T., 2000. Optical–biophysical relationships of vegetation spectra without background contamination. *Remote Sensing of Environment*. Vol. 74, no. 3, pp. 609–620. Available from DOI: 10.1016/S0034-4257(00)00150-4.

GAO, Z.; LUO, Z.; ZHANG, W.; LV, Z.; XU, Y., 2020. Deep Learning Application in Plant Stress Imaging: A Review. *AgriEngineering*. Vol. 2, no. 3, pp. 430–446. Available from DOI: 10.3390/agriengineering2030029.

GASPAR, T.; FRANCK, T.; BISBIS, B.; KEVERS, C.; JOUVE, L.; HAUSMAN, J.; DOMMES, J., 2002. Concepts in plant stress physiology. Application to plant tissue cultures. *Plant Growth Regulation*. Vol. 37, pp. 263–285. Available from DOI: 10.1023/A:1020835304842.

GE, Y.; BAI, G.; STOERGER, V.; SCHNABLE, J. C., 2016. Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Computers and Electronics in Agriculture*. Vol. 127, pp. 625–632. Available from DOI: 10.1016/j.compag.2016.07.028.

GE, Y.; THOMASSON, J. A.; SUI, R., 2011. Remote sensing of soil properties in precision agriculture: A review. *Frontiers of Earth Science*. Vol. 5, no. 3, pp. 229–238. Available from DOI: 10.1007/s11707-011-0175-0.

GELADI, P.; BURGER, J.; LESTANDER, T., 2004. Hyperspectral imaging: calibration problems and solutions. *Chemometrics and Intelligent Laboratory Systems*. Vol. 72, no. 2, pp. 209–217. Available from DOI: 10.1016/j.chemolab.2004.01.023.

GELFAND, A. E., 2000. Gibbs Sampling. *Journal of the American Statistical Association*. Vol. 95, no. 452, pp. 1300–1304. Available from DOI: 10.1080/01621459.2000.10474335.

GHOLIZADEH, A.; KOPAČKOVÁ, V., 2019. Detecting vegetation stress as a soil contamination proxy: a review of optical proximal and remote sensing techniques. *International Journal of Environmental Science and Technology*. Vol. 16, pp. 2511–2524. Available from DOI: 10.1007/s13762-019-02310-w.

GHOLIZADEH, A.; BORŮVKA, L.; SABERIOON, M.; VAŠÁT, R., 2013. Visible, Near-Infrared, and Mid-Infrared Spectroscopy Applications for Soil Assessment with Emphasis on Soil Organic Matter Content and Quality: State-of-the-Art and Key Issues. *Applied Spectroscopy*. Vol. 67, no. 12, pp. 1349–1362. Available from DOI: 10.1366/13-07288.

GILBERT, J.; HABER, S., 2013. Overview of some recent research developments in fusarium head blight of wheat. *Canadian Journal of Plant Pathology*. Vol. 35, no. 2, pp. 149–174. Available from DOI: 10.1080/07060661.2013.772921.

GILBERT, M. E.; ZWIENIECKI, M. A.; HOLBROOK, N. M., 2011. Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use efficiency in 11 soybean genotypes before and during mild drought. *Journal of Experimental Botany*. Vol. 62, no. 8, pp. 2875–2887. Available from DOI: 10.1093/jxb/erq461.

GILL, S. S.; TUTEJA, N., 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiology and Biochemistry*. Vol. 48, no. 12, pp. 909–930. Available from DOI: 10.1016/j.plaphy.2010.08.016.

GITELSON, A.; MERZLYAK, M. N., 1994. Spectral Reflectance Changes Associated with Autumn Senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *Journal of Plant Physiology*. Vol. 143, no. 3, pp. 286–292. Available from DOI: 10.1016/S0176-1617(11)81633-0.

GITELSON, A. A., 2016. Remote Sensing Estimation of Crop Biophysical Characteristics at Various Scales. In: THENKABAIL, P. S.; LYON, J. G. (eds.). *Hyperspectral Remote Sensing of Vegetation*. Boca Raton: CRC press, pp. 329–358.

GNYP, M.; PANITZKI, M.; REUSCH, S., 2015. In: *Precision agriculture '15*. Wageningen Academic Publishers. Chap. Proximal nitrogen sensing by off-nadir and nadir measurements in winter wheat canopy, pp. 43–50. Available from DOI: 10.3920/978-90-8686-814-8_4.

GOVENDER, M.; GOVENDER, P. J.; WEIERSBYE, I. M.; WITKOWSKI, E. T. F.; AHMED, F., 2009. Review of commonly used remote sensing and ground-based technologies to measure plant water stress. *Water SA*. Vol. 35, no. 5, pp. 741–752. Available from DOI: 10.4314/wsa.v35i5.49201.

GREDILLA, A.; VALLEJUELO, S. F.-O. de; ELEJOSTE, N.; DIEGO, A. de; MADARIAGA, J. M., 2016. Non-destructive Spectroscopy combined with chemometrics as a tool for Green Chemical Analysis of environmental samples: A review. *TrAC Trends in Analytical Chemistry*. Vol. 76, pp. 30–39. Available from DOI: 10.1016/j.trac.2015.11.011.

GUERRERO, C.; STENBERG, B.; WETTERLIND, J.; VISCARRA ROSSEL, R.; MAESTRE, F.; MOUAZEN, A. M.; ZORNOZA, R.; RUIZ-SINOGA, J.; KUANG, B., 2014. Assessment of soil organic carbon at local scale with spiked NIR calibrations: effects of selection and extra-weighting on the spiking subset. *European Journal of Soil Science*. Vol. 65, no. 2, pp. 248–263. Available from DOI: 10.1111/ejss.12129.

GUERRERO, C.; ZORNOZA, R.; GÓMEZ, I.; MATAIX-BENEYTO, J., 2010. Spiking of NIR regional models using samples from target sites: Effect of model size on prediction accuracy. *Geoderma*. Vol. 158, no. 1–2, pp. 66–77. Available from DOI: 10.1016/j.geoderma.2009.12.021.

GUTIERREZ, M.; REYNOLDS, M. P.; RAUN, W. R.; STONE, M. L.; KLATT, A. R., 2010. Spectral Water Indices for Assessing Yield in Elite Bread Wheat Genotypes under Well-Irrigated, Water-Stressed, and High-Temperature Conditions. *Crop Science*. Vol. 50, no. 1, pp. 197–214. Available from DOI: 10.2135/cropsci2009.07.0381.

GUYON, I.; ELISSEEFF, A., 2003. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*. Vol. 3, no. Mar, pp. 1157–1182.

HABOUDANE, D.; MILLER, J. R.; TREMBLAY, N.; ZARCO-TEJADA, P. J.; DEXTRAZE, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*. Vol. 81, no. 2-3, pp. 416–426. Available from DOI: 10.1016/S0034-4257(02)00018-4.

HABOUDANE, D.; TREMBLAY, N.; MILLER, J. R.; VIGNEAULT, P., 2008. Remote Estimation of Crop Chlorophyll Content Using Spectral Indices Derived from Hyperspectral Data. *IEEE Transactions on Geoscience and Remote Sensing*. Vol. 46, no. 2, pp. 423–437. Available from DOI: 10.1109/TGRS.2007.904836.

HAMID MUHAMMED, H., 2005. Hyperspectral Crop Reflectance Data for Characterising and Estimating Fungal Disease Severity in Wheat. *Biosystems Engineering.* Vol. 91, no. 1, pp. 9–20. Available from DOI: 10.1016/j.biosystemseng.2005.02.007.

HERNANDEZ, J.; LOBOS, G. A.; MATUS, I.; DEL POZO, A.; SILVA, P.; GALLEGUILLOS, M., 2015. Using Ridge Regression Models to Estimate Grain Yield from Field Spectral Data in Bread Wheat (*Triticum aestivum* L.) Grown under Three Water Regimes. *Remote Sensing.* Vol. 7, no. 2, pp. 2109–2126. Available from DOI: 10.3390/rs70202109.

HÖGY, P.; FRANZARING, J.; SCHWADORF, K.; BREUER, J.; SCHÜTZE, W.; FANGMEIER, A., 2010. Effects of free-air CO_2 enrichment on energy traits and seed quality of oilseed rape. *Agriculture, Ecosystems & Environment.* Vol. 139, no. 1–2, pp. 239–244. Available from DOI: 10.1016/j.agee.2010.08.009.

HUANG, L.; LI, T.; DING, C.; ZHAO, J.; ZHANG, D.; YANG, G., 2020. Diagnosis of the Severity of Fusarium Head Blight of Wheat Ears on the Basis of Image and Spectral Feature Fusion. *Sensors.* Vol. 20, no. 10, p. 2887. Available from DOI: 10.3390/s20102887.

HUANG, L.; WU, Z.; HUANG, W.; MA, H.; ZHAO, J., 2019. Identification of Fusarium Head Blight in Winter Wheat Ears Based on Fisher's Linear Discriminant Analysis and a Support Vector Machine. *Applied Sciences.* Vol. 9, no. 18, p. 3894. Available from DOI: 10.3390/app9183894.

HUANG, W.; GUAN, Q.; LUO, J.; ZHANG, J.; ZHAO, J.; LIANG, D.; HUANG, L.; ZHANG, D., 2014. New Optimized Spectral Indices for Identifying and Monitoring Winter Wheat Diseases. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* Vol. 7, no. 6, pp. 2516–2524. Available from DOI: 10.1109/JSTARS.2013.2294961.

HUANG, W.; LAMB, D. W.; NIU, Z.; ZHANG, Y.; LIU, L.; WANG, J., 2007. Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Precision Agriculture.* Vol. 8, pp. 187–197. Available from DOI: 10.1007/s11119-007-9038-9.

IHUOMA, S. O.; MADRAMOOTOO, C. A., 2017. Recent advances in crop water stress detection. *Computers and Electronics in Agriculture.* Vol. 141, pp. 267–275. Available from DOI: 10.1016/j.compag.2017.07.026.

INDAHL, U. G.; LILAND, K. H.; NÆS, T., 2009. Canonical partial least squares—a unified PLS approach to classification and regression problems. *Journal of Chemometrics*. Vol. 23, no. 9, pp. 495–504. Available from DOI: 10.1002/cem.1243.

JACONI, A.; POEPLAU, C.; RAMIREZ-LOPEZ, L.; VAN WESEMAEL, B.; DON, A., 2019. Log-ratio transformation is the key to determining soil organic carbon fractions with near-infrared spectroscopy. *European Journal of Soil Science*. Vol. 70, no. 1, pp. 127–139. Available from DOI: 10.1111/ejss.12761.

JACQUEMOUD, S.; USTIN, S. L., 2001. Leaf optical properties: A state of the art. In: *8th International Symposium of Physical Measurements & Signatures in Remote Sensing*. Aussois, pp. 223–332.

JAILLAIS, B.; ROUMET, P.; PINSON-GADAIS, L.; BERTRAND, D., 2015. Detection of *Fusarium* head blight contamination in wheat kernels by multivariate imaging. *Food Control*. Vol. 54, pp. 250–258. Available from DOI: 10.1016/j.foodcont.2015.01.048.

JANIK, L. J.; SKJEMSTAD, J.; SHEPHERD, K.; SPOUNCER, L., 2007. The prediction of soil carbon fractions using mid-infrared-partial least square analysis. *Soil Research*. Vol. 45, no. 2, pp. 73–81. Available from DOI: 10.1071/SR06083.

JANIK, L.; FORRESTER, S.; RAWSON, A., 2009. The prediction of soil chemical and physical properties from mid-infrared spectroscopy and combined partial least-squares regression and neural networks (PLS-NN) analysis. *Chemometrics and Intelligent Laboratory Systems*. Vol. 97, no. 2, pp. 179–188. Available from DOI: 10.1016/j.chemolab.2009.04.005.

JAY, S.; HADOUX, X.; GORRETTA, N.; RABATEL, G., 2014. Potential of hyperspectral imagery for nitrogen content retrieval in sugar beet leaves. In: *Proceedings International Conference of Agricultural Engineering*. Zurich, Switzerland, pp. 1–8.

JIN, X.; JIE, L.; WANG, S.; QI, H. J.; LI, S. W., 2018. Classifying Wheat Hyperspectral Pixels of Healthy Heads and *Fusarium* Head Blight Disease Using a Deep Neural Network in the Wild Field. *Remote Sensing*. Vol. 10, no. 3, p. 395. Available from DOI: 10.3390/rs10030395.

KAN, Z.-R.; LIU, W.-X.; LIU, W.-S.; LAL, R.; DANG, Y. P.; ZHAO, X.; ZHANG, H.-L., 2021. Mechanisms of soil organic carbon stability and its response to no-till: A global synthesis and perspective. *Global Change Biology*. Available from DOI: 10.1111/gcb.15968.

KARIMI, Y.; PRASHER, S.; MADANI, A.; KIM, S., 2008. Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations. *Canadian Biosystems Engineering*. Vol. 50, no. 7, pp. 13–20.

KHALEDI, N.; TAHERI, P.; RASTEGAR, M. F., 2017. Identification, virulence factors characterization, pathogenicity and aggressiveness analysis of *Fusarium* spp., causing wheat head blight in Iran. *European Journal of Plant Pathology*. Vol. 147, pp. 897–918. Available from DOI: 10.1007/s10658-016-1059-7.

KHAN, M. J.; KHAN, H. S.; YOUSAF, A.; KHURSHID, K.; ABBAS, A., 2018. Modern Trends in Hyperspectral Image Analysis: A Review. *IEEE Access*. Vol. 6, pp. 14118–14129. Available from DOI: 10.1109/access.2018.2812999.

KIM, Y.; GLENN, D. M.; PARK, J.; NGUGI, H. K.; LEHMAN, B. L., 2011. Hyperspectral image analysis for water stress detection of apple trees. *Computers and Electronics in Agriculture*. Vol. 77, no. 2, pp. 155–160. Available from DOI: 10.1016/j.compag.2011.04.008.

KNIPLING, E. B., 1970. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment*. Vol. 1, no. 3, pp. 155–159. Available from DOI: 10.1016/S0034-4257(70)80021-9.

KONG, W.; LIU, F.; ZHANG, C.; ZHANG, J.; FENG, H., 2016. Non-destructive determination of Malondialdehyde (MDA) distribution in oilseed rape leaves by laboratory scale NIR hyperspectral imaging. *Scientific Reports*. Vol. 6, no. 1. Available from DOI: 10.1038/srep35393.

KÖRSCHENS, M.; SCHULZ, E.; BEHM, R., 1990. Heißwasserlöslicher C und N im Boden als Kriterium für das N-Nachlieferungsvermögen. *Zentralblatt für Mikrobiologie*. Vol. 145, no. 4, pp. 305–311. Available from DOI: 10.1016/S0232-4393(11)80045-4.

KOSOVÁ, K.; KLÍMA, M.; VÍTAMVÁS, P.; PRÁŠIL, I. T., 2018. Odezva vybraných odrůd řepky na sucho a následná regenerace [Response of selected oilseed rape cultivars to drought and subsequent recovery]. *Úroda*. Vol. 66, pp. 19–25.

KRUSCHKE, J. K.; LIDDELL, T. M., 2017. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, pp. 1–29. Available from DOI: 10.3758/s13423-016-1221-4.

KUANG, B.; MAHMOOD, H. S.; QURAISHI, M. Z.; HOOGMOED, W. B.; MOUAZEN, A. M.; HENTEN, E. J. van, 2012. Sensing Soil Properties in the Laboratory, *In Situ*, and On-Line: A Review. *Advances in Agronomy*. Vol. 114, pp. 155–223. Available from DOI: 10.1016/B978-0-12-394275-3.00003-1.

KUCKENBERG, J.; TARTACHNYK, I.; NOGA, G., 2009. Temporal and spatial changes of chlorophyll fluorescence as a basis for early and precise detection of leaf rust and powdery mildew infections in wheat leaves. *Precision Agriculture*. Vol. 10, pp. 34–44. Available from DOI: 10.1007/s11119-008-9082-0.

KUHN, M.; JOHNSON, K., 2013. *Applied Predictive Modeling*. New York: Springer. Available from DOI: 10.1007/978-1-4614-6849-3.

KUMAR, A.; BHARTI, V.; KUMAR, V.; MEENA, P.; SURESH, G., 2017. Hyperspectral imaging applications in rapeseed and mustard farming. *Journal of Oilseeds Research*. Vol. 34, no. 1, pp. 1–8.

KYNČLOVÁ, P.; FILZMOSER, P.; HRON, K., 2015. Modeling compositional time series with vector autoregressive models. *Journal of Forecasting*. Vol. 34, no. 4, pp. 303–314. Available from DOI: 10.1002/for.2336.

LIDDELL, T. M.; KRUSCHKE, J. K., 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*. Vol. 79, pp. 328–348. Available from DOI: 10.1016/j.jesp.2018.08.009.

LINKE, R.; RICHTER, K.; HAUMANN, J.; SCHNEIDER, W.; WEIHS, P., 2008. Occurrence of repeated drought events: can repetitive stress situations and recovery from drought be traced with leaf reflectance? *Periodicum biologorum*. Vol. 110, no. 3, pp. 219–229.

LIU, L.; DONG, Y.; HUANG, W.; DU, X.; MA, H., 2020. Monitoring Wheat Fusarium Head Blight Using Unmanned Aerial Vehicle Hyperspectral Imagery. *Remote Sensing*. Vol. 12, no. 22, p. 3811. Available from DOI: 10.3390/rs12223811.

LIU, L.; DONG, Y.; HUANG, W.; DU, X.; REN, B.; HUANG, L.; ZHENG, Q.; MA, H., 2020. A Disease Index for Efficiently Detecting Wheat Fusarium Head Blight Using Sentinel-2 Multispectral Imagery. *IEEE Access*. Vol. 8, pp. 52181–52191. Available from DOI: 10.1109/ACCESS.2020.2980310.

LOBELL, D. B.; SCHLENKER, W.; COSTA-ROBERTS, J., 2011. Climate Trends and Global Crop Production Since 1980. *Science*. Vol. 333, no. 6042, pp. 616–620. Available from DOI: 10.1126/science.1204531.

LOWE, A.; HARRISON, N.; FRENCH, A. P., 2017. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods*. Vol. 13, p. 80. ISSN 1746-4811. Available from DOI: 10.1186/s13007-017-0233-z.

LU, B.; HE, Y.; DAO, P. D., 2019. Comparing the Performance of Multispectral and Hyperspectral Images for Estimating Vegetation Properties. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. Vol. 12, no. 6, pp. 1784–1797. Available from DOI: 10.1109/JSTARS.2019.2910558.

MA, H.; HUANG, W.; JING, Y.; PIGNATTI, S.; LANEVE, G.; DONG, Y.; YE, H.; LIU, L.; GUO, A.; JIANG, J., 2020. Identification of *Fusarium* Head Blight in Winter Wheat Ears Using Continuous Wavelet Analysis. *Sensors*. Vol. 20, no. 1, p. 20. Available from DOI: 10.3390/s20010020.

MAC ARTHUR, A.; ROBINSON, I., 2015. A critique of field spectroscopy and the challenges and opportunities it presents for remote sensing for agriculture, ecosystems, and hydrology. In: NEALE, C. M.; MALTESE, A. (eds.). *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVII*. Toulouse, France. Vol. 9637, p. 963705. Available from DOI: 10.1117/12.2201046.

MADARI, B. E.; REEVES III, J. B.; COELHO, M. R.; MACHADO, P. L.; DE-POLLI, H.; COELHO, R. M.; BENITES, V. M.; SOUZA, L. F.; MCCARTY, G. W., 2005. Mid- and Near-Infrared Spectroscopic Determination of Carbon in a Diverse Set of Soils from the Brazilian National Soil Collection. *Spectroscopy Letters*. Vol. 38, no. 6, pp. 721–740. Available from DOI: 10.1080/00387010500315876.

MAHLEIN, A.; KUSKA, M.; THOMAS, S.; BOHNENKAMP, D.; ALISAAC, E.; BEHMANN, J.; WAHABZADA, M.; KERSTING, K., 2017. Plant disease detection by hyperspectral imaging: from the lab to the field. *Advances in Animal Biosciences*. Vol. 8, no. Special Issue 2, pp. 238–243. Available from DOI: 10.1017/S2040470017001248.

MAHLEIN, A.-K., 2016. Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*. Vol. 100, no. 2, pp. 241–251. Available from DOI: 10.1094/PDIS-03-15-0340-FE.

MAHLEIN, A.-K.; ALISAAC, E.; AL MASRI, A.; BEHMANN, J.; DEHNE, H.-W.; OERKE, E.-C., 2019. Comparison and Combination of Thermal, Fluorescence, and Hyperspectral Imaging for Monitoring *Fusarium* Head Blight of Wheat on Spikelet Scale. *Sensors*. Vol. 19, no. 10, p. 2281. Available from DOI: 10.3390/s19102281.

MAHLEIN, A.-K.; OERKE, E.-C.; STEINER, U.; DEHNE, H.-W., 2012. Recent advances in sensing plant diseases for precision crop protection. *European Journal of Plant Pathology.* Vol. 133, pp. 197–209. Available from DOI: 10.1007/s10658-011-9878-z.

MAIN, R.; CHO, M. A.; MATHIEU, R.; O'KENNEDY, M. M.; RAMOELO, A.; KOCH, S., 2011. An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS Journal of Photogrammetry and Remote Sensing.* Vol. 66, no. 6, pp. 751–761. Available from DOI: 10.1016/j.isprsjprs.2011.08.001.

MAJIDI, M.; RASHIDI, F.; SHARAFI, Y., 2015. Physiological traits related to drought tolerance in *Brassica. International Journal of Plant Production.* Vol. 9, no. 4.

MARIOTTO, I.; THENKABAIL, P. S.; HUETE, A.; SLONECKER, E. T.; PLATONOV, A., 2013. Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HyspIRI mission. *Remote Sensing of Environment.* Vol. 139, pp. 291–305. Available from DOI: 10.1016/j.rse.2013.08.002.

MARSHALL, M.; THENKABAIL, P., 2015. Advantage of hyperspectral EO-1 Hyperion over multispectral IKONOS, GeoEye-1, WorldView-2, Landsat ETM+, and MODIS vegetation indices in crop biomass estimation. *ISPRS Journal of Photogrammetry and Remote Sensing.* Vol. 108, pp. 205–218. Available from DOI: 10.1016/j.isprsjprs.2015.08.001.

MARTINELLI, F.; SCALENGHE, R.; DAVINO, S.; PANNO, S.; SCUDERI, G.; RUISI, P.; VILLA, P.; STROPPIANA, D.; BOSCHETTI, M.; GOULART, L. R.; DAVIS, C. E.; DANDEKAR, A. M., 2015. Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development.* Vol. 35, no. 1, pp. 1–25. Available from DOI: 10.1007/s13593-014-0246-1.

MCCARTY, G.; REEVES, J.; REEVES, V.; FOLLETT, R.; KIMBLE, J., 2002. Mid-Infrared and Near-Infrared Diffuse Reflectance Spectroscopy for Soil Carbon Measurement. *Soil Science Society of America Journal.* Vol. 66, no. 2, pp. 640–646. Available from DOI: doi.org/10.2136/sssaj2002.6400a.

MCCULLAGH, P., 1980. Regression Models for Ordinal Data. *Journal of the Royal Statistical Society: Series B (Methodological).* Vol. 42, no. 2, pp. 109–127. Available from DOI: 10.1111/j.2517-6161.1980.tb01109.x.

MCMULLEN, M.; BERGSTROM, G.; DE WOLF, E.; DILL-MACKY, R.; HERSH-MAN, D.; SHANER, G.; VAN SANFORD, D., 2012. A Unified Effort to Fight an Enemy of Wheat and Barley: Fusarium Head Blight. *Plant Disease*. Vol. 96, no. 12, pp. 1712–1728. Available from DOI: 10.1094/PDIS-03-12-0291-FE.

MEER, F. van der, 2018. Near-infrared laboratory spectroscopy of mineral chemistry: A review. *International Journal of Applied Earth Observation and Geoinformation*. Vol. 65, pp. 71–78. ISSN 0303-2434. Available from DOI: 10.1016/j.jag.2017. 10.004.

MEHMOOD, T.; LILAND, K. H.; SNIPEN, L.; SÆBØ, S., 2012. A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*. Vol. 118, pp. 62–69. Available from DOI: 10.1016/j.chemolab.2012.07.010.

MEIJ, B. van der; KOOISTRA, L.; SUOMALAINEN, J.; BAREL, J. M.; DEYN, D. B. D., 2018. *Data from: Remote sensing of plant trait responses to field-based plant–soil feedback using UAV-based optical sensors*. Dryad Digital Repository. Available from DOI: 10.5061/dryad.75k1d.

MEIJ, B. van der; KOOISTRA, L.; SUOMALAINEN, J.; BAREL, J. M.; DEYN, G. B. D., 2017. Remote sensing of plant trait responses to field-based plant–soil feedback using UAV-based optical sensors. *Biogeosciences*. Vol. 14, no. 3, pp. 733–749. Available from DOI: 10.5194/bg-14-733-2017.

MERZLYAK, M. N.; GITELSON, A. A.; CHIVKUNOVA, O. B.; RAKITIN, V. Y., 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum*. Vol. 106, no. 1, pp. 135–141. Available from DOI: 10.1034/j.1399-3054.1999.106119.x.

MEVIK, B.-H.; WEHRENS, R.; LILAND, K. H., 2019. *pls: Partial Least Squares and Principal Component Regression*. Available also from: https://CRAN.R-project.org/package=pls. R package version 2.7-2.

MEWES, T.; FRANKE, J.; MENZ, G., 2011. Spectral requirements on airborne hyperspectral remote sensing data for wheat disease detection. *Precision Agriculture*. Vol. 12, p. 795. Available from DOI: 10.1007/s11119-011-9222-9.

MIAO, Y.; MULLA, D. J.; RANDALL, G. W.; VETSCH, J. A.; VINTILA, R., 2009. Combining chlorophyll meter readings and high spatial resolution remote sensing images for in-season site-specific nitrogen management of corn. *Precision Agriculture*. Vol. 10, no. 1, pp. 45–62. Available from DOI: 10.1007/s11119-008-9091-z.

MIELNICZUK, E.; SKWARYŁO-BEDNARZ, B., 2020. Fusarium Head Blight, Mycotoxins and Strategies for Their Reduction. *Agronomy*. Vol. 10, no. 4, p. 509. Available from DOI: 10.3390/agronomy10040509.

MILTON, E. J.; SCHAEPMAN, M. E.; ANDERSON, K.; KNEUBÜHLER, M.; FOX, N., 2009. Progress in field spectroscopy. *Remote Sensing of Environment*. Vol. 113, S92–S109. Available from DOI: 10.1016/j.rse.2007.08.001.

MINASNY, B.; MCBRATNEY, A. B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*. Vol. 32, no. 9, pp. 1378–1388. Available from DOI: 10.1016/j.cageo.2005.12.009.

MISHRA, P.; ASAARI, M. S. M.; HERRERO-LANGREO, A.; LOHUMI, S.; DIEZMA, B.; SCHEUNDERS, P., 2017. Close range hyperspectral imaging of plants: A review. *Biosystems Engineering*. Vol. 164, pp. 49–67. Available from DOI: 10.1016/j.biosystemseng.2017.09.009.

MOHARANA, S.; DUTTA, S., 2016. Spatial variability of chlorophyll and nitrogen content of rice from hyperspectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*. Vol. 122, pp. 17–29. Available from DOI: 10.1016/j.isprsjprs.2016.09.002.

MONTESINOS-LÓPEZ, O. A.; MONTESINOS-LÓPEZ, A.; CROSSA, J.; DE LOS CAMPOS, G.; ALVARADO, G.; SUCHISMITA, M.; RUTKOSKI, J.; GONZÁLEZ-PÉREZ, L.; BURGUEÑO, J., 2017. Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods*. Vol. 13, p. 4. Available from DOI: 10.1186/s13007-016-0154-2.

MORIN, M.; LAWRENCE, R.; REPASKY, K. S.; STERLING, T. M.; MCCANN, C.; POWELL, S., 2017. Agreement analysis and spatial sensitivity of multispectral and hyperspectral sensors in detecting vegetation stress at management scales. *Journal of Applied Remote Sensing*. Vol. 11, no. 4, p. 046025. Available from DOI: 10.1117/1.JRS.11.046025.

MULLA, D. J., 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*. Vol. 114, no. 4, pp. 358–371. Available from DOI: 10.1016/j.biosystemseng.2012.08.009.

MUNNÉ-BOSCH, S.; ALEGRE, L., 2004. Die and let live: leaf senescence contributes to plant survival under drought stress. *Functional Plant Biology*. Vol. 31, no. 3, pp. 203–216. Available from DOI: 10.1071/FP03236.

NAKHFOROOSH, A.; BODEWEIN, T.; FIORANI, F.; BODNER, G., 2016. Identification of Water Use Strategies at Early Growth Stages in Durum Wheat from Shoot Phenotyping and Physiological Measurements. *Frontiers in Plant Science.* Vol. 7, p. 1155. Available from DOI: 10.3389/fpls.2016.01155.

NANSEN, C., 2012. Use of Variogram Parameters in Analysis of Hyperspectral Imaging Data Acquired from Dual-Stressed Crop Leaves. *Remote Sensing.* Vol. 4, no. 1, pp. 180–193. Available from DOI: 10.3390/rs4010180.

NAUMANN, G.; ALFIERI, L.; WYSER, K.; MENTASCHI, L.; BETTS, R.; CARRAO, H.; SPINONI, J.; VOGT, J.; FEYEN, L., 2018. Global Changes in Drought Conditions Under Different Levels of Warming. *Geophysical Research Letters.* Vol. 45, no. 7, pp. 3285–3296. Available from DOI: 10.1002/2017GL076521.

NG, W.; MINASNY, B.; MALONE, B.; FILIPPI, P., 2018. In search of an optimum sampling algorithm for prediction of soil properties from infrared spectra. *PeerJ.* Vol. 6, e5722. Available from DOI: 10.7717/peerj.5722.

NOCITA, M.; STEVENS, A.; WESEMAEL, B. van; AITKENHEAD, M.; BACHMANN, M.; BARTHÈS, B.; DOR, E. B.; BROWN, D. J.; CLAIROTTE, M.; CSORBA, A.; DARDENNE, P.; DEMATTÊ, J. A. M.; GENOT, V.; CESAR, G.; MARIA, K.; MONTANARELLA, L.; NOON, C.; RAMIREZ-LOPEZ, L.; ROBERTSON, J.; SAKAI, H.; SORIANO-DISLA, H. M.; SHEPHERD, K. D.; STENBERG, B.; TOWETT, E. K.; VARGAS, R.; WETTERLIND, J., 2015. Chapter Four - Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring. *Advances in Agronomy.* Vol. 132, pp. 139–159. Available from DOI: 10.1016/bs.agron.2015.02.002.

NOSEK, B. A.; EBERSOLE, C. R.; DEHAVEN, A.; MELLOR, D., 2018. The preregistration revolution. *PNAS.* Vol. 115, no. 11, pp. 2600–2606. Available from DOI: 10.1073/pnas.1708274114.

OLLINGER, S. V., 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist.* Vol. 189, no. 2, pp. 375–394. Available from DOI: 10.1111/j.1469-8137.2010.03536.x.

PAGE, K.; DALAL, R.; DANG, Y., 2013. How useful are MIR predictions of total, particulate, humus, and resistant organic carbon for examining changes in soil carbon stocks in response to different crop management? A case study. *Soil Research.* Vol. 51, no. 8, pp. 719–725. Available from DOI: 10.1071/SR13064.

PAUSTIAN, K.; COLLIER, S.; BALDOCK, J.; BURGESS, R.; CREQUE, J.; DELONGE, M.; DUNGAIT, J.; ELLERT, B.; FRANK, S.; GODDARD, T.; GOVAERTS, B.; GRUNDY, M.; HENNING, M.; IZAURRALDE, R. C.; MADARAS, M.; MCCONKEY, B.; PORZIG, E.; RICE, C.; SEARLE, R.; SEAVY, N.; SKALSKY, R.; MULHERN, W.; JAHN, M., 2019. Quantifying carbon for agricultural soil management: from the current status toward a global soil information system. *Carbon Management*. Vol. 10, no. 6, pp. 567–587. Available from DOI: 10.1080/17583004.2019.1633231.

PEDERSEN, S.; LIND, K., 2017. Precision Agriculture – From Mapping to Site-Specific Application. In: PEDERSEN, S. M.; LIND, K. M. (eds.). *Precision Agriculture: Technology and Economic Perspectives*. Cham: Springer International Publishing AG, pp. 1–20.

PEÑUELAS, J.; BARET, F.; FILELLA, I., 1995. Semi-empirical indices to assess carotenoids/chlorophyll *a* ratio from leaf spectral reflectance. *Photosynthetica*. Vol. 31, no. 2, pp. 221–230.

PEÑUELAS, J.; GAMON, J.; FREDEEN, A.; MERINO, J.; FIELD, C., 1994. Reflectance indices associated with physiological changes in nitrogen-and water-limited sunflower leaves. *Remote sensing of Environment*. Vol. 48, no. 2, pp. 135–146. Available from DOI: 10.1016/0034-4257(94)90136-8.

PERRY, E. M.; DAVENPORT, J. R., 2007. Spectral and spatial differences in response of vegetation indices to nitrogen treatments on apple. *Computers and Electronics in Agriculture*. Vol. 59, no. 1–2, pp. 56–65. Available from DOI: 10.1016/j.compag.2007.05.002.

PERRY, E. M.; ROBERTS, D. A., 2008. Sensitivity of Narrow-Band and Broad-Band Indices for Assessing Nitrogen Availability and Water Stress in an Annual Crop. *Agronomy Journal*. Vol. 100, no. 4, pp. 1211–1219. Available from DOI: 10.2134/agronj2007.0306.

PETEINATOS, G. G.; KORSAETH, A.; BERGE, T. W.; GERHARDS, R., 2016. Using Optical Sensors to Identify Water Deprivation, Nitrogen Shortage, Weed Presence and Fungal Infection in Wheat. *Agriculture*. Vol. 6, no. 2, p. 24. Available from DOI: 10.3390/agriculture6020024.

PICCOLO, S. R.; FRAMPTON, M. B., 2016. Tools and techniques for computational reproducibility. *GigaScience*. Vol. 5, no. 1, p. 30. ISSN 2047-217X. Available from DOI: 10.1186/s13742-016-0135-4.

PIEKARCZYK, J.; SULEWSKA, H.; SZYMAŃSKA, G., 2011. Winter oilseed-rape yield estimates from hyperspectral radiometer measurements. *Quaestiones Geographicae*. Vol. 30, no. 1, pp. 77–84. Available from DOI: 10.2478/v10117-011-0007-z.

PINTER Jr, P. J.; JACKSON, R. D.; ELAINE EZRA, C.; GAUSMAN, H. W., 1985. Sun-angle and canopy-architecture effects on the spectral reflectance of six wheat cultivars. *International Journal of Remote Sensing*. Vol. 6, no. 12, pp. 1813–1825. Available from DOI: 10.1080/01431168508948330.

QIU, R.; YANG, C.; MOGHIMI, A.; ZHANG, M.; STEFFENSON, B. J.; HIRSCH, C. D., 2019. Detection of Fusarium Head Blight in Wheat Using a Deep Neural Network and Color Imaging. *Remote Sensing*. Vol. 11, no. 22, p. 2658. Available from DOI: 10.3390/rs11222658.

RAPAPORT, T.; HOCHBERG, U.; SHOSHANY, M.; KARNIELI, A.; RACHMILEVITCH, S., 2015. Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. *ISPRS Journal of Photogrammetry and Remote Sensing*. Vol. 109, pp. 88–97. Available from DOI: 10.1016/j.isprsjprs.2015.09.003.

RAZA, M. A. S.; SHAHID, A. M.; SALEEM, M. F.; KHAN, I. H.; AHMAD, S.; ALI, M.; IQBAL, R., 2017. Effects and management strategies to mitigate drought stress in oilseed rape (*Brassica napus* L.): a review. *Zemdirbyste-Agriculture*. Vol. 104, no. 1, pp. 85–94. Available from DOI: 10.13080/z-a.2017.104.012.

REBETZKE, G. J.; FISCHER, R. (A.; HERWAARDEN, A. F. van; BONNETT, D. G.; CHENU, K.; RATTEY, A. R.; FETTELL, N. A., 2014. Plot size matters: interference from intergenotypic competition in plant phenotyping studies. *Functional Plant Biology*. Vol. 41, no. 2, pp. 107–118. Available from DOI: 10.1071/FP13177.

REEVES, D., 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. *Soil and Tillage Research*. Vol. 43, no. 1–2, pp. 131–167. Available from DOI: 10.1016/S0167-1987(97)00038-X.

REEVES, J. B.; MCCARTY, G. W.; REEVES, V. B., 2001. Mid-infrared Diffuse Reflectance Spectroscopy for the Quantitative Analysis of Agricultural Soils. *Journal of Agricultural and Food Chemistry*. Vol. 49, no. 2, pp. 766–772. Available from DOI: 10.1021/jf0011283.

REEVES III, J. B., 2010. Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done? *Geoderma.* Vol. 158, no. 1–2, pp. 3–14. Available from DOI: 10.1016/j.geoderma.2009.04.005.

RINNAN, Å.; BERG, F. van den; ENGELSEN, S. B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry.* Vol. 28, no. 10, pp. 1201–1222. Available from DOI: 10.1016/j.trac.2009.07.007.

ROCHE, D., 2015. Stomatal conductance is essential for higher yield potential of C_3 crops. *Critical Reviews in Plant Sciences.* Vol. 34, no. 4, pp. 429–453. Available from DOI: 10.1080/07352689.2015.1023677.

RODRÍGUEZ-PÉREZ, J. R.; RIAÑO, D.; CARLISLE, E.; USTIN, S.; SMART, D. R., 2007. Evaluation of Hyperspectral Reflectance Indexes to Detect Grapevine Water Status in Vineyards. *American Journal of Enology and Viticulture.* Vol. 58, pp. 302–317.

RÖMER, C.; WAHABZADA, M.; BALLVORA, A.; PINTO, F.; ROSSINI, M.; PANIGADA, C.; BEHMANN, J.; LÉON, J.; THURAU, C.; BAUCKHAGE, C.; KERSTING, K.; RASCHER, U.; PLÜMER, L., 2012. Early drought stress detection in cereals: simplex volume maximisation for hyperspectral image analysis. *Functional Plant Biology.* Vol. 39, no. 11, pp. 878–890. Available from DOI: 10.1071/FP12060.

ROSENTHAL, R., 1979. The file drawer problem and tolerance for null results. *Psychological Bulletin.* Vol. 86, no. 3, pp. 638–641. Available from DOI: 10.1037/0033-2909.86.3.638.

SABAGH, A. E.; HOSSAIN, A.; BARUTÇULAR, C.; ISLAM, M. S.; RATNASEKERA, D.; KUMAR, N.; MEENA, R. S.; GHARIB, H. S.; SANEOKA, H.; SILVA, J. A. T. da, 2019. Drought and salinity stress management for higher and sustainable canola ('Brassica napus' L.) production: A critical review. *Australian Journal of Crop Science.* Vol. 13, no. 1, pp. 88–96. Available from DOI: 10.21475/ajcs.19.13.01.p1284.

SACCON, F. A.; PARCEY, D.; PALIWAL, J.; SHERIF, S. S., 2017. Assessment of *Fusarium* and Deoxynivalenol Using Optical Methods. *Food and Bioprocess Technology.* Vol. 10, pp. 34–50. Available from DOI: 10.1007/s11947-016-1788-9.

SAEYS, W.; MOUAZEN, A. M.; RAMON, H., 2005. Potential for Onsite and Online Analysis of Pig Manure Using Visible and Near Infrared Reflectance Spectroscopy. *Biosystems Engineering*. Vol. 91, no. 4, pp. 393–402. Available from DOI: 10.1016/j.biosystemseng.2005.05.001.

SALVATIER, J.; WIECKI, T. V.; FONNESBECK, C., 2016. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*. Vol. 2, e55. Available from DOI: 10.7717/peerj-cs.55.

SANDVE, G. K.; NEKRUTENKO, A.; TAYLOR, J.; HOVIG, E., 2013. Ten Simple Rules for Reproducible Computational Research. *PLoS Computational Biology*. Vol. 9, no. 10, e1003285. Available from DOI: 10.1371/journal.pcbi.1003285.

SANKARAN, S.; KHOT, L. R.; ESPINOZA, C. Z.; JAROLMASJED, S.; SATHUVALLI, V. R.; VANDEMARK, G. J.; MIKLAS, P. N.; CARTER, A. H.; PUMPHREY, M. O.; KNOWLES, N. R.; PAVEK, M. J., 2015. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *European Journal of Agronomy*. Vol. 70, pp. 112–123. Available from DOI: 10.1016/j.eja.2015.07.004.

SANKEY, J. B.; BROWN, D. J.; BERNARD, M. L.; LAWRENCE, R. L., 2008. Comparing local vs. global visible and near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. *Geoderma*. Vol. 148, no. 2, pp. 149–158. Available from DOI: 10.1016/j.geoderma.2008.09.019.

SEYBOLD, C. A.; FERGUSON, R.; WYSOCKI, D.; BAILEY, S.; ANDERSON, J.; NESTER, B.; SCHOENEBERGER, P.; WILLS, S.; LIBOHOVA, Z.; HOOVER, D.; THOMAS, P., 2019. Application of Mid-Infrared Spectroscopy in Soil Survey. *Soil Science Society of America Journal*. Vol. 83, no. 6, pp. 1746–1759. Available from DOI: 10.2136/sssaj2019.06.0205.

SHEN, G.; CAO, Y.; YIN, X.; DONG, F.; XU, J.; SHI, J.; LEE, Y.-W., 2022. Rapid and nondestructive quantification of deoxynivalenol in individual wheat kernels using near-infrared hyperspectral imaging and chemometrics. *Food Control*. Vol. 131, p. 108420. Available from DOI: 10.1016/j.foodcont.2021.108420.

SIMS, D. A.; GAMON, J. A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*. Vol. 81, no. 2–3, pp. 337–354. Available from DOI: 10.1016/S0034-4257(02)00010-X.

SMITH, G. M.; MILTON, E. J., 1999. The use of the empirical line method to calibrate remotely sensed data to reflectance. *International Journal of Remote Sensing.* Vol. 20, no. 13, pp. 2653–2662. Available from DOI: 10.1080/014311699211994.

SMITH, P.; SOUSSANA, J.-F.; ANGERS, D.; SCHIPPER, L.; CHENU, C.; RASSE, D. P.; BATJES, N. H.; EGMOND, F. van; MCNEILL, S.; KUHNERT, M.; ARIAS-NAVARRO, C.; OLESEN, J. E.; CHIRINDA, N.; FORNARA, D.; WOLLENBERG, E.; ÁLVARO-FUENTES, J.; SANZ-COBENA, A.; KLUMPP, K., 2019. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Global Change Biology.* Available from DOI: 10.1111/gcb.14815.

SÖDERSTRÖM, M.; BÖRJESSON, T., 2013. Within-field variation in deoxynivalenol (DON) contents in oats. In: STAFFORD, J. V. (ed.). *Precision agriculture '13.* Lleida, Spain: Wageningen Academic Publishers, pp. 329–334.

SORIANO-DISLA, J. M.; JANIK, L. J.; VISCARRA ROSSEL, R. A.; MACDONALD, L. M.; MCLAUGHLIN, M. J., 2014. The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. *Applied Spectroscopy Reviews.* Vol. 49, no. 2, pp. 139–186. Available from DOI: 10.1080/05704928.2013.811081.

STAFFORD, J. V., 2000. Implementing Precision Agriculture in the 21st Century. *Journal of Agricultural Engineering Research.* Vol. 76, no. 3, pp. 267–275. Available from DOI: 10.1006/jaer.2000.0577.

STEINER, B.; BUERSTMAYR, M.; MICHEL, S.; SCHWEIGER, W.; LEMMENS, M.; BUERSTMAYR, H., 2017. Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Tropical Plant Pathology.* Vol. 42, pp. 165–174. Available from DOI: 10.1007/s40858-017-0127-7.

STEINER, U.; BÜRLING, K.; OERKE, E.-C., 2008. Sensorik für einen präzisierten Pflanzenschutz. *Gesunde Pflanzen.* Vol. 60, pp. 131–141. Available from DOI: 10.1007/s10343-008-0194-2.

STENBERG, B.; ROSSEL, R. A. V.; MOUAZEN, A. M.; WETTERLIND, J., 2010. Visible and Near Infrared Spectroscopy in Soil Science. *Advances in Agronomy.* Vol. 107, pp. 163–215. Available from DOI: 10.1016/S0065-2113(10)07005-7.

STUMPE, B.; WEIHERMÜLLER, L.; MARSCHNER, B., 2011. Sample preparation and selection for qualitative and quantitative analyses of soil organic carbon with

mid-infrared reflectance spectroscopy. *European Journal of Soil Science*. Vol. 62, no. 6, pp. 849–862. Available from DOI: 10.1111/j.1365-2389.2011.01401.x.

SU, W.-H.; ZHANG, J.; YANG, C.; PAGE, R.; SZINYEI, T.; HIRSCH, C. D.; STEFFENSON, B. J., 2021. Automatic Evaluation of Wheat Resistance to Fusarium Head Blight Using Dual Mask-RCNN Deep Learning Frameworks in Computer Vision. *Remote Sensing*. Vol. 13, no. 1, p. 26. Available from DOI: 10.3390/rs13010026.

SUN, C.; LI, C.; ZHANG, C.; HAO, L.; SONG, M.; LIU, W.; ZHANG, Y., 2018. Reflectance and biochemical responses of maize plants to drought and re-watering cycles. *Annals of Applied Biology*. Vol. 172, no. 3, pp. 332–345. Available from DOI: 10.1111/aab.12423.

SZUCS, D.; IOANNIDIS, J., 2017. When Null Hypothesis Significance Testing Is Unsuitable for Research: A Reassessment. *Frontiers in Human Neuroscience*. Vol. 11, p. 390. Available from DOI: 10.3389/fnhum.2017.00390.

TESFAMARIAM, E. H.; ANNANDALE, J. G.; STEYN, J. M., 2010. Water Stress Effects on Winter Canola Growth and Yield. *Agronomy Journal*. Vol. 102, no. 2, pp. 658–666. Available from DOI: 10.2134/agronj2008.0043.

THENKABAIL, P.; ANEECE, I.; TELUGUNTLA, P.; OLIPHANT, A., 2021. Hyperspectral Narrowband Data Propel Gigantic Leap in the Earth Remote Sensing. *Photogrammetric Engineering and Remote Sensing*. Vol. 87, no. 7, pp. 461–467. Available from DOI: 10.14358/PERS.87.7.461.

THOMAS, B. W.; WHALEN, J. K.; SHARIFI, M.; CHANTIGNY, M.; ZEBARTH, B. J., 2016. Labile organic matter fractions as early-season nitrogen supply indicators in manure-amended soils. *Journal of Plant Nutrition and Soil Science*. Vol. 179, no. 1, pp. 94–103. Available from DOI: 10.1002/jpln.201400532.

THOMAS, S.; KUSKA, M. T.; BOHNENKAMP, D.; BRUGGER, A.; ALISAAC, E.; WAHABZADA, M.; BEHMANN, J.; MAHLEIN, A.-K., 2018. Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. *Journal of Plant Diseases and Protection*. Vol. 125, no. 1, pp. 5–20. Available from DOI: 10.1007/s41348-017-0124-6.

URBAN, M. O.; VAŠEK, J.; KLÍMA, M.; KRTKOVÁ, J.; KOSOVÁ, K.; PRÁŠIL, I. T.; VÍTÁMVÁS, P., 2017. Proteomic and physiological approach reveals drought-induced changes in rapeseeds: Water-saver and water-spender strategy. *Journal of Proteomics*. Vol. 152, pp. 188–205. Available from DOI: 10.1016/j.jprot.2016.11.004.

VAUGHAN, M.; BACKHOUSE, D.; DEL PONTE, E. M., 2016. Climate change impacts on the ecology of *Fusarium graminearum* species complex and susceptibility of wheat to *Fusarium* head blight: a review. *World Mycotoxin Journal*. Vol. 9, no. 5, pp. 685–700. Available from DOI: 10.3920/WMJ2016.2053.

VERRELST, J.; ALONSO, L.; CAMPS-VALLS, G.; DELEGIDO, J.; MORENO, J., 2012. Retrieval of Vegetation Biophysical Parameters Using Gaussian Process Techniques. *IEEE Transactions on Geoscience and Remote Sensing*. Vol. 50, no. 5, pp. 1832–1843. Available from DOI: 10.1109/TGRS.2011.2168962.

VIRNODKAR, S. S.; PACHGHARE, V. K.; PATIL, V.; JHA, S. K., 2020. Remote sensing and machine learning for crop water stress determination in various crops: a critical review. *Precision Agriculture*. Vol. 21, pp. 1–35. Available from DOI: 10.1007/s11119-020-09711-9.

VISCARRA ROSSEL, R.; WALVOORT, D.; MCBRATNEY, A.; JANIK, L. J.; SKJEMSTAD, J., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*. Vol. 131, no. 1-2, pp. 59–75. Available from DOI: 10.1016/j.geoderma.2005.03.007.

VISSER, M. D.; MCMAHON, S. M.; MEROW, C.; DIXON, P. M.; RECORD, S.; JONGEJANS, E., 2015. Speeding Up Ecological and Evolutionary Computations in R; Essentials of High Performance Computing for Biologists. *PLoS Computational Biology*. Vol. 11, no. 3, e1004140. Available from DOI: 10.1371/journal.pcbi.1004140.

WERFF, H. van der; MEIJDE, M. van der; JANSMA, F.; MEER, F. van der; GROOTHUIS, G. J., 2008. A Spatial-Spectral Approach for Visualization of Vegetation Stress Resulting from Pipeline Leakage. *Sensors*. Vol. 8, no. 6, pp. 3733–3743. Available from DOI: 10.3390/s8063733.

WEST, J. S.; BRAVO, C.; OBERTI, R.; LEMAIRE, D.; MOSHOU, D.; MCCARTNEY, H. A., 2003. The Potential of Optical Canopy Measurement for Targeted Control of Field Crop Diseases. *Annual Review of Phytopathology*. Vol. 41, pp. 593–614. Available from DOI: 10.1146/annurev.phyto.41.121702.103726.

WEST, J. S.; CANNING, G. G.; PERRYMAN, S. A.; KING, K., 2017. Novel Technologies for the detection of Fusarium head blight disease and airborne inoculum. *Tropical Plant Pathology*. Vol. 42, pp. 203–209. Available from DOI: 10.1007/s40858-017-0138-4.

WHETTON, R. L.; HASSALL, K. L.; WAINE, T. W.; MOUAZEN, A. M., 2018. Hyperspectral measurements of yellow rust and fusarium head blight in cereal crops: Part 1: Laboratory study. *Biosystems Engineering*. Vol. 166, pp. 101–115. Available from DOI: 10.1016/j.biosystemseng.2017.11.008.

WHETTON, R. L.; WAINE, T. W.; MOUAZEN, A. M., 2018. Hyperspectral measurements of yellow rust and fusarium head blight in cereal crops: Part 2: Online field measurement. *Biosystems Engineering*. Vol. 167, pp. 144–158. Available from DOI: 10.1016/j.biosystemseng.2018.01.004.

XIA, J.; YANG, Y. W.; CAO, H. X.; ZHANG, W.; XU, L.; WANG, Q.; KE, Y.; ZHANG, W.; GE, D.; HUANG, B., 2018. Hyperspectral Identification and Classification of Oilseed Rape Waterlogging Stress Levels Using Parallel Computing. *IEEE Access*. Vol. 6, pp. 57663–57675. Available from DOI: 10.1109/ACCESS.2018.2873689.

XIAO, Y.; DONG, Y.; HUANG, W.; LIU, L.; MA, H.; YE, H.; WANG, K., 2020. Dynamic Remote Sensing Prediction for Wheat Fusarium Head Blight by Combining Host and Habitat Conditions. *Remote Sensing*. Vol. 12, no. 18, p. 3046. Available from DOI: 10.3390/rs12183046.

XUE, J.; SU, B., 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors*. Vol. 2017, p. 1353691. Available from DOI: 10.1155/2017/1353691.

YANG, W.; DUAN, L.; CHEN, G.; XIONG, L.; LIU, Q., 2013. Plant phenomics and high-throughput phenotyping: accelerating rice functional genomics using multidisciplinary technologies. *Current Opinion in Plant Biology*. Vol. 16, no. 2, pp. 180–187. Available from DOI: 10.1016/j.pbi.2013.03.005.

YANG, X.; XIE, H.; DRURY, C.; REYNOLDS, W.; YANG, J.; ZHANG, X., 2012. Determination of organic carbon and nitrogen in particulate organic matter and particle size fractions of Brookston clay loam soil using infrared spectroscopy. *European Journal of Soil Science*. Vol. 63, no. 2, pp. 177–188. Available from DOI: 10.1111/j.1365-2389.2011.01421.x.

YU, K.-Q.; ZHAO, Y.-R.; LI, X.-L.; SHAO, Y.-N.; LIU, F.; HE, Y., 2014. Hyperspectral Imaging for Mapping of Total Nitrogen Spatial Distribution in Pepper Plant. *PLoS ONE*. Vol. 9, no. 12, e116205. Available from DOI: 10.1371/journal.pone.0116205.

ZARCO-TEJADA, P.; GONZÁLEZ-DUGO, V.; BERNI, J., 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment*. Vol. 117, pp. 322–337. Available from DOI: 10.1016/j.rse.2011.10.007.

ŻELAZNY, W. R., 2020. Application of feature selection for predicting leaf chlorophyll content in oats (*Avena sativa* L.) from hyperspectral imagery. *Agronomy Research*. Vol. 18, no. 4, pp. 2665–2676. Available from DOI: 10.15159/ar.20.174.

ŻELAZNY, W. R.; CHRPOVÁ, J.; HAMOUZ, P., 2021. Fusarium head blight detection from spectral measurements in a field phenotyping setting—A preregistered study. *Biosystems Engineering*. Vol. 211, pp. 97–113. Available from DOI: 10.1016/j.biosystemseng.2021.08.019.

ŻELAZNY, W. R.; LUKÁŠ, J., 2020. Drought Stress Detection in Juvenile Oilseed Rape Using Hyperspectral Imaging with a Focus on Spectra Variability. *Remote Sensing*. Vol. 12, no. 20, p. 3462. Available from DOI: 10.3390/rs12203462.

ŻELAZNY, W. R.; ŠIMON, T., 2022. Calibration Spiking of MIR-DRIFTS Soil Spectra for Carbon Predictions Using PLSR Extensions and Log-Ratio Transformations. *Agriculture*. Vol. 12, no. 5, p. 682. Available from DOI: 10.3390/agriculture12050682.

ZHANG, C.; KOVACS, J. M., 2012. The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture*. Vol. 13, no. 6, pp. 693–712. ISSN 1573-1618. Available from DOI: 10.1007/s11119-012-9274-5.

ZHANG, D.-Y.; CHEN, G.; YIN, X.; HU, R.-J.; GU, C.-Y.; PAN, Z.-G.; ZHOU, X.-G.; CHEN, Y., 2020. Integrating spectral and image data to detect Fusarium head blight of wheat. *Computers and Electronics in Agriculture*. Vol. 175, p. 105588. Available from DOI: 10.1016/j.compag.2020.105588.

ZHANG, D.; CHEN, G.; ZHANG, H.; JIN, N.; GU, C.; WENG, S.; WANG, Q.; CHEN, Y., 2020. Integration of spectroscopy and image for identifying fusarium damage in wheat kernels. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. Vol. 236, p. 118344. Available from DOI: 10.1016/j.saa.2020.118344.

ZHANG, D.; WANG, D.; GU, C.; JIN, N.; ZHAO, H.; CHEN, G.; LIANG, H.; LIANG, D., 2019. Using Neural Network to Identify the Severity of Wheat *Fusarium* Head Blight in the Field Environment. *Remote Sensing*. Vol. 11, no. 20, p. 2375. Available from DOI: 10.3390/rs11202375.

ZHANG, D.; WANG, Q.; LIN, F.; WENG, S.; LEI, Y.; CHEN, G.; GU, C.; ZHENG, L., 2020. New Spectral Classification Index for Rapid Identification of Fusarium Infection in Wheat Kernel. *Food Analytical Methods*. Vol. 13, pp. 2165–2175. Available from DOI: 10.1007/s12161-020-01829-w.

ZHANG, D.; WANG, Q.; LIN, F.; YIN, X.; GU, C.; QIAO, H., 2020. Development and Evaluation of a New Spectral Disease Index to Detect Wheat Fusarium Head Blight Using Hyperspectral Imaging. *Sensors*. Vol. 20, no. 8, p. 2260. Available from DOI: 10.3390/s20082260.

ZHANG, D.; WANG, X.; MA, W.; ZHAO, C., 2012. Research Vertical Distribution of Chlorophyll Content of Wheat Leaves Using Imaging Hyperspectra. *Intelligent Automation & Soft Computing*. Vol. 18, no. 8, pp. 1111–1120. Available from DOI: 10.1080/10798587.2008.10643315.

ZHANG, J.; WANG, N.; YUAN, L.; CHEN, F.; WU, K., 2017. Discrimination of winter wheat disease and insect stresses using continuous wavelet features extracted from foliar spectral measurements. *Biosystems Engineering*. Vol. 162, pp. 20–29. Available from DOI: 10.1016/j.biosystemseng.2017.07.003.

ZHANG, L.; YANG, X.; DRURY, C.; CHANTIGNY, M.; GREGORICH, E.; MILLER, J.; BITTMAN, S.; REYNOLDS, W. D.; YANG, J., 2018. Infrared spectroscopy estimation methods for water-dissolved carbon and amino sugars in diverse Canadian agricultural soils. *Canadian Journal of Soil Science*. Vol. 98, pp. 484–499. Available from DOI: 10.1139/cjss-2018-0027.

ZHANG, X.; HE, Y., 2013. Rapid estimation of seed yield using hyperspectral images of oilseed rape leaves. *Industrial Crops Products*. Vol. 42, pp. 416–420. Available from DOI: 10.1016/j.indcrop.2012.06.021.

ZHANG, X.; LU, G.; LONG, W.; ZOU, X.; LI, F.; NISHIO, T., 2014. Recent progress in drought and salt tolerance studies in *Brassica* crops. *Breeding Science*. Vol. 64, no. 1, pp. 60–73. Available from DOI: 10.1270/jsbbs.64.60.

ZIMMERMANN, M.; LEIFELD, J.; FUHRER, J., 2007. Quantifying soil organic carbon fractions by infrared-spectroscopy. *Soil Biology and Biochemistry*. Vol. 39, no. 1, pp. 224–231. Available from DOI: 10.1016/j.soilbio.2006.07.010.

ZOVKO, M.; ŽIBRAT, U.; KNAPIČ, M.; KOVAČIĆ, M. B.; ROMIĆ, D., 2019. Hyperspectral remote sensing of grapevine drought stress. *Precision Agriculture*. Vol. 20, pp. 335–347. Available from DOI: 10.1007/s11119-019-09640-2.

ZYL, C. J. van, 2018. Frequentist and Bayesian inference: A conceptual primer. *New Ideas in Psychology.* Vol. 51, pp. 44–49. Available from DOI: 10.1016/j. newideapsych.2018.06.004.

ZYPHUR, M. J.; OSWALD, F. L., 2015. Bayesian estimation and inference: A user's guide. *Journal of Management.* Vol. 41, no. 2, pp. 390–420. Available from DOI: 10.1177/0149206313501200.

Appendix A

Wiktor R. Żelazny and Jan Lukáš: Drought Stress Detection in Juvenile Oilseed Rape Using Hyperspectral Imaging with a Focus on Spectra Variability

remote sensing

MDPI

Article

Drought Stress Detection in Juvenile Oilseed Rape Using Hyperspectral Imaging with a Focus on Spectra Variability

Wiktor R. Żelazny [1],* and Jan Lukáš [2]

[1] Crop Research Institute Prague-Ruzyně, Division of Crop Management Systems, Drnovská 507/73, CZ161 06 Praha 6 Ruzyně, Czech Republic

[2] Crop Research Institute Prague-Ruzyně, Division of Crop Protection and Plant Health, Drnovská 507/73, CZ161 06 Praha 6 Ruzyně, Czech Republic; lukas@vurv.cz

* Correspondence: wzelazny@vurv.cz; Tel.: +420-233-022-491

check for updates

Abstract: Hyperspectral imaging (HSI) has been gaining recognition as a promising proximal and remote sensing technique for crop drought stress detection. A modelling approach accounting for the treatment effects on the stress indicators' standard deviations was applied to proximal images of oilseed rape—a crop subjected to various HSI studies, with the exception of drought. The aim of the present study was to determine the spectral responses of two cultivars, 'Cadeli' and 'Viking', representing distinctive water management strategies, to three types of watering regimes. Hyperspectral data cubes were acquired at the leaf level using a 2D frame camera. The influence of the experimental factors on the extent of leaf discolorations, vegetation index values, and principal component scores was investigated using Bayesian linear models. Clear treatment effects were obtained primarily for the vegetation indexes with respect to the watering regimes. The mean values of RGI, MTCI, RNDVI, and GI responded to the difference between the well-watered and water-deprived plants. The RGI index excelled among them in terms of effect strengths, which amounted to $-0.96\,[-2.21, 0.21]$ and $-0.71\,[-1.97, 0.49]$ units for each cultivar. A consistent increase in the multiple index standard deviations, especially RGI, PSRI, TCARI, and TCARI/OSAVI, was associated with worsening of the hydric regime. These increases were captured not only for the dry treatment but also for the plants subjected to regeneration after a drought episode, particularly by PSRI (a multiplicative effect of $0.33\,[0.16, 0.68]$ for 'Cadeli'). This result suggests a higher sensitivity of the vegetation index variability measures relative to the means in the context of the oilseed rape drought stress diagnosis and justifies the application of HSI to capture these effects. RGI is an index deserving additional scrutiny in future studies, as both its mean and standard deviation were affected by the watering regimes.

Keywords: imaging spectroscopy; Rikola; irradiance; *Brassica napus* L.; pot experiments; reproducibility

1. Introduction

Given the continuing increases in average temperatures [1] and projections of more frequent and severe droughts in agricultural regions [2,3], water deficiency has been among the most extensively studied crop stress factors [4]. In pot experiments, crop responses to drought can be investigated by varying the watering regime and comparing the obtained plant reactions across the treatments [5–8]. An alternative approach is to exploit the variability of water management strategies exhibited by individual genotypes [9,10].

Several dehydration avoidance mechanisms have been described in crops [11,12]. Plants can rapidly respond to a water deficit by closing their stomata, which reduces the leaf transpiration. As a

trade-off, this reduction leads to a simultaneous decrease in the photosynthesis rate, related to limited CO_2 assimilation [13,14]. Differentiation of crop cultivars with respect to their stomatal conductance regulation has been proposed. Plants that manage their water resources in a conservative way and maintain a steady CO_2 fixation rate, affected by moisture availability to a limited extent, have been termed as water-savers. Water-spenders, on the other hand, maximise their CO_2 assimilation, depleting the available water resources at the onset of a drought due to a delayed closure of the stomata [15,16]. Cultivars with a high baseline stomatal conductance tend to not exhibit a mid-day depression in photosynthetic rates. They are capable of sustaining a high photosynthesis rate and can avoid heat stress due to the cooling action of the transpiration, provided that water is available [17].

Stressed and healthy vegetation differ with respect to their spectral reflectances, with the effects of stress detectable before they become apparent to the naked eye [18–20]. The visible spectral region is affected by stress-induced changes in pigment concentrations and activities. These changes include anthocyanin and (relative) carotenoid accumulation aimed at protecting the photosynthetic apparatus and pigment breakdown, which accompanies chloroplast deterioration caused by oxidative stress [13,21–23]. As leaf mesophyll cells lose their turgor and shrink, there can be a temporary increase in the near-infrared reflectance, eventually followed by a decrease below the normal level [19]. The red-edge shifts towards shorter wavelengths and becomes less steep [20,24,25].

Among the spectral methods, imaging spectrometry (hyperspectral imaging) has been gaining recognition as a promising proximal and remote sensing technique for crop status assessment [26–28]. Its important advantage over the more traditional point spectrometry is the availability of precise spatial information [26], which can address the mixed spectra problem in close-range applications. This advantage is accomplished by using spectral segmentation methods [29], which enable the separation of the object and background pixels [30], or the identification of pixels affected by unfavourable illumination effects [6]. Furthermore, the presence and distribution of geometric features can be analysed in the image [29,31].

Studies devoted to drought effects on crop hyperspectra have been primarily focused on the species that dominate the global commodity market. Those species include maize [6,30] and other staple cereals [7,32,33]. A relatively large amount of attention has also been given to fruit crops [18,34,35]. On the other hand, numerous other species have so far been largely neglected by the studies, including those of regional importance.

Due to its nutritional [22,36,37] and technical [37,38] value, oilseed rape (*Brassica napus* L.; hereafter, OSR) is an important crop in many parts of the world. It is widespread in North America [39,40], China [40], Europe [39], and India [41]. OSR is susceptible to drought [12,22] and, along with other brassicas, the future cultivation of this species is endangered by dry spells [39,42].

The reproductive phase of OSR has been associated with an especially high sensitivity [22,43], but it can also be permanently affected by water deprivation earlier in its development. This possibility justifies extending studies to juvenile plants and to crop recovery after conclusion of the drought period, which is an underexplored research area [5]. Müller et al. [44] compared the physiological status of OSR plants that had been water-deprived at the shooting developmental stage and then rewatered with specimens receiving irrigation for the entire duration of the experiment. The treated plants exhibited reduced productivity and their physiological profiles were affected. The physiological changes in even younger plants were studied by Kosová et al. [45] and Urban et al. [15].

In addition to the physiological parameters, a trace of a drought episode can be detectable in a spectral signature of the affected crop. Such a possibility was demonstrated by Linke et al. [8] for wheat and by Sun et al. [5] for maize. The authors tested the changes of several vegetation indexes in plants exposed to repeated drought and recovery cycles. They observed a full recovery after the first cycle, but the second recovery was incomplete. As a possible cause, the authors suspected progressing cell deterioration due to oxygen radicals, which could not be neutralised in the absence of carotenoids due to their removal in the course of the preceding stress episode.

OSR has been the subject of various hyperspectral imaging studies. Based on field experiments, Piekarczyk et al. [46] and Zhang and He [47] attempted to predict its yield using vegetation indexes and partial least squares regression, respectively. Kumar et al. [41] cite several publications devoted to OSR pests and diseases. Xia et al. [37] analysed the imagery of water-logged plants. The effects of herbicide exposure were studied by Kong et al. [48]. In contrast to these stress factors, the possibilities of capturing the OSR response to drought using a hyperspectral camera remain unaddressed.

As highlighted by Kruschke and Liddell [49], "stressors [...] can increase the variance of a group because not everyone responds the same way to the stressor". In the context of close-range crop hyperspectral imaging, the "group" can refer to plant foliage or leaf tissue composed of individual leaves and cells, respectively, each responding to the change in the environment in a distinct way. Especially characteristic for stress-induced leaf senescence is the source–sink differentiation between the older and younger leaves [23]. The potential of imaging spectrometry to provide an insight into the spatial variation of stress symptoms across crop foliage was demonstrated for drought [7,32,33,50], nitrogen deficiency [51], pest infestation [50], and herbicide exposure [48]. OSR is characterised by relatively large leaves, even in early developmental phases. Hyperspectral imaging that captures leaf-level spectral variation may, therefore, prove to be a suitable approach for water deficiency detection in this crop [32].

Studies on crop responses to stress conditions frequently employ traditional experimental designs, such as a randomised block design, coupled with linear modelling for statistical inference [42,43,52]. The frequentist approach prevails in the fitting and evaluation of these models. Various authors noted the shortcomings of the frequentist statistics, and have advocated Bayesian methods as an alternative [49,53,54]. Historically, first the lack and then the high computational demands of suitable numerical methods posed obstacles towards a wider adoption of the Bayesian paradigm [55]. These hindrances have been largely removed by an increase in computer speeds [55,56], followed by improved accessibility of parallel computing [57], and the availability of software with capabilities suited to the needs of the scientific community [55,58–60].

One major appeal of Bayesian statistics is the ease with which interval estimates of model parameters can be quantified, even for complex models. Notably, it is possible to obtain estimates with respect to not only the mean values but also standard deviations, shape factors, or hurdle values—again, also for complex models [49]. In the context of stress detection with imaging spectroscopy, this capability can be readily exploited to assess the influence of the stressor on the spectral variation across the foliage of the affected plant.

The aim of the present study is to determine the spectral response of juvenile OSR representing two water management strategies to three types of watering regimes. The study is performed at the leaf level by employing a high-resolution hyperspectral camera. The influence of the OSR cultivars and watering regimes on the extent of leaf discolourations, vegetation indexes, and principal component scores are investigated. Bayesian statistics are used to obtain the interval estimates of the treatment effects with respect to the mean value and standard deviation differences.

2. Material and Methods

2.1. Plant Material and Experimental Factors

The experiment was based on winter OSR plants of the 'Cadeli' and 'Viking' cultivars. The two genotypes differ in terms of their drought-coping strategies, with 'Cadeli' representing the "water-saver strategy" and 'Viking' exhibiting the "water-spender strategy", as revealed by their physiological [15,45] and proteomic [15] profiles. This difference is related to the origin of the cultivars, which is France for 'Cadeli' and Germany for 'Viking' [15].

The study was conducted on the premises of the Crop Research Institute in Prague-Ruzyně (Czech Republic). The seeds of both cultivars were obtained from OSEVA PRO s.r.o. (Opava, Czech Republic). Each seed was started on 11th May 2018 by placing it in a thermostat (Biological thermostat BT-120, Laboratorní přístroje Praha, Prague, Czechoslovakia) for two days set to 20 °C.

The obtained seedlings were transplanted three days later to 14-cm diameter pots filled with 1.01 kg of potting mixture produced at the site. The seeds were topped with an additional 0.25 kg of the mixture. Each pot contained five seedlings of either of the cultivars. The potted plants were grown in a growth chamber (T-64, Tyler, Budapest, Hungary) in 18 °C to 20 °C, under a 16-h photo period, exposed to 400 µmol m^{-2} s^{-1} irradiance up to the second leaf (BBCH 12) developmental stage. At that point, the watering regime experimental factor was introduced.

The watering of the pots followed one of the three treatments: The control pots were watered daily to 70% of the substrate water capacity (SWC). For the pots in the dry treatment, the watering was reduced to 45% of the SWC, starting 14 days after the sowing, and 10 more days later the watering was stopped completely. The pots in the rewatered treatment were treated according to the same plan as the water-deprived plants, but after 5 days of suspended watering they were watered to 100% of the SWC to induct regeneration. The pots were grouped in the growth chamber according to the watering regime, which resulted in a lack of true replication of this factor.

A total of $n = 26$ pots were used in the study. Table 1 depicts the pot numbers according to the two experimental factors. The uneven numbers across the treatment combinations stem from the fact that the material was used in another experiment, which involved destructive sampling.

Table 1. Oilseed rape pot counts used in the experiment according to the experimental treatments. The parenthesised values refer to the counts after excluding a low-quality image.

Watering Regime	Cultivar	
	Cadeli	Viking
dry	6	6
rewatered	3 (2)	3 (2)
watered	5	3

2.2. Image Acquisition and Pre-processing

The imaging took place on 27 June 2018. By that time, the plants had attained the phase of 3 to 4 leaves (BBCH 13, 14), and pigmentation changes due to water deprivation were apparent (Figure 1a). The images were collected between 14:45 and 15:45, outdoors, in natural light conditions. The illumination was variable, as illustrated by the radiant exposure measurements from a meteorological station located at the site (Figure 2), and there were periods of no direct sunlight. A photo tent was used to obtain diffuse illumination and create a wind barrier.

The imager was a 2D frame hyperspectral camera (Rikola, Senop, Oulu, Finland), mounted on a tripod at the tent entrance (Figure 1b). An irradiance sensor was placed inside the tent to account for the variation in the illumination conditions. Its readings, expressed in relative units, varied between 82 and 181 (Figure 2), reflecting the unstable light conditions during the campaign. A dark reference was obtained prior to the acquisition of the OSR images with the aid of a 50-mm black masking tape (T743-2.0, Thorlabs Inc., Newton, NJ, USA). Four pots—two per cultivar and placed at the tent bottom in an alternated manner against a background of black non-woven textile—were captured in each image. Since the number of pots in the rewatered treatment was not divisible by four, some of the pots were captured for a second time. Those extra pot images were not included in the analysed dataset. First, the dry plants were imaged, followed by the watered plants, and then followed by the rewatered plants. This ordering reflected the pot grouping in the growth chamber. The images of the plants were interleaved with images of Spectralon tiles with 2, 9, 23, 44, and 75% reflectance factors. The internal camera temperature was stable, in the 32.56 °C to 33.81 °C range.
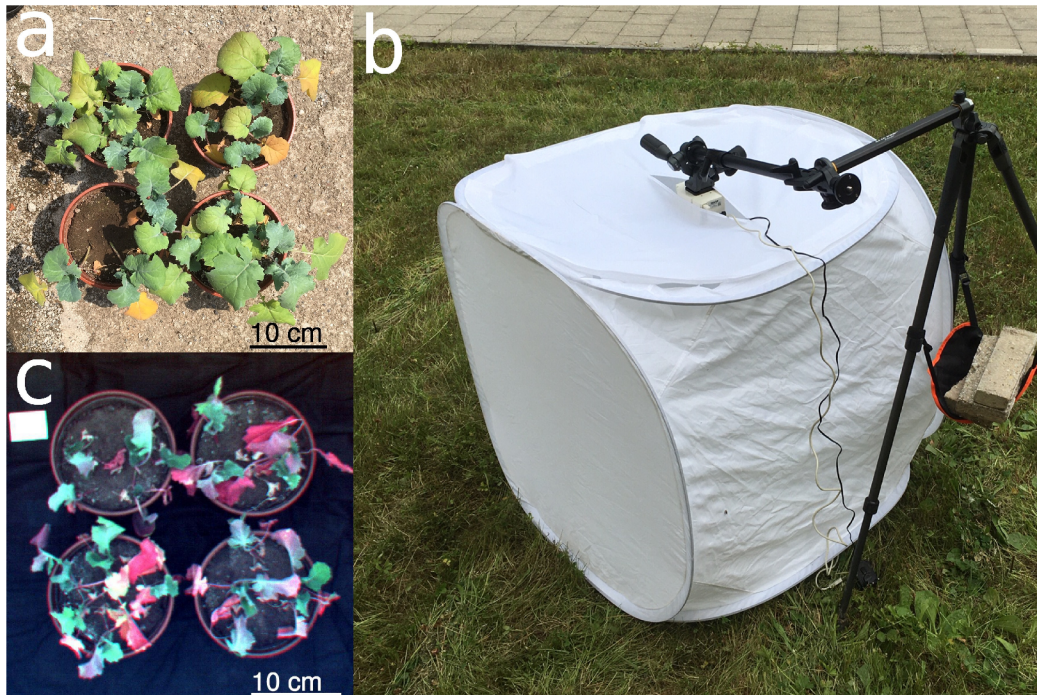
Figure 1. Material used in the experiment. (**a**) Oilseed rape in pots 16 days before the spectral data collection. (**b**) Imaging setup with the photo tent and the hyperspectral camera on the tripod. (**c**) False-colour composite rendering of one of the hyperspectral data cubes employed in the pixel labelling.
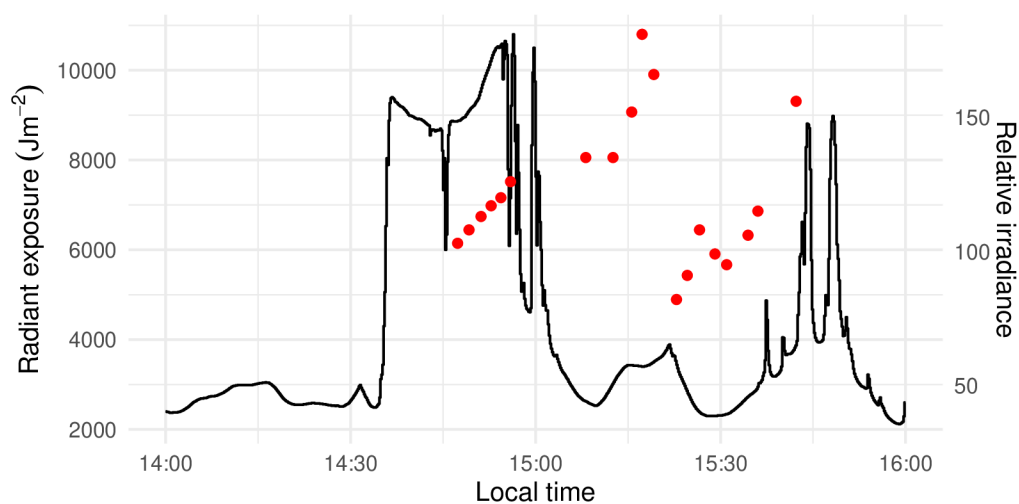

Figure 2. Radiometric measurements of ambient illumination during the acquisition of the hyperspectral imagery. The black line represents radiant exposure values integrated over 10-s intervals. The red dots represent readings from a relative irradiance sensor associated with the individual images.

The hyperspectral data cubes comprised 41 evenly spaced bands ranging from 503 to 903 nm. The spatial resolution was 1010×1010 pixels and the radiometric resolution was 12 bits. The integration time was set to 30 ms. The pot rims were approximately 0.70 m away from the camera lens, resulting in a GSD of approximately 0.43 mm per image pixel.

Figure 3 depicts the imagery pre-processing and processing workflow. After the conversion from the digital numbers, Spectralon radiance values were sampled on a 20 × 20 pixel grid. The grid sampling was intended to reduce the effect of spatial autocorrelation and the computation time. For every band, a mixed-effect empirical line model [61] was fitted with a reflectance logit as the dependent variable, radiance as the independent variable, and pixel as the grouping variable (`fit ELM` in Figure 3). The logit transformation accounted for the reflectance values being constrained between 0 and 100% [62], and the grouping variable was introduced for the possibility of an uneven illumination of the scene during the acquisition process. A transformation of the radiance values to the reflectance values was performed by applying the obtained models to the individual spectral bands (`pred ELM`).

In the next pre-processing step, the reflectance images were subjected to band registration (`reg`) to remove the effect of camera sensor misalignment [63]. The ORB algorithm was used for feature detection and description and was coupled with brute-force descriptor matching [64]. The band registration failed for one of the images of the rewatered plants, probably because of leaf movements from the wind. That image was excluded from the subsequent analyses (Table 1).

In each of the plant images, the coordinates of three points along each pot circumference were identified by hand. From these, the pot centre coordinates and radii were derived. The radii lengths were then reduced by a factor of 0.95 to exclude the pot rims from the regions of interest, which were delimited as inscribed squares (`crop`).

2.3. Pixel Classification and Evaluation of Class Size Proportions

A random sample of 200 pixels was drawn from across all regions of interest to train and validate a classification model aimed at distinguishing between healthy (fresh) leaf zones and those exhibiting discolouration, which was attributed to drought. Due to an uneven number of pixels across the pot images, stratified sampling was employed. First, a pot was sampled, followed by a pixel within. The sampled pixels were subsequently hand-classified as either background, fresh-leaf, dry-leaf, or edge pixel, based on pseudo-RGB (R: 647 nm, G: 563 nm, B: 503 nm) rendering of the pot images (Figure 1c).

The pixels at the leaf edges or zone boundaries were treated as missing data and dropped. The reflectance spectra of the remaining pixels ($n = 181$) were randomly partitioned into the training and test dataset at a 3:1 proportion. The partitioning was stratified with respect to the pixel class. Using the training dataset, a Support Vector Machine (SVM) classification model with the radial basis function kernel [65] was fitted to the pixel hyperspectra (`fit SVM`). The `cost` hyperparameter of the model was tuned to maximise the classification accuracy using 10-fold cross-validation and the Bayesian model-based optimisation search algorithm [66]. The performance of the obtained model was then assessed using the test dataset. Finally, the model was applied to classify every pixel in the pot images (`pred SVM`).

The dry-leaf and fresh-leaf classes were merged to create plant masks [67], which were subsequently subjected to 3-pixel erosion [63] to remove leaf pedicels and artefacts resulting from imperfect band registration. The eroded masks were then applied to the pot images (`mask`).

Dry-leaf pixels were counted in each masked pot image, and the effects of the experimental treatments on the dry-leaf pixel proportion were assessed using a Bayesian linear mixed-effect model (`fit LM1`). The model assumed a zero-inflated binomial data generating distribution of the response variable (with a logit link) and accounted for the grouping of the leaf pixels within the pots and of the pots within the individual images. In addition to reflecting the dataset structure, the inclusion of the grouping variables was intended to address the problem of variable illumination conditions during the hyperspectral data acquisition campaign. Conservative, yet meaningful priors [54,68] were assumed.
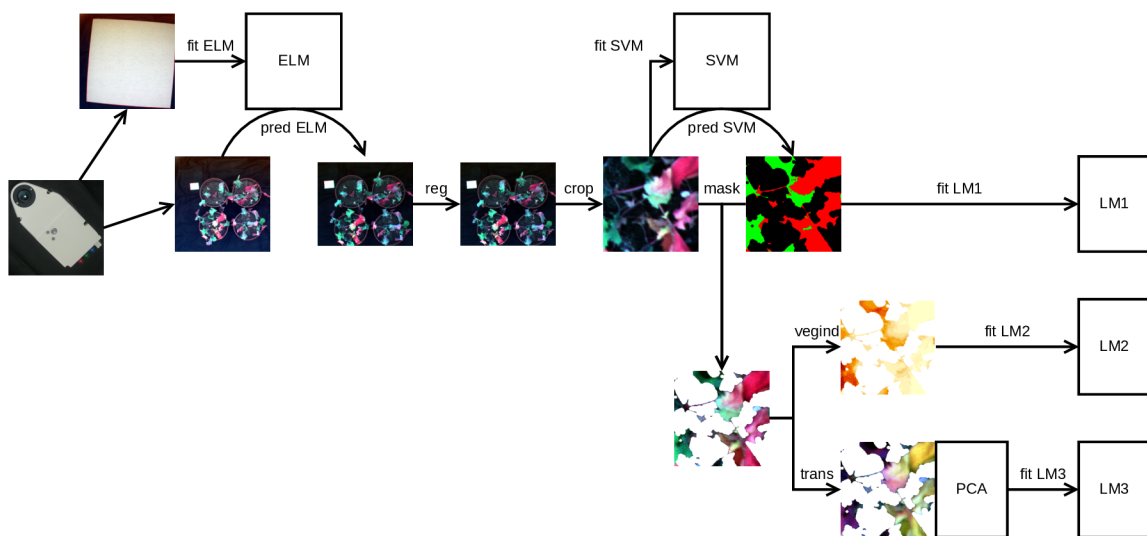
Figure 3. Pre-processing and processing workflow of the hyperspectral images. `fit ELM` = fitting of an empirical line model, `pred ELM` = application of the empirical line model to the spectra, `reg` = band registration, `crop` = cropping of the images to the regions of interest, `fit SVM` = fitting of a Support Vector Machine model, `pred SVM` = pixel classification using the Support Vector Machine model, `mask` = background masking, `fit LM1` = linear modelling of dry-leaf pixel proportions, `vegind` = deriving of vegetation indexes, `fit LM2` = linear modelling of the vegetation index values, `trans` = spectra pre-processing, `fit LM3` = linear modelling of the principal component scores of the pre-processed spectra.

2.4. Vegetation Index and Full-Spectrum Analyses

Twenty vegetation-index values (Table 2) were calculated for each pixel of the masked pot images (`vegind`). In cases where the wavelengths present in an index definition did not match the available image wavelengths, the closest wavelength was used, instead.

Table 2. Vegetation indexes used in the study, and their reported sensitivities to biochemical, physiological, and structural plant properties.

Vegetation Index	Formula	Sensitivity	Reference
SR	$\frac{r_{900}}{r_{680}}$	Chl, fIPAR, LAI	[69]
GI	$\frac{r_{554}}{r_{677}}$	Chl	[70]
RGI	$\frac{r_{690}}{r_{550}}$	Chl	[70]
DVI	$r_{900} - r_{680}$	LWC	[71]
NDVI	$\frac{r_{900} - r_{680}}{r_{900} + r_{680}}$	Chl, fIPAR, LAI	[69]
RDVI	$\sqrt{\mathrm{NDVI} \times \mathrm{DVI}}$	fAPAR	[72]
PSRI	$\frac{r_{678} - r_{500}}{r_{750}}$	Chl, Car	[73]
PSSR$_a$	$\frac{r_{800}}{r_{680}}$	Chl, Car	[74]
PSND$_a$	$\frac{r_{800} - r_{680}}{r_{800} + r_{680}}$	Chl, Car	[74]
RNDVI	$\frac{r_{750} - r_{705}}{r_{750} + r_{705}}$	Chl	[75]
PRI$_{570}$	$\frac{r_{570} - r_{531}}{r_{570} + r_{531}}$	ΔF/Fm′	[76]
PRI$_{512}$	$\frac{r_{512} - r_{531}}{r_{512} + r_{531}}$	G$_s$, Ψ, EPS	[77]
PRI$_{norm}$	$\frac{\mathrm{PRI}_{570}}{\mathrm{RDVI} \times \frac{r_{700}}{r_{670}}}$	G$_s$, Ψ	[78]
MTCI	$\frac{r_{753.75} - r_{708.75}}{r_{708.75} - r_{681.25}}$	Chl	[79]
MCARI	$\left[(r_{700} - r_{670}) - 0.2(r_{700} - r_{550})\right] \times (r_{700} - r_{670})$	Chl	[80]
TCARI	$3\left[(r_{700} - r_{670}) - 0.2(r_{700} - r_{550})\frac{r_{700}}{r_{670}}\right]$	Chl	[81]
OSAVI	$\frac{r_{800} - r_{670}}{r_{800} + r_{670}}$	LAI	[82]
TCARI/OSAVI	$\frac{\mathrm{TCARI}}{\mathrm{OSAVI}}$	Chl	[81]
CI$_{green}$	$\frac{r_{750}}{r_{550}} - 1$	Chl	[83]
CI$_{re}$	$\frac{r_{750}}{r_{710}} - 1$	Chl	[83]

Car = carotenoids, Chl = chlorophyll, EPS = violaxanthin:antheraxanthin:zeaxanthin balance, ΔF/Fm′ = fluorescence-based PSII light use efficiency, fAPAR = fractional absorbed PAR, fIPAR = fractional intercepted PAR, G$_s$ = stomatal conductance, LAI = leaf area index, LWC = leaf water content, Ψ = leaf water potential.

From each vegetation index pot image, 36 pixels were sampled on a regular grid to reduce the effect of spatial correlation while retaining information on the within-pot index value variation. Pixels located in the masked-out areas were discarded. To assess the effect of the experimental treatments on the index values, an ensemble of univariate Bayesian linear mixed-effect models was fitted (`fitLM2`). The individual model formulations took the sample–pot–image grouping hierarchy of the observations into account and relaxed the assumption of index value homoscedasticity across the treatment combinations. Because of the variety of the indexes, the modelling assumed uninformative priors [55,84].

In addition to the vegetation-index approach, an analysis based on the full-spectrum information from raw and pre-processed spectra was attempted. The pre-processing scenarios (`trans`) comprised the Savitzky–Golay filter (SGF), multiplicative scatter correction (MSC), finite differences derivation, and second derivation [85]. In the next step, the spectra were subjected to dimensionality reduction

using PCA to remove redundant radiometric information. The influence of the experimental treatments on the first four PCA loading values was then assessed using multivariate linear modelling (`fit LM3`), readily available in the Bayesian paradigm [55], to take the correlations between the loadings into account. The predictor part of the model was formulated in the same way as for the vegetation indexes.

2.5. Statistical Inference and Model Diagnostics

Accuracy of the SVM classification was determined using a confusion matrix. For the linear models, the posterior distributions of parameters were derived and visualised to assess the directions, magnitudes, and uncertainties of treatment effect estimates. Numerical summaries: posterior mode and a 95% credibility interval [49,54] were also computed. The estimated differences among the vegetation index means were additionally converted to Cohen's *d* relative effect sizes, with the watered treatment index standard deviations pooled across the cultivars as the standardiser [86]. The fits of the linear models were assessed using the *R-hat* statistics [59], by inspecting posterior trace plots [55], and performing predictive posterior checks [54].

2.6. Reproducing the Study

Pre-processed hyperspectral data cubes are available from a Zenodo repository along with the scripts that were employed for their analysis (doi:10.5281/zenodo.3975431). A GNU Guix [87] manifest file and definitions of extra software packages are also included to recreate the computational environment. A Makefile [88] describes and facilitates the execution of individual steps of data processing.

A major part of the analysis was programmed in the R language, and run in the 3.6.1 version of the interpreter [89]. The `e1071` package (version 1.7.2) [90] was employed to fit the SVM model, and `mlr` (2.15.0) [91] was used in combination with `mlrMBO` (1.1.2) [66] for its tuning. The Bayesian linear models were fitted with the aid of the `brms` (2.10.0) [58] interface to Stan (2.19.1) [59]. Tools available in SAGA GIS (6.3.0) [92], accessed from the `RSAGA` package (1.3.0) [93], enabled image masking and erosion. Band registration was performed using Python bindings to the OpenCV library (3.4.3) [94].

3. Results

3.1. Image Segmentation and Dry Pixel Occurrence

Accurate classification of the reflectance spectra was obtained with SVM, with all but 2 of the 47 pixels in the test set correctly assigned (Table 3). The spectra of ten randomly sampled pixels in each pot data cube are shown in Figure 4. As expected, the spectra of the pixels identified as dry exhibit a decreased red-edge slope and absent chlorophyll absorption features [24,25]. Their spectral variability for wavelengths below 700 nm appears higher relative to the fresh pixels. The background spectra form a slightly curved pattern, which is typically encountered for soil. Some pixels in this class are characterised by an increase in the near infrared reflectance, which can be attributed to organic debris and sub-pixel effects (spectral mixing).

Table 3. Validation results of the Support Vector Machine (SVM) pixel classification model.

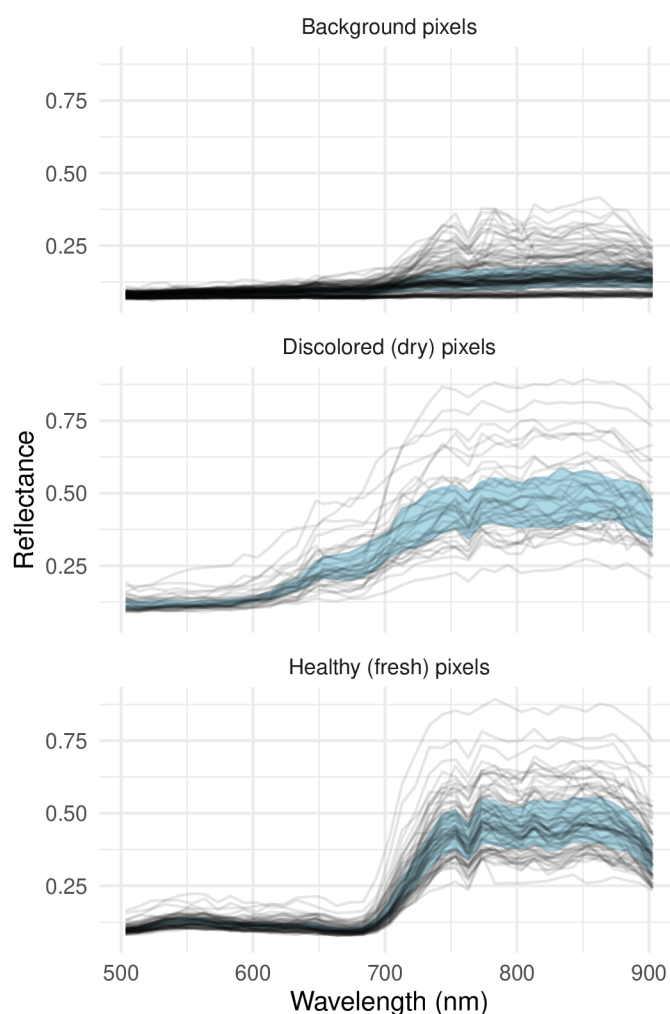| Observed Classes	Predicted Classes		
	b	d	f
background	26	0	1
dry	0	8	1
fresh	0	0	11

Figure 4. Representative hyperspectra of the pixel classes obtained from SVM segmentation of pot images. The blue areas delimit the interquartile reflectance ranges.

Figure 5 depicts the relationship between the experimental factors and the proportion of pixels identified by the SVM model as dry in a hyperspectral image. The narrowest posterior distribution was obtained for the cultivar contrast under the dry regime, meaning that the cultivar effect was estimated with the highest certainty in this analysis [49]. However, since the distribution is centred close to the value of 1, it fails to provide information on the sign of the difference. Wide posterior distributions were obtained for the two remaining comparisons in this group. The multiplicative effect size along with the 95% credibility interval is $\text{ES}_{\text{Viking/Cadeli}} = 1.06\,[0.23, 5.12]$. The contrasts involving the watering regime suggest the dry pixel occurrence having been affected by a restricted water supply, albeit with a high uncertainty. As expected, all but two comparisons indicate a lower dry leaf surface area with improved water availability, especially for 'Viking' ($\text{ES}_{\text{watered/dry}} = 0.009\,[0.00005, 1.6]$).
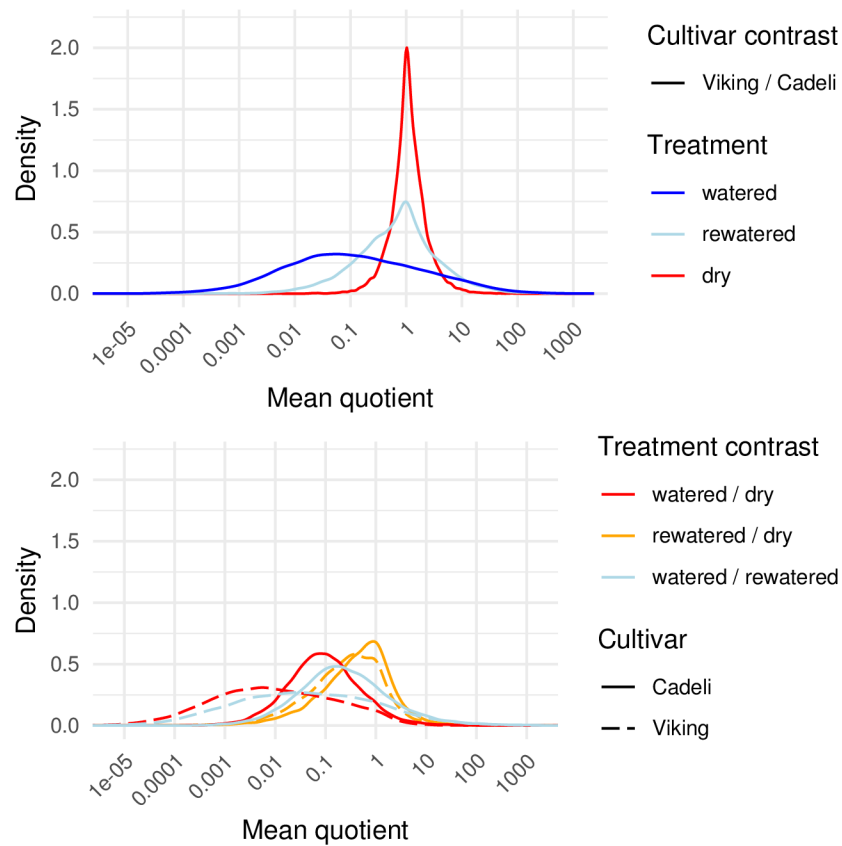
Figure 5. Posterior distributions of the mean dry pixel frequency differences between the experimental factors. Each curve represents one contrast. The differences are assumed to be multiplicative. The top subplot depicts comparisons between the cultivars and the bottom plot comparisons between watering regimes. The effect sizes are on a logarithmic scale centred at the value of 1 (lack of effect).

3.2. Vegetation Indexes

The influence of the experimental factors on the vegetation index values is shown in Figure 6. Because of the different numeric scales associated with individual formulations, the *x*-axis ranges pertaining to the index means are proportional to their standard deviations in the watered treatment, and the *y*-axis ranges are inversely proportional. In this way, not only can treatment effect directions and the strength of evidence be assessed for single indexes, but it is also possible to quantify the relative effect sizes [86] and to compare them across the formulations. Treatment effects with respect to index standard deviations were measured on a multiplicative scale. For this reason, fixed axis ranges were employed for the remaining subplots.
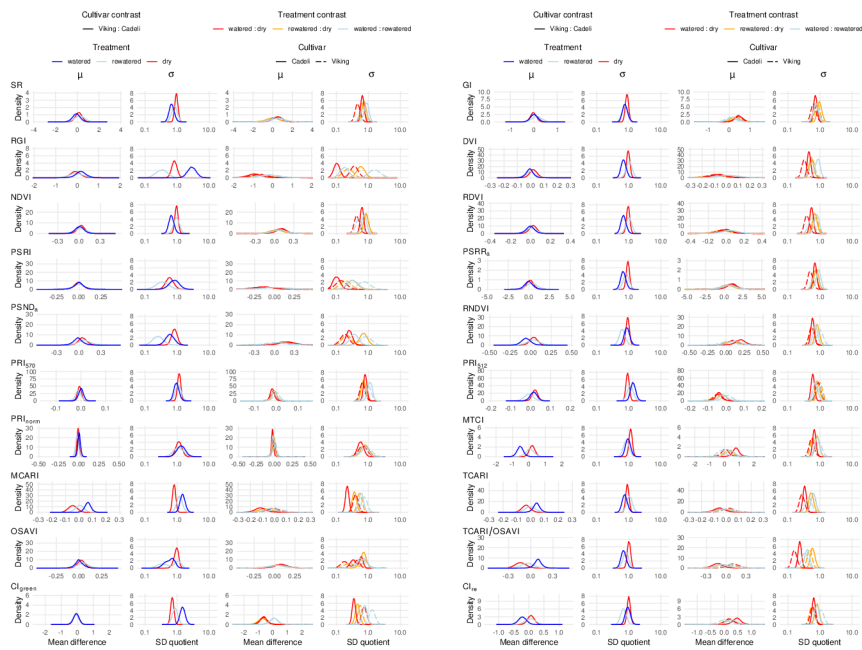
Figure 6. Posterior distributions of the mean (μ) and standard deviation (σ) differences between experimental factors for each vegetation index. The mean and standard deviation differences are assumed to be additive and multiplicative, respectively. The left-hand and right-hand side subplot columns depict comparisons between the cultivars and watering regimes, respectively. The latter are on a semi-logarithmic scale.

Regarding the cultivar effect (odd column pairs), 'Viking' and 'Cadeli' maintained under the watered treatment clearly differed with respect to the MCARI and MTCI mean index values (left-hand column in each odd column pair), as indicated by the value of zero being in the tail of the posterior density distribution. For MCARI, the relevant curve extends over the positive values of the estimated difference, which indicates that 'Viking' had, on average, higher values of this index. The raw effect size is $ES_{\mathrm{Viking-Cadeli}} = 0.06\,[0.01, 0.11]$, and Cohen's d is $d_{\mathrm{Viking-Cadeli}} = 1.16\,[0.22, 2.08]$. To a limited extent, the cultivars in the control treatment differed in terms of TCARI ($ES_{\mathrm{Viking-Cadeli}} = 0.04\,[-0.02, 0.10]$, $d_{\mathrm{Viking-Cadeli}} = 0.84\,[-0.33, 1.99]$) and TCARI/OSAVI ($ES_{\mathrm{Viking-Cadeli}} = 0.10\,[-0.02, 0.21]$, $d_{\mathrm{Viking-Cadeli}} = 1.20\,[-0.24, 2.67]$) PRI_{norm} appears to be insensitive to the cultivar differences, as indicated by the compressed mass of the posterior density centred around the value of zero ($ES_{\mathrm{Viking-Cadeli}} = 0.00\,[-0.03, 0.03]$). On the other hand, the Cohen's d credibility interval is wide ($d_{\mathrm{Viking-Cadeli}} = 0.02\,[-0.99, 1.04]$).

In addition to the vegetation index mean values, their standard deviations differed across the two cultivars (right-hand columns). Discernible differences occurred in a larger number of indexes, primarily for the control treatment. The density distributions of SR, DVI, NDVI, RDVI, $PSSR_a$, and TCARI/OSAVI extended below a ratio of one, indicating lower index value variations in watered 'Viking' than in 'Cadeli'. An opposite effect occurred for RGI, MCARI, and CI_{green}. Similar differentiation is not so apparent for the remaining treatments.

The influence of the watering regimes (even column pairs in Figure 6) on the leaf spectra was captured by the mean values of several vegetation indexes. Unsurprisingly, particularly large differences were obtained for the watered:dry contrast. The RGI index exhibited a high sensitivity in the 'Cadeli' cultivar, with its values lower in the control plants ($ES_{\mathrm{watered-dry}} = -0.96\,[-2.21, 0.21]$, $d_{\mathrm{watered-dry}} = -6.94\,[-18.01, 1.48]$). Moreover, the water availability had a positive influence on the MTCI, RNDVI, and GI indexes in the 'Cadeli' cultivar, with the effect not as strong as for RGI, but more precisely estimated, as indicated by the concentrated mass of the posterior density curve. Similarly to the cultivar effect, the PRI_{norm} mean appears to have been insensitive to the leaf spectra differences across the individual watering regimes (e.g., $ES_{\mathrm{watered-dry}} = -0.03\,[-0.06, 0.01]$, $d_{\mathrm{watered-dry}} = -0.89\,[-2.10, 0.25]$ for 'Cadeli'). The PRI index that appears to respond to the watering treatments is PRI_{512}, but this pattern is uncertain (e.g., $ES_{\mathrm{watered-dry}} = -0.04\,[-0.09, 0.01]$, $d_{\mathrm{watered-dry}} = -1.06\,[-2.42, 0.22]$ for 'Cadeli').

The variation in RGI and PSRI vegetation indexes exhibited sensitivity to the difference between the dry and control leaf spectra in 'Cadeli'. Not only is the observed treatment effect strong, but its estimate is fairly precise ($ES_{\mathrm{watered/dry},\sigma} = 0.10\,[0.07, 0.16]$ for RGI and $0.10\,[0.06, 0.17]$ for PSRI; note that the effects are multiplicative). The same indexes revealed an effect of drought on the 'Viking' spectra, albeit to a lesser degree ($ES_{\mathrm{watered/dry},\sigma} = 0.36\,[0.21, 0.64]$ and $0.14\,[0.07, 0.29]$). The variations in the majority of the remaining indexes were affected by the discussed treatment contrast for at least one of the cultivars. Several indexes revealed the difference between the rewatered and dry treatment, particularly PSRI ($ES_{\mathrm{rewatered/dry},\sigma} = 0.33\,[0.16, 0.68]$ for 'Cadeli'). Even more interestingly, the variations in TCARI and TCARI/OSAVI responded to the watered:rewatered contrast in both cultivars, with the latter index associated with a stronger effect ($ES_{\mathrm{watered/rewatered},\sigma} = 0.40\,[0.28, 0.55]$ for 'Cadeli'). What is striking it that all of the affected indexes exhibited the same direction of the water regime effect, namely, a variation decrease with an improving water availability (posterior distributions extending over values below one).

3.3. Full Spectrum Information

The distribution of the leaf pixel spectra in the principal component space did not reveal any differences between the investigated cultivars Figure 7. Regarding the watering regime, the observations representing the rewatered treatment occur in clusters. For the raw spectra, they form a line corresponding to positive PC1 or negative PC2 coordinates, and the values in-between. According to the loadings plot, both of these directions can be associated with a decreased NIR

reflectance. A similar pattern, with the PC2 axis reversed, was obtained for SGF. As a matter of fact, this pre-processing altered the spectra to a minimal degree. After MSC pre-processing, the rewatered pixel spectra become associated with high PC1 values, indicative of increased green and decreased red and red-edge reflectance, suggesting a red-edge shift towards longer wavelengths. The derivated spectra of the rewatered regime are associated with positive PC2 values, indicating a more descending slope to the left of the red absorption feature and a more ascending slope towards the longer wavelengths; thus more pronounced red light absorption. The double derivation did not result in any clustering; however, the pixels representing the dry watering regime appear to extend over a larger area of the principal component space, suggesting a higher spectral variation. An interesting pattern, though unrelated to any of the experimental treatments, can be discerned in the MSC PCA plot, in which the spectra are separated into two large clusters.

By using the obtained PCA coordinates of individual pixels as input data for linear modelling, the information on grouping of the observations could be incorporated into the analysis. With this additional step, patterns suggested by the PCA plots turned out to be largely spurious, but some new ones emerged (Figure 8). Regardless of the spectra pre-processing, no separation of the cultivars was obtained with respect to the means of the first principal component scores (first subplot column). However, the comparison of the PC1 score standard deviations (second column) revealed less varied values for the watered 'Viking' plants relative to the 'Cadeli' cultivar for MSC ($\text{ES}_{\text{Viking/Cadeli},\sigma} = 0.53\,[0.32, 0.96]$), indicating a higher variation of green, red, and red-edge reflectances in the latter (Figure 7).

Regarding the treatment contrasts, the raw and SGF-filtered spectra exhibit somewhat lower PC1 mean values of the leaf pixels in the control watering regime compared to the regeneration treatment ($\text{ES}_{\text{watered}-\text{rewatered}} = -0.33\,[-0.81, 0.15]$ and $-0.33\,[-0.84, 0.12]$, respectively for 'Cadeli'; third column in Figure 8). This outcome is in agreement with the clustering of the rewatered pixels in the right-hand part of the respective PCA plots (Figure 7), but the evidence is too weak to draw any conclusions.

A consistent pattern of treatment effect posterior distributions is apparent for the remaining pre-processing approaches. The variability of the PC1 scores was found to be higher in the watered regime than in both dry ('Cadeli') and rewatered plants (both cultivars). When the latter two treatments are compared, the dry spectra appear to be more variable. The treatments can, thus, be ordered as watered > dry > rewatered. The patterns are especially pronounced in the case of derivative spectra and spectra subjected to MSC.
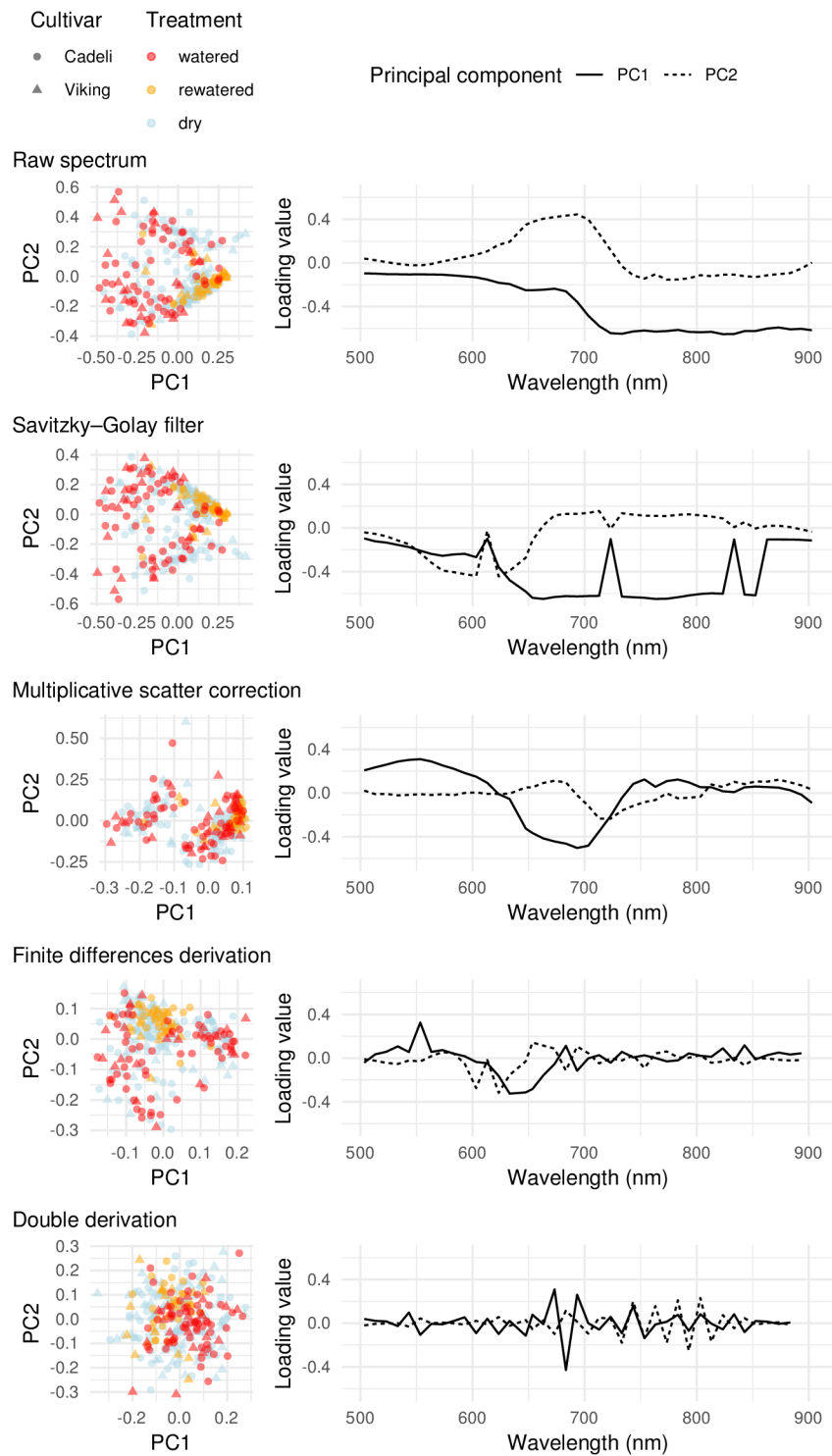
Figure 7. Principal component scores (left-hand column) and loadings (right-hand column) of the pixel reflectance values according to image pretreatment. Colour and shape differentiate observations with respect to the experimental treatments.
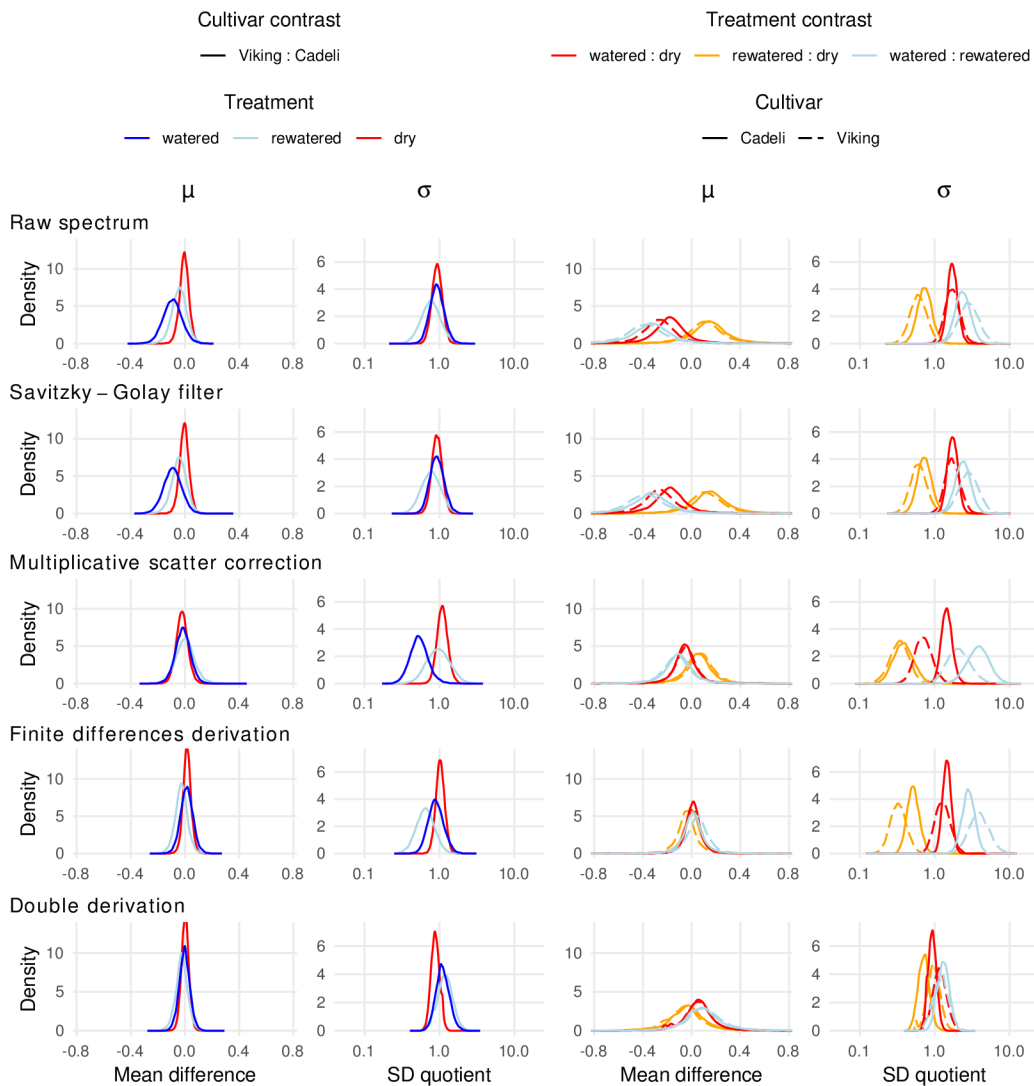
Figure 8. Posterior distributions of the mean (μ) and standard deviation (σ) differences between experimental factors for first principal component loading values according to spectra pre-processing. The mean and standard deviation differences are assumed to be additive and multiplicative, respectively. The left-hand and right-hand side subplot columns depict comparisons between the cultivars and watering regimes, respectively. The latter are on a semi-logarithmic scale.

4. Discussion

4.1. Image Quality and Patterns Related to Segmentation

The inconsistency in the irradiance sensor readings with regard to the radiant exposure measurements suggest that artefacts were introduced during the conversion of digital numbers to radiance values. This problem most likely stems from the directional sensitivity of the sensor, which was placed on the flexible photo tent construction. Directional sensitivity of the device delivered with a Rikola camera has been reported also by other authors [95]. Other potential nuisance factors include the distance from the meteorological station to the spot where the imagery was captured and the presence of a building and the camera operators in the proximity of the photo tent. Although the formulations of the linear models, employed at later stages of the image analysis, accounted for

radiometric differences between the individual data cubes, it is still preferable to acquire data of maximum quality in the first place. Therefore, for similar studies, rigid installation of a Rikola irradiance sensor is recommended.

SVMs are a versatile multivariate classification tool due to their non-parametric nature, robustness against outliers, reduced risk of training data over-fitting, quick and reliable convergence to a global optimum, and the availability of the kernel trick, which can yield non-linear hyperplanes [65,96]. The applicability of SVMs to assigning leaf pixels into drought stress classes was demonstrated by Asaari et al. [6] and Behmann et al. [7] for cereals. The high pixel classification accuracy and plausible spectral patterns that can be discerned in the obtained classes highlight the potential of SVMs to segment OSR images, a crop from a different agronomic group and a botanical family.

An obvious disadvantage of the adopted approach is the laborious pixel labelling. SVMs and their extensions give satisfactory predictions even when trained with small datasets. On the other hand, the modelling can fail when errors are present in the reference data [96]. Rather than reducing the size of the training pool, it would be more desirable to employ a solution that allows for dispensing of pixel labelling, especially considering the fact that it is challenging before stress symptoms are visible. In a maize drought phenotyping study, Asaari et al. [6] avoided this step by performing unsupervised classification on a reduced dataset and labelling the obtained clusters, rather than individual pixels, prior to SVM classification. Behmann et al. [7] devised a workflow based on ordinal clustering, which further facilitated the process, as only the extreme clusters needed to be labelled.

Scarce evidence of treatment effects was obtained from pixel counts representing fresh and dry pixel classes. The estimation uncertainty can, in part, be attributed to low pot counts in the control and regeneration treatments. The stronger reduction of the dry-pixel proportion in well-watered conditions estimated for 'Viking' relative to 'Cadeli' would be in agreement with the high drought sensitivity of this cultivar [15]. However, additional data are needed to confirm this pattern.

4.2. Vegetation Indexes

In a spring wheat experiment by Peteinatos et al. [52], water-stressed plants exhibited decreased MCARI values. One of the components on this index is the green reflectance. Consequently, the authors linked the observed effect to a reduced chlorophyll content, which is a common stress symptom in plants [24,97]. Similarly, in the present study, a higher MCARI in well-watered 'Viking' can be explained in terms of increased photosynthetic activity fostered by the favourable hydric conditions. The tendency towards minimising the periods of stomatal closure allows this water-spender to thrive in the control watering regime.

However, this interpretation can be questioned in light of the results obtained by Haboudane et al. [98] for maize. The authors reported a negative, rather than positive, relationship between the chlorophyll content and MCARI. At the same time, the relation was positive for MTCI, an index that employs reflectances around the red-edge [79]. The latter result was corroborated by Gitelson [99] for maize and soybean. The discrepancy pattern between MCARI and MTCI is in agreement with the present study findings. While watered 'Viking' exhibited higher MCARI than 'Cadeli' maintained under the same regime, the MTCI values were found to be lower ($ES_{Viking-Cadeli} = -0.53\,[-0.93, -0.12]$, $d_{Viking-Cadeli} = -1.18\,[-2.09, -0.26]$).

Contradictions of this kind reveal the problematic nature of relying on single vegetation indexes, at least as far as index means are concerned. In addition to the property of interest, the index value can be affected by additional confounding variables; in particular, the relationships tend to be crop-specific [98]. Interpretation of MCARI is especially challenging, given the erratic behaviour of this index for samples with a low chlorophyll content. It was shown that below a certain threshold, the relationship between MCARI and the pigment becomes reversed [81]. Such reports highlight the need for joint interpretation of multiple indexes, either in an informal fashion or by their further statistical processing [7,98].

The TCARI and TCARI/OSAVI indexes were originally developed in the context of chlorophyll content estimation [81], and their suitability to crop water status diagnosis can be linked to pigmentation changes in drought-affected tissues. Both were tested by Perry and Roberts [100] in a maize experiment, in which they discriminated between irrigated and unirrigated parts of the field. Just as for MCARI, the increased values of these indexes associated with the 'Viking' cultivar can be linked to its water-spender management strategy, but more data are needed to verify this finding.

Some of the index values varied more in values for well-watered 'Viking' than for well-watered 'Cadeli', while for certain others, the relationship was the opposite. An explanation linking these patterns to the differing water management strategies (a water-saver and a water-spender) seems dubious. More plausibly, the observed effects were determined by additional cultivar properties, particularly those related to the leaf surface and structure of the forming canopy [29]. Due to the dearth of studies comparing crop cultivars with respect to the variability of their spectral characteristics, the discussed results cannot be confronted with the findings of other authors. With regard to the lack of similar differences under the remaining watering regimes, it can be argued that the severity of pigmentation and structural (e.g., leaf shrinkage) changes caused by a drought episode [13,21,22] occluded the differences between the genotypes. An alternative explanation is the lower number of plant samples in the drought and regeneration treatments, making the effect estimates less precise, and treatment differences less likely detectable, as a further consequence.

RGI excelled among the vegetation indexes when evaluating their strength of response to restricted water availability. In their maize study, Sun et al. [5] associated the occurrence of a drought with an RGI increase of approximately 0.05 units on the index scale (point estimate inferred from the marginal estimates given in the paper), a value captured by the $ES_{watered-dry} = -0.96\,[-2.21, 0.21]$ and $-0.71\,[-1.97, 0.49]$ raw intervals obtained in the present study for 'Cadeli' and 'Viking', respectively. The potential usefulness of this index is further illustrated by its strong negative correlation to leaf water status indicators investigated by Rodríguez-Pérez et al. [101] in a commercial vineyard. Water availability revealed a positive influence on MTCI, RNDVI, and GI. The RNDVI difference ($ES_{watered-dry} = 0.19\,[-0.01, 0.41]$, $d_{watered-dry} = 2.10\,[-0.05, 4.47]$) is similar in magnitude to the spring wheat cultivars responses reported by Gutierrez et al. [102]. Depending on the crop developmental stage and the trial, RNDVI of the control plants exceeded the water-stressed treatment by 0.03 to 0.18 units (point estimates based on the marginal estimates mentioned in the paper). RNDVI is an NDVI-like index originally developed for woody species [75], and then employed to monitor cereal crops grown in areas with drought occurrence [102]. In light of the above findings, it seems to also be suited to OSR cultivation. The raw effect estimate obtained for GI, $ES_{watered-dry} = 0.44\,[-0.08, 0.95]$ ($d_{watered-dry} = 1.58\,[-0.29, 3.48]$), is in agreement with the difference between treatment means reported by Peteinatos et al. [52] for spring wheat (0.17 units). GI is a simple index combining green reflectance with the reflectance near the lower end of the red-edge. Despite its name ("greenness index"), in the present study its value seems to have been affected by the red-edge shift and flattening, rather than by changes in the green region, which appeared to be limited. Main et al. [103] published an extensive comparison of vegetation index performances with respect to the chlorophyll content prediction, which provides additional evidence of a weak GI response to the pigment signal.

One of the strengths of Bayesian statistics is the possibility of inferring an absence of a practical significance of an effect [49,54]. Surprisingly, the PRI_{norm} mean appeared to have been insensitive to the leaf spectra differences across both cultivars and individual watering regimes. The PRI family detects changes in crop photosynthetic radiation use efficiency by providing an insight into xanthophyll epoxidation processes [76,104]. According to Peñuelas et al. [105], this information is a better proxy of physiological status than total chlorophyll content. In a comparison of vegetation indexes by Rossini et al. [104], PRI_{570} turned out to be the best predictor of a range of maize water status indicators. In another maize study, the order of PRI_{570} values reflected the assignment of experimental plots to

91

irrigation levels and the timing of irrigation suppression [106]. The same index exhibited reliable correlations with several indicators of the winter wheat water status [107].

PRI is known to be sensitive to ambient illumination and other confounding factors [104,106]. Although in the present study the samples were placed in a photo tent to obtain diffuse illumination, a sensor was employed during the imagery acquisition to compensate for irradiance instability, and the linear model accounted for radiometric variability between the data cubes; the obtained correction might have been insufficient given the variable external conditions. In the case of the cultivar treatment, the examination of Cohen's *d* points to the overall small variability of this index as an alternative explanation of the obtained pattern with respect to PRI_{norm}. For a drought diagnosis based on proximal hyperspectral imaging and vegetation index means, we recommend avoiding days with unstable illumination conditions, unless artificial illumination is employed, or one or more calibration panels are included in every image.

Merzlyak et al. [73] proposed PSRI as an indicator index of leaf senescence, which can be triggered by water deprivation. This index was among the features discerning between barley drought senescence classes in the study by Behmann et al. [7]. Accordingly, the obtained PSRI standard deviation sensitivity to the contrasting watering regimes in 'Cadeli' can be linked to the source–sink character of the leaf senescence process [23]. TCARI and TCARI/OSAVI responded to the difference between the watered and rewatered treatments. Such a separation was not possible with the index means, suggesting that analysing the spectral variability is more suited to detecting a trace of a drought episode from which a crop did not necessarily fully recover. TCARI/OSAVI can perform better than TCARI and OSAVI by disentangling the effect of chlorophyll and LAI [81], as demonstrated by Haboudane et al. [98] and Perry and Roberts [100]. It was one of the indexes reported to reflect the maize physiological status in the Rossini et al. [104] drought experiment. It may seem that LAI plays a limited role in the present study, as the background is filtered out using segmentation. However, drought alters the structure of the foliage, leading to LAI modification accompanied by increased chlorophyll concentration in shrunken leaves [8,105], both affecting the reflectance spectrum. The remarkable overall consistency of the index standard deviations increasing with restricted watering corroborates the relationship between the stress level and symptom variability mentioned by Kruschke and Liddell [49]. In light of these findings, vegetation index standard deviations appear to be sensitive stress indicators in the context of drought diagnosis using proximal hyperspectral imaging, perhaps more so than the index means.

Biochemical and physiological parameters determined in the laboratory from leaf samples are reliable indicators of a crop status [108]. Drought stress occurrence is commonly assessed by analysing water content [15,24,44,104,107], pigment [5,15,24,108] and nutrient [5,8] concentrations, photosynthetic fluorescence [44] and photosynthetic [8,15] and transpiration [15] rates, or stomatal conductance [15]. Some measurements are possible in field conditions, such as leaf water potential [24,101], stomatal conductance [109], fluorescence [24,104,106,108], SPAD chlorophyll [98]; and leaf [104,106,109] and canopy [102] temperature.

Relationships between the parameter values and vegetation indexes were demonstrated in various drought studies. PRI_{570} and red edge position responded to chlorophyll and carotenoid concentrations in maize [5]. The former index was also sensitive to the changes in the pigment concentration ratio and leaf fluorescence [106]. In another maize study, PRI_{570} and TCARI/OSAVI exhibited strong relationships to chlorophyll fluorescence and leaf temperature [104]. Several published vegetation indexes predicted to a satisfactory degree leaf and canopy water contents of wheat, and further improvements were obtained by formulating custom indexes based on raw and derivative spectra [107]. Multiple indexes responded to water deficit in wheat caused by a powdery mildew infection [108]. Rodríguez-Pérez et al. [101] obtained high correlations between grapevine leaf water contents and indexes derived from spectra subjected to continuum removal.

OSR readily responds to drought stress in terms of biochemical and physiological indicators. Clear differentiation between control and stressed plants was obtained by Urban et al. [15], with the

differences especially pronounced for net photosynthetic rate, stomatal conductance, leaf transpiration rate, evapotranspiration change, and proline content. Pasban Eslam [109] reported a consistent modification of leaf relative water content, stomatal conductance, and temperature across five OSR cultivars over two years of his experiment. In another experiment, water deprivation was associated with decreased leaf fluorescence and an osmolarity increase [44]. The present study related the patterns of vegetation index values to the experimental treatments: the OSR cultivar and the watering regime. In the light of the cited results, it is plausible that a number of the obtained effects could be replicated in an observational study, in which the treatments would be replaced with biochemical and physiological parameter measurements at the linear modelling step. Further research is needed to verify this expectation.

4.3. Full Spectrum Information

The distribution of the leaf pixel spectra in the principal component space differentiated the regeneration treatment from the remaining investigated watering regimes. The indicated decrease in the NIR reflectance for the raw spectra and the spectra subjected to SGF can be linked to leaf cell structure alteration by stress [19]. However, this interpretation is contradicted by the redshift revealed by the MSC transformation, indicative of good hydration [20]. Similarly, the steep red-edge pattern obtained for the finite differences derivation can be associated with an increased chlorophyll concentration [24]. These patterns need to be approached with caution, considering the fact that PCA does not account for the experimental design resulting in the hierarchical structure of the dataset. Moreover, only one data cube representing the rewatered treatment was analysed in the present study, as registration failed for the other one, which had to be discarded. The obtained pixel clusters could be the result of specific illumination conditions at the moment of image capture, which dominated the spectral signal [29].

Regarding the double derivation, the apparent differences in pixel extents are in agreement with the preceding part of the analysis, which revealed higher standard deviations of vegetation indexes derived for the drought treatment relative to the control plants. One could suspect that the occurrence of two large clusters for the MSC-transformed spectra is related to changing ambient illumination conditions or an uneven distribution of radiant energy inside the photo tent. In that case, each cluster would contain pixels associated with individual images or pot positions, respectively. However, none of those hypotheses was confirmed by consulting the dataset.

The little-varied PC1 scores obtained for 'Viking' compared to 'Cadeli' can be explained in terms of the higher stress level of the latter. 'Cadeli' tends to restrict stomatal conductance [45], which is a suboptimal strategy in the conditions of high water availability, as photosynthesis is impaired [13]. The discussed treatment separation was apparent only after subjecting the spectra to MSC. This pre-processing is known to remove some scatter and baseline shift artefacts [85]. In the present study, it might have mitigated the influence of variable illumination conditions on the captured hyperspectral data cubes. A question arises whether a similar improvement would have been achieved in the vegetation index part of the analysis if they had been derived from the MSC-pre-processed rather than the raw spectra.

The high variation of the dry pixel spectra subjected to double derivation suggested by the PCA analysis is absent from the results of the linear modelling. A possible explanation might be a high noise characterising derivative spectra [85], and subject to compounding when the operation is repeated [105]. The derivation might have also been negatively affected by the low spectral resolution of the analysed hyperspectral data cubes, which precluded a detailed reconstruction of the spectra shapes. Peñuelas et al. [105] reported an improved relationship between the second order derivative indexes and sunflower leaf water potential relative to the principal components and indexes derived from the raw spectra, but their data were acquired with a fivefold higher spectral resolution than in the present study. Finally, the double derivation linear model posed problems for the MCMC sampler [59], with detrimental consequences for the reliability of the obtained posterior distributions. In future

studies of this kind, it is recommended that the derivation be combined with smoothing [85] and that the spectral resolution of the imagery be maximised, even at a price of an increased data volume [110] and information redundancy [7].

The proximity of the extreme watering regimes in terms of the PC1 standard deviations is counterintuitive and in disagreement with the results of the vegetation index part of the study. The fact that each pre-processing resulted in distinctive PC loading vectors precludes a straightforward interpretation in terms of the spectral regions. PCA is an unsupervised dimensionality reduction method. Compared to the SVM approach, it does not require pixel labelling, and compared to the vegetation index approach, it does not involve an arbitrary choice of indexes, the performance of which is site-specific. On the other hand, the obtained principal axes do not necessarily need to be related to factors of interest. The obtained result is problematic, but nevertheless interesting. Similarly to the TCARI and TCARI/OSAVI standard deviations, it may point to a way of detecting a trace of a severe drought episode in a seemingly healthy and well-hydrated crop. The signal attenuation obtained for the finite differences derivation can be linked to the capability of this transformation to filter out the effects of structural differences between crop cultivars [111]. For MSC, it can be associated with the removal of illumination artefacts [85].

5. Conclusions

We investigated the feasibility of a 2D frame hyperspectral camera as a proximal sensor to detect drought stress of juvenile plants of two oilseed rape cultivars with different water management strategies in semi-controlled, outdoor conditions. A support vector machine accurately distinguished between normal leaf pixels and those bearing drought symptoms. Only 2 of the 47 model validation pixels were misclassified, though time-consuming labelling was required to train the classifier. Based on the pixel assignment, some evidence of leaf discolouration was obtained for the drought-stressed 'Viking', in accord with the provenance of this cultivar. The ratio between the number of dry-labelled pixels in the control and stress watering regimes was estimated as $0.009 \, [0.00005, 1.6]$.

Several vegetation index means responded to the difference between the control and water-deprived plants, especially RGI, MTCI, RNDVI, and GI; while none of the tested PRI indexes distinguished among the treatments. RGI excelled among the vegetation indexes in terms of effect strengths, which amounted to $-0.96 \, [-2.21, 0.21]$ and $-0.71 \, [-1.97, 0.49]$ units for each cultivar with respect to the watered–dry treatment contrast.

The most striking finding was a consistent increase in the multiple index standard deviations to worsening of the hydric regime. The increases occurred not only in the dry treatment but also for plants subjected to regeneration after a drought episode. This result suggests a higher sensitivity of the vegetation index variability measures relative to the means for oilseed rape drought stress diagnosis. It also justifies the application of imaging spectroscopy to capture these effects. Especially clear responses were obtained for RGI, PSRI, TCARI, and TCARI/OSAVI. Some of the patterns involved also the regeneration watering regime. In particular, PSRI standard deviation for 'Cadeli' differed by a factor of $0.33 \, [0.16, 0.68]$ between the rewatered and dry treatments. It seems worthwhile to include RGI in similar studies in the future given the fact that both the mean and standard deviation (a multiplicative effect of $0.10 \, [0.07, 0.16]$ for the watered–dry contrast in the case of 'Cadeli') of this index were affected by the water availability.

The drought stress could be discerned in the spectral signatures when regeneration was still possible. On the other hand, the symptoms were already visible to the naked eye. Additional factors can be introduced in follow-up studies to verify the robustness of the findings and their application to earlier drought stress detection. A single campaign could be replaced by a time series to capture the temporal development of the drought stress and of the spectral responses. Another modification would be to restrict the watering of the plants at an earlier developmental phase and investigate which of the spectral stress indicators remain viable for younger plants. Additional insights could be obtained by augmenting the new dataset with biochemical and physiological measurements.

Despite the unstable light conditions during the imaging campaign, the experimental treatments had strong and consistent effects on some of the examined spectral indicators and can be interpreted in terms of their robustness. However, although several measures were taken to mitigate the variable illumination effects, it cannot be ruled out that the observed patters were artefacts caused by the external conditions, instead. For this reason, regardless of the study extensions, the obtained results need to be replicated in an independent experiment with a larger sample, an improved design, and stricter precautions with respect to illumination stability during imagery acquisition.

References

1. Lobell, D.B.; Schlenker, W.; Costa-Roberts, J. Climate Trends and Global Crop Production Since 1980. *Science* **2011**, *333*, 616–620. [CrossRef] [PubMed]
2. Naumann, G.; Alfieri, L.; Wyser, K.; Mentaschi, L.; Betts, R.; Carrao, H.; Spinoni, J.; Vogt, J.; Feyen, L. Global Changes in Drought Conditions Under Different Levels of Warming. *Geophys. Res. Lett.* **2018**, *45*, 3285–3296. [CrossRef]
3. Arnell, N.; Lowe, J.; Challinor, A.; Osborn, T. Global and regional impacts of climate change at different levels of global temperature increase. *Clim. Chang.* **2019**, *155*, 377–391. [CrossRef]
4. Daryanto, S.; Wang, L.; Jacinthe, P.A. Global synthesis of drought effects on cereal, legume, tuber and root crops production: A review. *Agric. Water Manag.* **2017**, *179*, 18–33. [CrossRef]
5. Sun, C.; Li, C.; Zhang, C.; Hao, L.; Song, M.; Liu, W.; Zhang, Y. Reflectance and biochemical responses of maize plants to drought and re-watering cycles. *Ann. Appl. Biol.* **2018**, *172*, 332–345. [CrossRef]
6. Asaari, M.S.M.; Mertens, S.; Dhondt, S.; Inzé, D.; Wuyts, N.; Scheunders, P. Analysis of hyperspectral images for detection of drought stress and recovery in maize plants in a high-throughput phenotyping platform. *Comput. Electron. Agric.* **2019**, *162*, 749–758. [CrossRef]
7. Behmann, J.; Steinrücken, J.; Plümer, L. Detection of early plant stress responses in hyperspectral images. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 98–111. [CrossRef]
8. Linke, R.; Richter, K.; Haumann, J.; Schneider, W.; Weihs, P. Occurrence of repeated drought events: Can repetitive stress situations and recovery from drought be traced with leaf reflectance? *Period. Biol.* **2008**, *110*, 219–229.
9. Buezo, J.; Sanz-Saez, Á.; Moran, J.F.; Soba, D.; Aranjuelo, I.; Esteban, R. Drought tolerance response of high-yielding soybean varieties to mild drought: Physiological and photochemical adjustments. *Physiol. Plant.* **2019**, *166*, 88–104. [CrossRef]
10. Gilbert, M.E.; Zwieniecki, M.A.; Holbrook, N.M. Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use efficiency in 11 soybean genotypes before and during mild drought. *J. Exp. Bot.* **2011**, *62*, 2875–2887. [CrossRef]
11. Blum, A. Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive? *Aust. J. Agric. Res.* **2005**, *56*, 1159–1168. [CrossRef]
12. Raza, M.A.S.; Shahid, A.M.; Saleem, M.F.; Khan, I.H.; Ahmad, S.; Ali, M.; Iqbal, R. Effects and management strategies to mitigate drought stress in oilseed rape (*Brassica napus* L.): A review. *Zemdirb. Agric.* **2017**, *104*, 85–94. [CrossRef]
13. Ashraf, M.; Harris, P. Photosynthesis under stressful environments: An overview. *Photosynthetica* **2013**, *51*, 163–190. [CrossRef]

14. Flexas, J.; Bota, J.; Loreto, F.; Cornic, G.; Sharkey, T. Diffusive and metabolic limitations to photosynthesis under drought and salinity in C$_3$ plants. *Plant Biol.* **2004**, *6*, 269–279. [CrossRef] [PubMed]

15. Urban, M.O.; Vašek, J.; Klíma, M.; Krtková, J.; Kosová, K.; Prášil, I.T.; Vítámvás, P. Proteomic and physiological approach reveals drought-induced changes in rapeseeds: Water-saver and water-spender strategy. *J. Proteom.* **2017**, *152*, 188–205. [CrossRef] [PubMed]

16. Nakhforoosh, A.; Bodewein, T.; Fiorani, F.; Bodner, G. Identification of Water Use Strategies at Early Growth Stages in Durum Wheat from Shoot Phenotyping and Physiological Measurements. *Front. Plant Sci.* **2016**, *7*, 1155. [CrossRef] [PubMed]

17. Roche, D. Stomatal conductance is essential for higher yield potential of C$_3$ crops. *Crit. Rev. Plant Sci.* **2015**, *34*, 429–453. [CrossRef]

18. Kim, Y.; Glenn, D.M.; Park, J.; Ngugi, H.K.; Lehman, B.L. Hyperspectral image analysis for water stress detection of apple trees. *Comput. Electron. Agric.* **2011**, *77*, 155–160. [CrossRef]

19. Knipling, E.B. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sens. Environ.* **1970**, *1*, 155–159. [CrossRef]

20. Carter, G.A.; Knapp, A.K. Leaf optical properties in higher plants: Linking spectral characteristics to stress and chlorophyll concentration. *Am. J. Bot.* **2001**, *88*, 677–684. [CrossRef]

21. Gill, S.S.; Tuteja, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem.* **2010**, *48*, 909–930. [CrossRef] [PubMed]

22. Din, J.; Khan, S.; Ali, I.; Gurmani, A. Physiological and agronomic response of canola varieties to drought stress. *J. Anim. Plant Sci.* **2011**, *21*, 78–82.

23. Munné-Bosch, S.; Alegre, L. Die and let live: Leaf senescence contributes to plant survival under drought stress. *Funct. Plant Biol.* **2004**, *31*, 203–216. [CrossRef] [PubMed]

24. Govender, M.; Govender, P.J.; Weiersbye, I.M.; Witkowski, E.T.F.; Ahmed, F. Review of commonly used remote sensing and ground-based technologies to measure plant water stress. *Water SA* **2009**, *35*, 741–752. [CrossRef]

25. van der Werff, H.; van der Meijde, M.; Jansma, F.; van der Meer, F.; Groothuis, G.J. A Spatial-Spectral Approach for Visualization of Vegetation Stress Resulting from Pipeline Leakage. *Sensors* **2008**, *8*, 3733–3743. [CrossRef]

26. Adão, T.; Hruška, J.; Pádua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J.J. Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sens.* **2017**, *9*, 1110. [CrossRef]

27. Lowe, A.; Harrison, N.; French, A.P. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods* **2017**, *13*, 80. [CrossRef]

28. Khan, M.J.; Khan, H.S.; Yousaf, A.; Khurshid, K.; Abbas, A. Modern Trends in Hyperspectral Image Analysis: A Review. *IEEE Access* **2018**, *6*, 14118–14129. [CrossRef]

29. Mishra, P.; Asaari, M.S.M.; Herrero-Langreo, A.; Lohumi, S.; Diezma, B.; Scheunders, P. Close range hyperspectral imaging of plants: A review. *Biosyst. Eng.* **2017**, *164*, 49–67. [CrossRef]

30. Ge, Y.; Bai, G.; Stoerger, V.; Schnable, J.C. Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Comput. Electron. Agric.* **2016**, *127*, 625–632. [CrossRef]

31. Kumar, A.; Bharti, V.; Kumar, V.K.U.; Meena, P. Hyperspectral imaging: A potential tool for monitoring crop infestation, crop yield and macronutrient analysis, with special emphasis to Oilseed Brassica. *J. Oilseed Brassica* **2016**, *7*, 113–125.

32. Bruning, B.; Liu, H.; Brien, C.; Berger, B.; Lewis, M.; Garnett, T. The Development of Hyperspectral Distribution Maps to Predict the Content and Distribution of Nitrogen and Water in Wheat (*Triticum aestivum*). *Front. Plant Sci.* **2019**, *10*. [CrossRef] [PubMed]

33. Römer, C.; Wahabzada, M.; Ballvora, A.; Pinto, F.; Rossini, M.; Panigada, C.; Behmann, J.; Léon, J.; Thurau, C.; Bauckhage, C.; et al. Early drought stress detection in cereals: Simplex volume maximisation for hyperspectral image analysis. *Funct. Plant Biol.* **2012**, *39*, 878–890. [CrossRef] [PubMed]

34. Zarco-Tejada, P.; González-Dugo, V.; Berni, J. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sens. Environ.* **2012**, *117*, 322–337. [CrossRef]

35. Zovko, M.; Žibrat, U.; Knapič, M.; Kovačić, M.B.; Romić, D. Hyperspectral remote sensing of grapevine drought stress. *Precis. Agric.* **2019**, *20*, 335–347. [CrossRef]

36. Sabagh, A.E.; Hossain, A.; Barutçular, C.; Islam, M.S.; Ratnasekera, D.; Kumar, N.; Meena, R.S.; Gharib, H.S.; Saneoka, H.; da Silva, J.A.T. Drought and salinity stress management for higher and sustainable canola ('*Brassica napus*' L.) production: A critical review. *Aust. J. Crop. Sci.* **2019**, *13*, 88–96. [CrossRef]

37. Xia, J.; Yang, Y.W.; Cao, H.X.; Zhang, W.; Xu, L.; Wang, Q.; Ke, Y.; Zhang, W.; Ge, D.; Huang, B. Hyperspectral Identification and Classification of Oilseed Rape Waterlogging Stress Levels Using Parallel Computing. *IEEE Access* **2018**, *6*, 57663–57675. [CrossRef]

38. Högy, P.; Franzaring, J.; Schwadorf, K.; Breuer, J.; Schütze, W.; Fangmeier, A. Effects of free-air CO_2 enrichment on energy traits and seed quality of oilseed rape. *Agric. Ecosyst. Environ.* **2010**, *139*, 239–244. [CrossRef]

39. Zhang, X.; Lu, G.; Long, W.; Zou, X.; Li, F.; Nishio, T. Recent progress in drought and salt tolerance studies in *Brassica* crops. *Breed. Sci.* **2014**, *64*, 60–73. [CrossRef] [PubMed]

40. Bonjean, A.P.; Dequidt, C.; Sang, T. Rapeseed in China. *OCL* **2016**, *23*, D605. [CrossRef]

41. Kumar, A.; Bharti, V.; Kumar, V.; Meena, P.; Suresh, G. Hyperspectral imaging applications in rapeseed and mustard farming. *J. Oilseeds Res.* **2017**, *34*, 1–8.

42. Majidi, M.; Rashidi, F.; Sharafi, Y. Physiological traits related to drought tolerance in *Brassica*. *Int. J. Plant Prod.* **2015**, *9*, 4.

43. Tesfamariam, E.H.; Annandale, J.G.; Steyn, J.M. Water Stress Effects on Winter Canola Growth and Yield. *Agron. J.* **2010**, *102*, 658–666. [CrossRef]

44. Müller, T.; Lüttschwager, D.; Lentzsch, P. Recovery from drought stress at the shooting stage in oilseed rape (*Brassica napus*). *J. Agron. Crop. Sci.* **2010**, *196*, 81–89. [CrossRef]

45. Kosová, K.; Klíma, M.; Vítamvás, P.; Prášil, I.T. Odezva vybraných odrůd řepky na sucho a následná regenerace [Response of selected oilseed rape cultivars to drought and subsequent recovery]. *Úroda* **2018**, *66*, 19–25.

46. Piekarczyk, J.; Sulewska, H.; Szymańska, G. Winter oilseed-rape yield estimates from hyperspectral radiometer measurements. *Quaest. Geogr.* **2011**, *30*, 77–84. [CrossRef]

47. Zhang, X.; He, Y. Rapid estimation of seed yield using hyperspectral images of oilseed rape leaves. *Ind. Crop. Prod.* **2013**, *42*, 416–420. [CrossRef]

48. Kong, W.; Liu, F.; Zhang, C.; Zhang, J.; Feng, H. Non-destructive determination of Malondialdehyde (MDA) distribution in oilseed rape leaves by laboratory scale NIR hyperspectral imaging. *Sci. Rep.* **2016**, *6*. [CrossRef] [PubMed]

49. Kruschke, J.K.; Liddell, T.M. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychon. Bull. Rev.* **2017**, 1–29. [CrossRef] [PubMed]

50. Nansen, C. Use of Variogram Parameters in Analysis of Hyperspectral Imaging Data Acquired from Dual-Stressed Crop Leaves. *Remote Sens.* **2012**, *4*, 180–193. [CrossRef]

51. Jay, S.; Hadoux, X.; Gorretta, N.; Rabatel, G. Potential of hyperspectral imagery for nitrogen content retrieval in sugar beet leaves. In Proceedings of the Proceedings International Conference of Agricultural Engineering, Zurich, Switzerland, 6–10 July 2014; pp. 1–8.

52. Peteinatos, G.G.; Korsaeth, A.; Berge, T.W.; Gerhards, R. Using Optical Sensors to Identify Water Deprivation, Nitrogen Shortage, Weed Presence and Fungal Infection in Wheat. *Agriculture* **2016**, *6*, 24. [CrossRef]

53. van Zyl, C.J. Frequentist and Bayesian inference: A conceptual primer. *New Ideas Psychol.* **2018**, *51*, 44–49. [CrossRef]

54. Zyphur, M.J.; Oswald, F.L. Bayesian estimation and inference: A user's guide. *J. Manag.* **2015**, *41*, 390–420. [CrossRef]

55. Che, X.; Xu, S. Bayesian data analysis for agricultural experiments. *Can. J. Plant Sci.* **2010**, *90*, 575–603. [CrossRef]

56. Gelfand, A.E. Gibbs Sampling. *J. Am. Stat. Assoc.* **2000**, *95*, 1300–1304. [CrossRef]

57. Visser, M.D.; McMahon, S.M.; Merow, C.; Dixon, P.M.; Record, S.; Jongejans, E. Speeding Up Ecological and Evolutionary Computations in R; Essentials of High Performance Computing for Biologists. *PLoS Comput. Biol.* **2015**, *11*, e1004140. [CrossRef] [PubMed]

58. Bürkner, P.C. Advanced Bayesian Multilevel Modeling with the R Package brms. *R J.* **2018**, *10*, 395–411. [CrossRef]

59. Carpenter, B.; Gelman, A.; Hoffman, M.D.; Lee, D.; Goodrich, B.; Betancourt, M.; Brubaker, M.; Guo, J.; Li, P.; Riddell, A. Stan: A Probabilistic Programming Language. *J. Stat. Softw.* **2017**, *76*, 1–32. [CrossRef]

60. Salvatier, J.; Wiecki, T.V.; Fonnesbeck, C. Probabilistic programming in Python using PyMC3. *PeerJ Comput. Sci.* **2016**, *2*, e55. [CrossRef]

61. Smith, G.M.; Milton, E.J. The use of the empirical line method to calibrate remotely sensed data to reflectance. *Int. J. Remote Sens.* **1999**, *20*, 2653–2662. [CrossRef]

62. Filzmoser, P.; Hron, K.; Reimann, C. Univariate statistical analysis of environmental (compositional) data: Problems and possibilities. *Sci. Total Environ.* **2009**, *407*, 6100–6108. [CrossRef] [PubMed]

63. Roeder, A.H.; Cunha, A.; Burl, M.C.; Meyerowitz, E.M. A computational image analysis glossary for biologists. *Development* **2012**, *139*, 3071–3080. [CrossRef] [PubMed]

64. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571. [CrossRef]

65. Karimi, Y.; Prasher, S.; Madani, A.; Kim, S. Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations. *Can. Biosyst. Eng.* **2008**, *50*, 13–20.

66. Bischl, B.; Richter, J.; Bossek, J.; Horn, D.; Thomas, J.; Lang, M. mlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions. *arXiv* **2018**, arXiv:1703.03373.

67. Amigo, J.M.; Babamoradi, H.; Elcoroaristizabal, S. Hyperspectral image analysis. A tutorial. *Anal. Chim. Acta* **2015**, *896*, 34–51. [CrossRef] [PubMed]

68. Gelman, A.; Jakulin, A.; Pittau, M.G.; Su, Y.S.; others. A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* **2008**, *2*, 1360–1383. [CrossRef]

69. Serrano, L.; Filella, I.; Peñuelas, J. Remote Sensing of Biomass and Yield of Winter Wheat under Different Nitrogen Supplies. *Crop Sci.* **2000**, *40*, 723–731. [CrossRef]

70. Zarco-Tejada, P.J.; Berjón, A.; López-Lozano, R.; Miller, J.R.; Martín, P.; Cachorro, V.; González, M.; De Frutos, A. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sens. Environ.* **2005**, *99*, 271–287. [CrossRef]

71. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]

72. Roujean, J.L.; Breon, F.M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [CrossRef]

73. Merzlyak, M.N.; Gitelson, A.A.; Chivkunova, O.B.; Rakitin, V.Y. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* **1999**, *106*, 135–141. [CrossRef]

74. Blackburn, G.A. Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *Int. J. Remote Sens.* **1998**, *19*, 657–675. [CrossRef]

75. Gitelson, A.; Merzlyak, M.N. Spectral Reflectance Changes Associated with Autumn Senescence of *Aesculus hippocastanum* L. and *Acer* platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *J. Plant Physiol.* **1994**, *143*, 286–292. [CrossRef]

76. Gamon, J.A.; Serrano, L.; Surfus, J.S. The photochemical reflectance index: An optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia* **1997**, *112*, 492–501. [CrossRef] [PubMed]

77. Hernández-Clemente, R.; Navarro-Cerrillo, R.M.; Suárez, L.; Morales, F.; Zarco-Tejada, P.J. Assessing structural effects on PRI for stress detection in conifer forests. *Remote Sens. Environ.* **2011**, *115*, 2360–2375. [CrossRef]

78. Zarco-Tejada, P.J.; González-Dugo, V.; Williams, L.; Suárez, L.; Berni, J.A.; Goldhamer, D.; Fereres, E. 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.* **2013**, *138*, 38–50. [CrossRef]

79. Dash, J.; Curran, P. The MERIS terrestrial chlorophyll index. *Int. J. Remote Sens.* **2004**, *25*, 5403–5413. [CrossRef]

80. Daughtry, C.; Walthall, C.; Kim, M.; De Colstoun, E.B.; McMurtrey, J., III. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [CrossRef]

81. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [CrossRef]

82. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [CrossRef]

83. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [CrossRef] [PubMed]

84. Gelman, A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* **2006**, *1*, 515–534. [CrossRef]

85. Rinnan, Å.; van den Berg, F.; Engelsen, S.B. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal. Chem.* **2009**, *28*, 1201–1222. [CrossRef]

86. Cumming, G. The New Statistics: Why and How. *Psychol. Sci.* **2014**, *25*, 7–29. [CrossRef] [PubMed]

87. Courtès, L.; Wurmus, R. *Reproducible and User-Controlled Software Environments in HPC with Guix*; Euro-Par 2015: Parallel Processing Workshops; Hunold, S., Costan, A., Giménez, D., Alexandru, I., Ricci, L., Gómez Requena, M.E., Scarano, V., Verbanescu, A.L., Scott, S.L., Lankes, S., Eds.; Vienna University of Technology: Vienna, Austria, 2015; pp. 579–591. [CrossRef]

88. Feldman, S.I. Make—A program for maintaining computer programs. *Softw. Pract. Exp.* **1979**, *9*, 255–265. [CrossRef]

89. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.

90. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, 2019. Available online: https://cran.r-project.org/package=e1071 (accessed on 21 October 2020).

91. Bischl, B.; Lang, M.; Kotthoff, L.; Schiffner, J.; Richter, J.; Studerus, E.; Casalicchio, G.; Jones, Z.M. mlr: Machine Learning in R. *J. Mach. Learn. Res.* **2016**, *17*, 1–5.

92. Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L.; Wehberg, J.; Wichmann, V.; Böhner, J. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev. Discuss.* **2015**, *8*. [CrossRef]

93. Brenning, A.; Bangs, D.; Becker, M. RSAGA: SAGA Geoprocessing and Terrain Analysis. 2018. Available online: https://cran.r-project.org/package=RSAGA (accessed on 21 October 2020).

94. Bradski, G. The OpenCV Library. *Dobb'S J. Softw. Tools* **2000**, *25*, 120–125.

95. Honkavaara, E.; Khoramshahi, E. Radiometric Correction of Close-Range Spectral Image Blocks Captured Using an Unmanned Aerial Vehicle with a Radiometric Block Adjustment. *Remote Sens.* **2018**, *10*, 256. [CrossRef]

96. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [CrossRef]

97. Jacquemoud, S.; Ustin, S.L. Leaf optical properties: A state of the art. In Proceedings of the 8th International Symposium of Physical Measurements & Signatures in Remote Sensing, Aussois, France, 8–12 January 2001; pp. 223–332.

98. Haboudane, D.; Tremblay, N.; Miller, J.R.; Vigneault, P. Remote Estimation of Crop Chlorophyll Content Using Spectral Indices Derived from Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 423–437. [CrossRef]

99. Gitelson, A.A. *Hyperspectral Remote Sensing of Vegetation*; chapter Remote Sensing Estimation of Crop Biophysical Characteristics at Various Scales; Thenkabail, P.S.; Lyon, J.G., Eds.; CRC Press: Boca Raton, FL, USA, 2016; pp. 329–358.

100. Perry, E.M.; Roberts, D.A. Sensitivity of Narrow-Band and Broad-Band Indices for Assessing Nitrogen Availability and Water Stress in an Annual Crop. *Agron. J.* **2008**, *100*, 1211–1219. [CrossRef]

101. Rodríguez-Pérez, J.R.; Riaño, D.; Carlisle, E.; Ustin, S.; Smart, D.R. Evaluation of Hyperspectral Reflectance Indexes to Detect Grapevine Water Status in Vineyards. *Am. J. Enol. Vitic.* **2007**, *58*, 302–317.

102. Gutierrez, M.; Reynolds, M.P.; Raun, W.R.; Stone, M.L.; Klatt, A.R. Spectral Water Indices for Assessing Yield in Elite Bread Wheat Genotypes under Well-Irrigated, Water-Stressed, and High-Temperature Conditions. *Crop Sci.* **2010**, *50*, 197–214. [CrossRef]

103. Main, R.; Cho, M.A.; Mathieu, R.; O'Kennedy, M.M.; Ramoelo, A.; Koch, S. An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 751–761. [CrossRef]

104. Rossini, M.; Fava, F.; Cogliati, S.; Meroni, M.; Marchesi, A.; Panigada, C.; Giardino, C.; Busetto, L.; Migliavacca, M.; Amaducci, S.; et al. Assessing canopy PRI from airborne imagery to map water stress in maize. *ISPRS J. Photogramm. Remote Sens.* **2013**, *86*, 168–177. [CrossRef]

105. Peñuelas, J.; Gamon, J.; Fredeen, A.; Merino, J.; Field, C. Reflectance indices associated with physiological changes in nitrogen-and water-limited sunflower leaves. *Remote Sens. Environ.* **1994**, *48*, 135–146. [CrossRef]

106. Chou, S.; Chen, J.M.; Yu, H.; Chen, B.; Zhang, X.; Croft, H.; Khalid, S.; Li, M.; Shi, Q. Canopy-Level Photochemical Reflectance Index from Hyperspectral Remote Sensing and Leaf-level Non-photochemical Quenching as Early Indicators of Water Stress in Maize. *Remote Sens.* **2017**, *9*, 794. [CrossRef]

107. Sun, H.; Feng, M.; Xiao, L.; Yang, W.; Wang, C.; Jia, X.; Zhao, Y.; Zhao, C.; Muhammad, S.K.; Li, D. Assessment of plant water status in winter wheat (*Triticum aestivum* L.) based on canopy spectral indices. *PLoS ONE* **2019**, *14*, e0216890. [CrossRef]

108. Feng, W.; Qi, S.; Heng, Y.; Zhou, Y.; Wu, Y.; Liu, W.; He, L.; Li, X. Canopy Vegetation Indices from *In situ* Hyperspectral Data to Assess Plant Water Status of Winter Wheat under Powdery Mildew Stress. *Front. Plant Sci.* **2017**, *8*, 1219. [CrossRef]

109. Pasban Eslam, B. Evaluation of Physiological Indices, Yield and its Components as Screening Techniques for Water Deficit Tolerance in Oilseed Rape Cultivars. *J. Agric. Sci. Technol.* **2009**, *11*, 413–422.

110. Salvador, M.; Resmini, R. *Data Mining for Geoinformatics*; chapter Computation in Hyperspectral Imagery (HSI) Data Analysis: Role and Opportunities; Cervone, G., Lin, J., Waters, N., Eds.; Springer: New York, NY, USA, 2014; pp. 1–27.

111. Danson, F.; Steven, M.; Malthus, T.; Clark, J. High-spectral resolution data for determining leaf water content. *Int. J. Remote Sens.* **1992**, *13*, 461–470. [CrossRef]

Appendix B

Wiktor R. Żelazny, Jana Chrpová, and Pavel Hamouz: Fusarium head blight detection from spectral measurements in a field phenotyping setting — a pre-registered study

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Research Paper

Fusarium head blight detection from spectral measurements in a field phenotyping setting — A pre-registered study

Wiktor R. Żelazny [a,b,*], *Jana Chrpová* [c], *Pavel Hamouz* [d]

[a] *Division of Crop Management Systems, Crop Research Institute, Drnovská 507/73, 161 06 Praha 6 Ruzyně, Czech Republic*
[b] *Faculty of Engineering, Czech University of Life Sciences Prague, Kamýcká 129, 165 00 Praha 6 Suchdol, Czech Republic*
[c] *Division of Crop Genetics and Breeding, Crop Research Institute, Drnovská 507/73, 161 06 Praha 6 Ruzyně, Czech Republic*
[d] *Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Kamýcká 129, 165 00 Praha 6 Suchdol, Czech Republic*

Spectroscopic methods can contribute to addressing the field phenotyping bottleneck problem in crop breeding programs. In disease resistance phenotyping, spectral signatures can be analysed to derive infection severity scores and to screen breeding lines. Hyperspectra of winter wheat spikes were acquired in a Fusarium head blight phenotyping trial at the milk- and wax-ripening phenological phases. Disease severity ratings were simultaneously performed by an expert on a 9-point visual scale. Ordinal support vector machine models were then trained to assign hill plots to the individual severity levels. The predictive models' performance was evaluated for data collection timing, spectral preprocessing and permitted rating-error tolerance. The models trained to spectra acquired at the milk-ripening phase were sufficiently accurate to reliably distinguish between low, medium and high symptom severity; with accuracy approaching 100% for two-point error tolerance. However, deterioration in prediction quality was noted for the wax-ripening campaign, presumably due to spike-drying. After aggregation of the spectra using the median function no gain could be associated with further pre-processing. Modest performance improvements obtained with two schemes do not justify the additional data acquisition costs involved, but standard normal variate could be advantageous for some scenarios with mean-aggregated spectra. In addition to phenotyping, the results are discussed in relation to large-scale farming applications. Elevated infection risk detection prior to anthesis is recommended for fungicide treatment, considering the pathogen biology. The study is accompanied by a publicly-available dataset and the computational scripts employed to obtain the results.

* *Corresponding author*. Division of Crop Management Systems, Crop Research Institute, Drnovská 507/73, 161 06 Praha 6 Ruzyně, Czech Republic. Fax: +420 233 310 638.
E-mail addresses: wzelazny@vurv.cz (W.R. Żelazny), chrpova@vurv.cz (J. Chrpová), hamouzp@af.czu.cz (P. Hamouz).

Nomenclature	
2D	two-dimensional
dai	days after inoculation
FHB	Fusarium head blight
GS	BBCH growth stage
GSD	ground sample distance
NIR	near infrared
NPK	nitrogen, phosphorus and potassium
RGB	red, green and blue
SNV	standard normal variate
SVM	support vector machine
SWIR	short-wave infrared
UAV	unmanned aerial vehicle
Vis-NIR	visible and near infrared
VSS	visual symptom score

1. Introduction

Fusarium head blight (FHB, scab) is a cereal fungal disease caused by *Fusarium* spp. (Jaillais et al., 2015; Khaledi et al., 2017; Saccon et al., 2017; Mielniczuk & Skwaryło-Bednarz, 2020). The pathogen severely impairs yield, and the grain quality is greatly diminished by mycotoxin action (Bauriegel & Herppich, 2014; Cambaza et al., 2019; Mielniczuk & Skwaryło-Bednarz, 2020). As in numerous other diseases, FHB has patchy distribution in crops (Jin et al., 2018; Liu et al., 2020b), and this hinders its monitoring at the field spatial scale. Infection occurrences have been increasing because of the high prevalence of maize in crop rotations (Dammer et al., 2011; Bauriegel & Herppich, 2014) and due to the adoption of reduced tillage systems by farmers (McMullen et al., 2012; Gilbert & Haber, 2013; Mielniczuk & Skwaryło-Bednarz, 2020). Climate change may also aggravate this problem in the future (Gilbert & Haber, 2013; Vaughan et al., 2016).

FHB control is based on cultural preventive measures and fungicide applications (McMullen et al., 2012; Mielniczuk & Skwaryło-Bednarz, 2020), the latter of which are scheduled according to the disease risk levels predicted from weather forecasts (McMullen et al., 2012; Xiao et al., 2020). There is also ongoing research into biological control agents (McMullen et al., 2012; Mielniczuk & Skwaryło-Bednarz, 2020). The development of resistant cultivars is a further major influence in decreasing FHB occurrence (Gilbert & Haber, 2013; Buerstmayr et al., 2020; Mielniczuk & Skwaryło-Bednarz, 2020).

Progress in breeding programs is constrained by the field phenotyping bottleneck (McMullen et al., 2012; Steiner et al., 2017). In particular, the traditional visual rating basis of evaluating breeding line resistance is a laborious and subjective process (Bauriegel & Herppich, 2014; Deery et al., 2014; Mahlein, 2016; Su et al., 2021). Spectroscopic methods can be superior to visual rating by providing rapid and unbiased assessment in early infection stages (Thomas et al., 2018; West et al., 2017), and their suitability to detect rust and powdery mildew has been investigated in cereal pathology studies (as in Franke & Menz, 2007; Huang et al., 2007; Kuckenberg et al., 2009; Cao et al., 2013; Huang et al., 2014; Feng et al., 2016; Zhang et al., 2017).

The primary focus of FHB spectroscopy studies has been post-harvest mycotoxin detection in extracted kernels under laboratory conditions. The research has been reviewed by Saccon et al. (2017) and Femenias et al. (2020), and findings on this topic continue to be published (Zhang et al., 2020a,b; Shen et al., 2022). Less attention, however, has been given to whole spikes (Bauriegel et al., 2011; Alisaac et al., 2018; Whetton et al., 2018a; Huang et al., 2019, 2020; Mahlein et al., 2019; Zhang et al., 2020c).

Field research has also been limited until recent years, as highlighted by Whetton et al. (2018b) and Huang et al. (2019). Whetton et al. (2018b) employed a proximal push-broom hyperspectral imager coupled with an artificial light source for FHB occurrence density estimation in winter wheat and barley. Söderström and Börjesson (2013) predicted the deoxynivalenol Fusarium mycotoxin in oats from data captured by three different sensors, including a multispectral satellite imager. Further research included the Liu et al. (2020b) proposal of a novel broad-band vegetation index to detect the disease from Sentinel-2 multispectral imagery. Xiao et al. (2020) augmented a time series of satellite observations with meteorological data for the same purpose, and Liu et al. (2020a) investigated the feasibility of FHB monitoring at the field scale using UAV equipped with hyperspectral camera.

Whetton et al. (2018b) considered that the small number of FHB field studies may stem from the difficult detection of the disease in a crop stand. Unlike rust and powdery mildew, FHB symptoms affect only the spikes, and these constitute a small fraction of the total biomass. More recently, Liu et al. (2020a) proposed the identification of infected field areas based on whole canopy characteristics, but the feasibility of this novel and potentially controversial approach requires further scrutiny.

There is a limited transfer ability of disease detection methods developed for large-scale farming applications to field phenotyping. Simultaneous screening of multiple genotypes constrains the size of experimental plots to a level below the spatial resolution attainable with many remote or even proximal sensing systems (Rebetzke et al., 2014; Barmeier & Schmidhalter, 2016). Artificial inoculation may be limited to a small subset of plants in each plot, thus further limiting the choice of spectral data acquisition techniques. The plot size reduction can also modify its spectral properties (Barmeier & Schmidhalter, 2016), and this invalidates models which assume a continuous canopy.

Moreover, disease severity must be estimated when searching for tolerant cultivars, rather than mere disease occurrence (Qiu et al., 2019). This task has only recently been attempted (Bock et al., 2010; Huang et al., 2019; Qiu et al., 2019; Zhang et al., 2019). Moreover, the predictions must consider masking the disease signal by genotype differences (Pinter et al., 1985). All these challenges create the need for research programs specifically aimed at field phenotyping.

This study therefore investigates the accuracy of winter wheat FHB infection severity assessments based on reflectance spectroscopy measurements in a field phenotyping setting. The specific aims are: (1) to estimate absolute accuracy of ordinal support vector machine (SVM) predictive models by comparing their outputs to visual ratings by an expert, (2) to examine the obtained accuracy according to the

crop phenological phase at which data were acquired and (3) to analyse the influence of spectra pre-processing on infection severity prediction quality. We hypothesised that the predictions at a late stage of the disease would be satisfactory for genotype screening. We expected inferior predictions with spectral data obtained at an earlier stage.

2. Material and methods

2.1. Site and experiment description

The experiment was conducted at the premises of the Crop Research Institute in Praha Ruzyně, Czech Republic, at 50.085° N, 14.300° E. Ruzyně lies at 340 m asl and is a predominantly sugar beet growing region. The soil is Orthic Luvisol with a silty clay loam texture to the depth of 0.7 m. The annual average air temperatures registered in the 1996–2020 period ranged between 7.6 °C and 11.0 °C, and the annual precipitation sums varied between 344.6 mm and 769.1 mm.

The study was based on a trial evaluating winter wheat cultivar resistance to FHB. The trial field was maintained in a crop rotation of cereals, pea and cereal-legume cover crops. The seedbed preparation comprised shallow tillage of the preceding crop, NPK fertilisation (45 kg ha^{-1} of N, P$_2$O$_5$ and K$_2$O) and ploughing. The soil was then levelled, which was followed by a pass of a trailed combination seedbed cultivator.

The winter wheat was sown by hand between the 1st and the 3rd of October 2019 in 12 parallel blocks 1-metre-wide. Each block comprised two rows of hill plots, facing each other. The grain in these plots was sown every 0.25 m in circular areas of 0.10–0.15 m diameter (Fig. 1). The hill−plot pairs differed depending on the sown cultivar, and these were arranged according to earliness in the first three blocks, and in alphabetic order in the remainder. The cultivars included the following; genotypes submitted to pre-registration trials, standard cultivars recommended for growing in Czech Republic conditions, breeding stock selection, breeding lines subjected to genotyping and genetic material evaluated under the European Fusarium ring test framework. The cultivar arrangements were in triplicate, with the same cultivar order used in each replication.

Plot maintenance after sowing comprised 0.5 l ha^{-1} Cougar Forte herbicide treatment (Bayer AG, Leverkusen, Germany) and 0.1 l ha^{-1} Markate 50 insecticide (Sumiagro, Praha, Czech Republic). These were applied on the 25th of October 2019, at the second-leaf stage of the plants (BBCH GS 12), which were again treated with pesticides on the 23rd of April 2020, during the flag leaf development (GS 37−39). These comprised 1.2 l ha^{-1} Agritox 50 SL (Nufarm UK Limited, Bradford, UK) and 0.4 l ha^{-1} Starane Forte (Dow AgroSciences s.r.o., Praha, Czech Republic) herbicides and 0.08 l ha^{-1} Nexide insecticide (Cheminova A/S, Harboør, Denmark). The paths separating the blocks were maintained free of vegetation with a mower
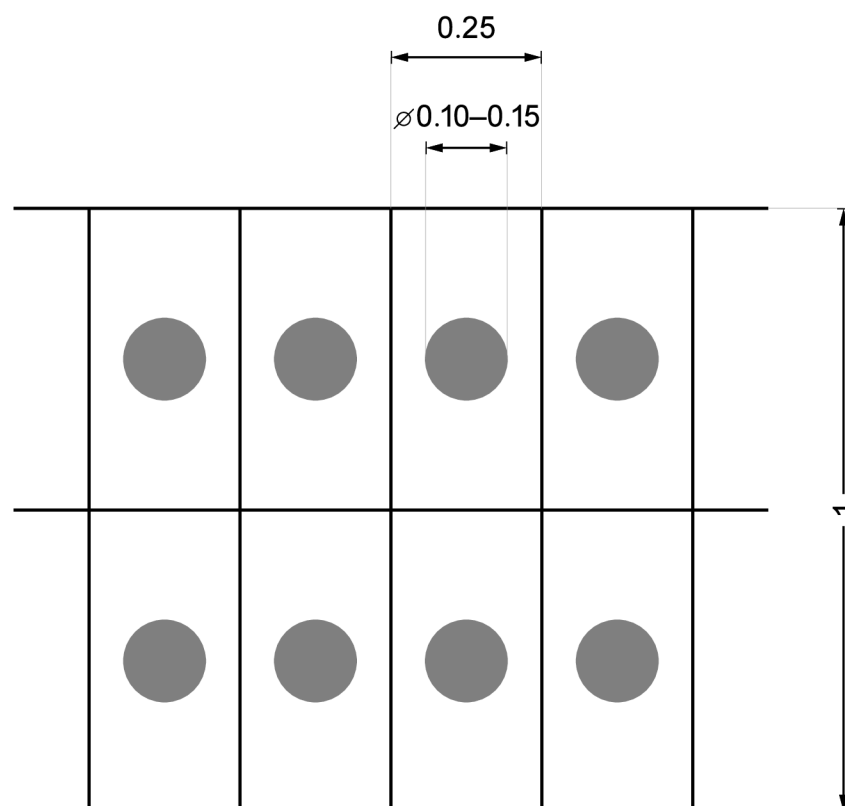

Fig. 1 − The arrangement of sowing areas within the hill plots. The dimensions are in metres.

Fig. 2 — A row of hill-plot pairs with tied and bagged infected spikes facing non-infected spikes.

blade and a harrow driven by compact utility tractor; with three passes made during the vegetation period.

One row in each block of the hill plots was inoculated with isolate B of *Fusarium culmorum* when they reached flowering (GS 65). This inoculation was by hand sprayer filled with 0.8×10^7 spores ml^{-1} suspension. Ten spikes were tied in a cluster, and the application performed from all sides. Excess suspension was intercepted using a folded aluminium foil screen held behind the clusters to minimise uncontrolled spreading of the inoculum by wind currents. The clusters were then kept for 24 h in 0.4 by 0.6 m polyethylene bags to create stable inoculation conditions (Fig. 2). This was followed by treatment in the blocks facing those already treated. Here, ten spikes were again tied and covered with a bag for comparison. This procedure differed to the Khaledi et al. (2017) study, because these were left dry, and not previously sprayed with water. All plots were maintained under irrigation in all weather conditions except during and after rain to foster infection development.

The study focused on hill plots inoculated on a single day and their facing plots in order to avoid the confounding effect of external conditions variation on the multiple inoculation and data collection dates. The available sample size was maximised by choosing the 3rd of June 2020 for plot inoculation. Anthesis was recorded in 176 hill plots on that day.

2.2. Proximal sensing campaigns

Spectral data were collected by ASD Fieldspec 4 Hi-Res spectroradiometer with contact probe (Malvern Panalytical, Malvern, UK). The device operation involved three detectors, and provided radiometric measurements with a 16-bit resolution. The spectral range was 350—2500 nm, with 3-nm resolution in the Vis—NIR and 8-nm in the SWIR region. The respective sampling intervals were 1.4 nm and 1.8 nm, with measurements re-sampled to 1 nm. The contact probe had a built-in halogen light source to provide measurements unaffected by ambient illumination when full contact with the sample was attained.

Two spectral data acquisition campaigns were planned: the first took place between 11.45 and 15.30 local time on the 17th of June 2020, 14 dai, soon after initial infection symptoms were visible to the naked eye. The second was performed between 09.15 and 12.40 on the 1st of July 2020, 28 dai, when the symptoms became severe. These dates coincided with the GS 77 milk- and 87 wax-ripening plant developmental phases.

Wiktor R. Żelazny, Jana Chrpová, and Pavel Hamouz: Fusarium head blight...

BIOSYSTEMS ENGINEERING 211 (2021) 97−113   **101**

Each spike cluster was subjected to a sequence of five measurements, during which three spikes were positioned between the probe and a black non-woven textile to eliminate accidental light penetration from the background. This approach was similar to the Ma et al. (2020) study, but adapted to contact measurements. The fabric had low reflectance of less than 6% throughout the measured spectrum, and as a result very limited effect on the measured spectral curves. The block order was randomised, but the measurements in each block were performed systematically to minimise the time between scans. Non-inoculated plots were measured in a batch first, in order to avoid their contamination with the pathogen. Two measurement runs were performed with the probe disinfected in-between, and this avoided running out of time before both hill plots in each pair were scanned. The device was calibrated using a white Spectralon when entering a new row, and then approximately every ten minutes. The reflectance factor of the Spectralon was >99% for visible light and >98% for NIR. The resultant dataset comprised complete spectral data for 67 hill−plot pairs out of the available pool of 176 pairs, thus equivalent to 1340 spectra. With two exceptions, no cultivar appeared more than once in the dataset. All included cultivars are listed in Supplement S1.

Collection of imaging data supplemented point spectrometry. The images were obtained from a tripod-mounted hyperspectral Rikola model Fabry−Pérot 2D frame camera (Senop, Oulu, Finland). The data cubes comprised 45 bands with spatial resolution of 1010×648 px, 12-bit radiometric resolution and covering 500−900 nm spectral range. The distance between the spikes and the camera lens was 0.6−0.7 m, resulting in GSD of approximately 0.37−0.43 mm px^{-1}. The integration time was set to 5 ms. Dark current noise and irradiance were measured and a Spectralon panel with 75% reflectance factor positioned next to the spikes for radiometric correction. The acquisition process was slow because of difficult aiming of the device's narrow field of view of 36.5°, the absence of a viewfinder and slow system response to positional adjustments. There was also additional time waiting for appropriate light conditions without cloud obstruction. In the end, these data were discarded because the experimental spike clusters were impossible to distinguish from the remaining spikes in most of the captured images. Therefore we do not recommend Rikola for field phenotyping applications when individual, pre-determined plant organs require precise capture with high spatial resolution. A specialised camera, such as that used by Behmann et al. (2018), may be more suitable for this purpose.

2.3. Reference data collection

A visual symptom score (VSS) was assigned to each hill plot of the infected rows 14, 21, 28 and 35 dai. This score was based on the following 9-point scale for percentage of infected spikelets in the spike cluster; 9 points: <5%, 8: 5−17%, 7: 18−30%, 6: 31−43%, 5: 44−56%, 4: 57−69%, 3: 70−82%, 2: 83−95% and 1: >95%. Note that the higher point values here denote lower severity level. Plots were examined linearly, not at random. The two assessments closest in time to spectral data

acquisitions were kept for further analysis. They were the 17th of June 2020, 14 dai, and the 1st of July 2020, 28 dai; matching exactly the dates of the spectral campaigns.

FHB symptoms were noted in some uninoculated spike bundles during the late proximal sensing campaign, and therefore the facing hill plots still in the GS 87 wax-ripening were also rated on the 2nd of July 2020, 29 dai. The hill plot pairs with facing-plot rating below 9 were excluded from further analysis, leaving the final number of 48 retained hill plot pairs. The amount of discarded data is thought to have been higher if the non-inoculated spikes had been sprayed with water prior to their bagging, and the deviation from the Khaledi et al. (2017) procedure is fortunate, all other things being equal.

2.4. Hill plot partitioning and spectra pre-processing

The hill plots were randomly split into training and validation data partitions in 3:1 proportion. The split was repeated 10 times in a jack-knife manner (Westad & Marini, 2015). The spectra from each acquisition campaign were then independently pre-processed. They were initially subjected to parabolic correction to smooth transitions between the three Fieldspec detectors (Hueni & Bialek, 2017) and the 'noisy' regions below 425 nm and above 2275 nm were visually identified and removed.

The median reflectance was derived for each combination of partitioning, pre-processing, hill plot and spectral band. The choice of the median function was motivated by lack of distributional assumptions and robustness towards outliers. This, however, could also result in spectra from uninfected parts of some spike bundles dominating the spectra affected by disease and causing a lack of infection signal in the median signature. The performance of mean values was therefore explored. The following seven pre-processing schemes were tested after the aggregation step: (1) no further pre-processing, which provided a 'raw spectrum', (2) subtraction of an aggregated non-inoculated counterpart spectrum, (3) division by an aggregated non-inoculated counterpart spectrum, (4) standard normal variate (SNV), (5) subtraction followed by SNV, (6) division followed by SNV and (7) maximum normalisation followed by the Savitzky−Golay derivative and smoothing to approximate the approach used in Whetton et al. (2018b).

2.5. Modelling

The VSS assessments were performed on a discrete interval scale (Bock et al., 2010), and the computational complexity of the task can be reduced by applying an ordinal predictive model. Here, $n − 1$ independent functions are sufficient to discriminate between n ordered classes (Behmann et al., 2014a). According to Thomas et al. (2018), SVM modelling is a suitable method of detecting early plant disease symptoms. Consequently, ordinal classification SVMs with the radial basis function kernel were employed to predict the VSSs; as in Behmann et al. (2014b). Separate models were fitted for each combination of the spectral acquisition campaign, reference

VSS dataset, plot partitioning, spectra aggregation function and the spectra pre-processing scheme. This provided a total of 560 model ensembles. The C and gamma tuning hyper-parameters were determined by Bayesian optimisation algorithm (Bischl et al., 2018), according to the leave-one-out cross-validation maximum accuracy criterion. The validation accuracies of each model were then estimated; allowing for mis-classification of zero, one and two class differences; similar to Behmann et al. (2014b).

The joint effects of the spectra acquisition campaign and ground-truth data collection timings, spectra pre-processing scenarios and magnitudes of error tolerances on the probability of a correct VSS rating were modeled using mixed-effect Bayesian generalised linear models (Bürkner, 2018) separately for median and mean aggregation. A Bernoulli generating distribution with a logit link was assumed to comply with the binary type of the dependent variable (Kruschke, 2015) — a correct or incorrect VSS prediction adjusted for error tolerance. The grouping variables in the model reflected the dataset structure, and these comprised the data partition identifier crossed with the hill plot pair identifier. The draws from the posterior distributions associated with individual levels of the independent variables and their selected contrasts were then visualised, and summarised using the medians and 95% credibility intervals in the visualisations.

2.6. Study pre-registration, data availability and computational reproducibility

This was a pre-registered study (Nosek et al., 2018), with the pre-registration form (Żelazny et al., 2020) hosted by the Center of Open Science. The dataset is available from a Zenodo repository (https://doi.org/10.5281/zenodo.4536881), excluding the hyperspectral data cubes because of their excessive size and the fact that they were not analysed.

The analysis was coded in the R language (R Core Team, 2020). The `prospectr` package (Stevens & Ramirez-Lopez, 2013) was employed for the spectra pre-processing, `e1071` (Meyer et al., 2020) for SVM classification and `mlr` (Bischl et al., 2016) for model tuning. The linear modelling was performed by `brms` (Bürkner, 2018) interface to Stan (Carpenter et al., 2017), and the draws from the posterior distributions were further processed using the `tidybayes` package (Kay, 2020). The dataset is accompanied by computational scripts and documentation for reproducing the analysis. A GNU Guix manifest file and a channel specification (Courtès & Wurmus, 2015) are also provided to recreate the original software analysis environment.

3. Results

3.1. Exploratory data analysis

FHB infection is readily discernible in the measured spectra, as depicted in the top panel of Fig. 3 for median aggregation. The lower VSSs denote higher severity, and these are associated with increased red reflectance and the lowered NIR shoulder next to the red edge region. The remainder of the NIR and the SWIR spectra have increased reflectance. There is a successive blue shift discernable as the ratings decrease towards the value of 4, and a qualitative change in the signature shape for the three lowest VSSs. The latter involves a further increase in red reflectance, to the extent that the red absorption feature almost disappears, and the curves assuming an almost linear shape in the visible and red-edge regions and beyond. The VSSs in the middle of the scale were assigned to the hill plots during both early and late reference data collection, thus enabling comparison over time. The wax-ripening later phenological phase is characterised by increased green and red reflectance relative to the milk-ripening reflectance pattern. Similar curves were obtained for spectra derived using mean aggregation (Supplement S2).

The bottom panels in Fig. 3 highlight that most pre-processing schemes retained the possibility of discerning at least the curves corresponding to the lowest or highest infection levels. However, the character and degree of the VSSs separation differ across scenarios. As expected, subtraction and division resulted in the top-rated spectra following the horizontal line centred at 0 and 1 value respectively, whereas high disease severities are associated with increasingly pronounced deviations. The timing effect was preserved. The SNV yielded three bundles of curves associated with all early-date, low late-date and high late-date ratings. Pre-pending it with subtraction amplified the curve separation in the red-edge region, while suppressing the SWIR signal and introducing noise. Pre-pending the SNV with division also gave 'noisy' signatures, and transformation adapted from Whetton et al. (2018b) resulted in a tight gradient. Under this scheme, plant health deterioration is associated with less pronounced features, except for the longest wavelengths, where the relationship is reversed. This pre-processing accentuated the effect of the infection on the left shoulder of the NIR plateau. All of these patterns occurred also after transforming mean-aggregated spectra (Supplement S2).

Figure 4 shows that the late-date VSSs tended to be lower than those obtained from the first visual assessment, thus reflecting infection spread. They also exhibit a slightly wider range. A small number of the test data-points extend beyond the model calibration domains. The affected partitions are 1, 3 and 9; the latter in late assessment, only.

3.2. SVM model performances

While 100% correct VSS ratings within two-point error tolerance were predicted for large fractions of the test data partitions, predictions were poor in all zero-tolerance scenarios (Fig. 5). However, relatively high performance is associated with the early-date predictions regardless of spectral data timing provided one-point VSS error is permitted. This does not hold for the late-date ratings, where the accuracies seldom exceeded 75%. The effects of the individual spectra pre-processing schemes are less clear than those of the timings. The patterns are inconsistent and no obviously advantageous or failing strategy can be identified.

Aggregated confusion matrices were derived to explore the possible determinants of accuracy differences at the data-point level. The confusion matrices corresponding to individual partitionings were extracted and subjected to mathematical addition of their element contents. Figure 6 depicts
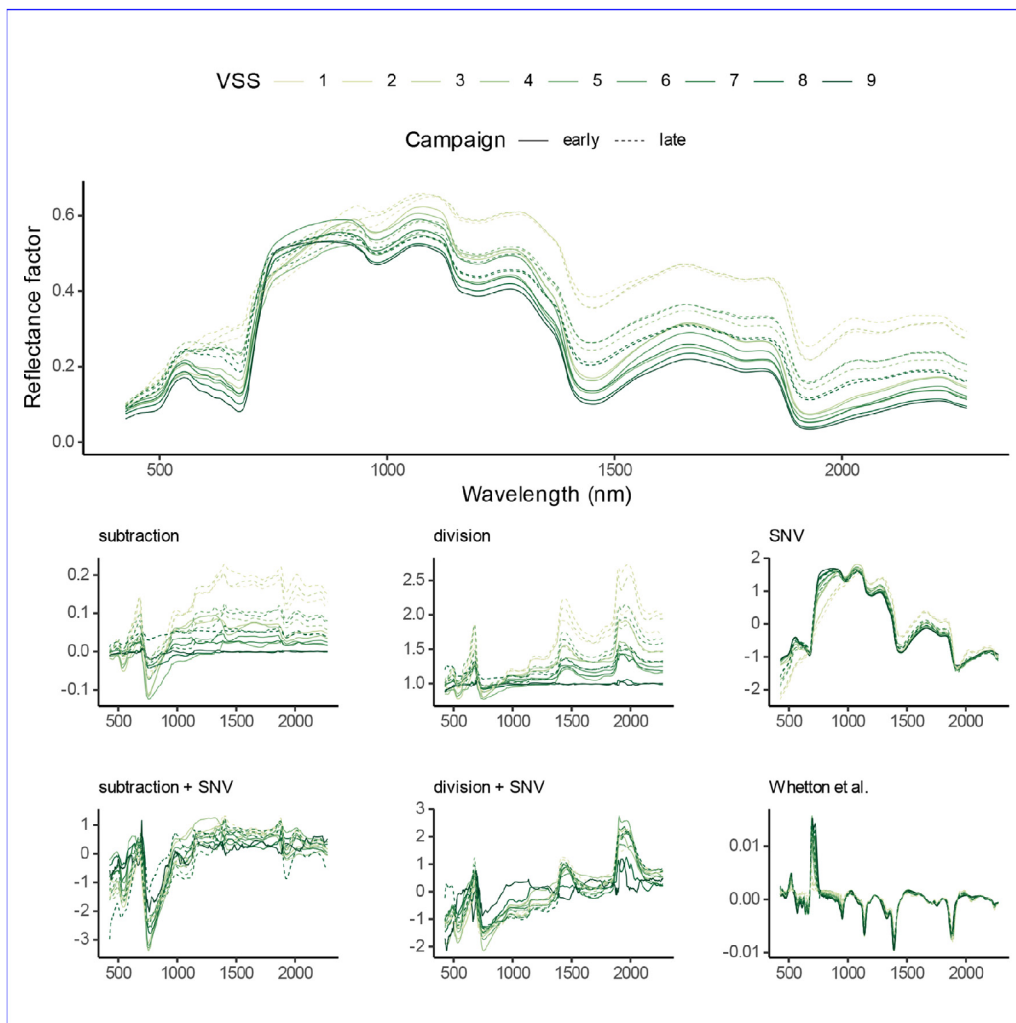
Fig. 3 – **Spectral signatures from the two measurement campaigns according to VSS, and the influence of the pre-processing strategies. A median of median-aggregated spectra is displayed for each rating. A lower rating denotes higher infection severity.**
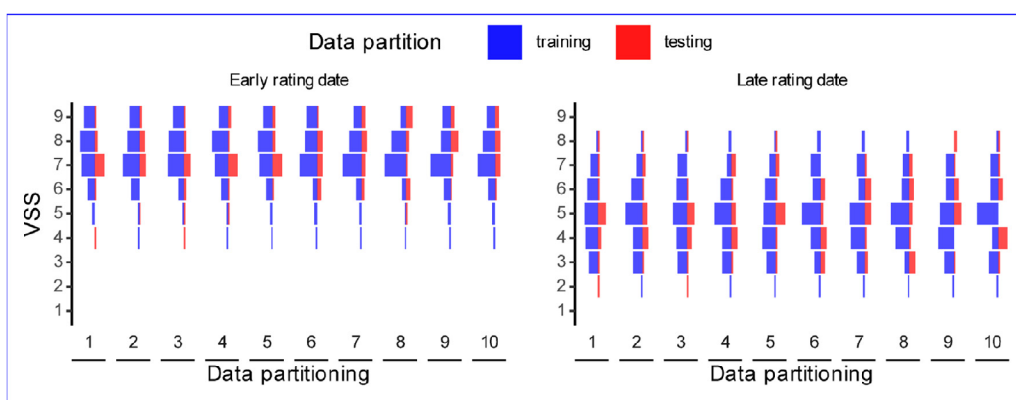

Fig. 4 – **The distributions of VSSs across the training and test data subsets according to the dataset partitioning and VSS rating timing. A lower rating denotes higher infection severity.**
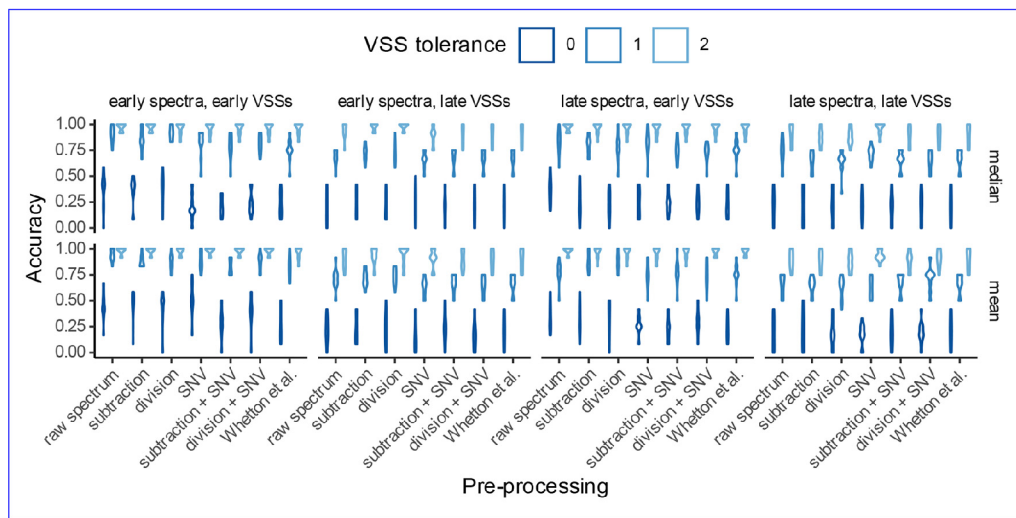
Fig. 5 – Distributions of SVM model accuracies across the dataset partitions according to the timing of spectral and ground-truth data collection, spectra pre-processing, spectra aggregation and allowed prediction error.

that the predictions appear to spread symmetrically from the matrix diagonals, thus suggesting that the observed accuracy variation should be attributed primarily to different model precisions, rather than biases (Bock et al., 2010). For certain pre-processing schemes the range of the early VSS predictions is compressed relative to those generated for the later date, despite only slight differences in the actual rating ranges (Fig. 4). This pattern may be explained by the early-date SVM models returning the same rating regardless of input spectra differences. However, among the affected models the single-prediction occurrence for spectra subjected to subtraction or division is 25%, and no model trained to the raw spectra behaves in this manner.

3.3. Linear modelling

Figure 7 depicts posterior draws derived for the linear model fitted to the SVM model prediction hits and misses. It presents generalised inferences on predictive model population properties, rather than summarising the performances of a particular sample of predictive models as seen in Fig. 5. The meaning of the y axis shifts from calculated prediction accuracy to potential probability of correct prediction. The patterns, however, are similar for the both figures, as expected. Credibility intervals corresponding to experimental treatment combinations can be derived from the draws. According to this criterion, performing no spectra transformation can result in satisfactory early-date ratings, even if the spectra are acquired at a later date. Moreover, SNV can be advantageous when the mean function is employed to aggregate the spectra and all data are colleted on the earlier date.

A fitted linear model enables effect size estimation for experimental treatment contrasts (Kruschke & Liddell, 2017). The prediction success probabilities corresponding to the different scanning dates were compared with the same VSS

rating dates maintained to resolve the hypothesis of deteriorated model quality when late-date spectral signatures are replaced with early-date spectra. The probability of a correct rating by a SVM model is virtually unaffected by the spectral campaign timing if an error of two VSS points is accepted. In contrast, there is positive effect of an earlier spectral campaign when subtraction or division pre-processing is applied after median aggregation to determine infection severity with single-point tolerance (Table 1). The classification success probability increase is below 30% when a predictive model is trained to the VSSs collected on the early date. This contrasts with a possible greater than 50% increase for the late-date ratings, characterised by lower baseline estimates (Figs. 5 and 7). Although the effect sizes are highly uncertain when no prediction error is allowed, positive early spectra influence can be inferred for the same transformations when early VSS ratings are considered. This pattern extends to additional pre-processing schemes when the mean function is employed for spectra aggregation. In particular, the prediction performance increases by more than 100% with SNV (Table 2).

The effects of pre-processing schemes on the SVM prediction quality were examined, with the raw-spectra scenarios forming the baseline for comparison. The prediction performance remains almost the same if an error of two VSS points is allowed (Tables 3 and 4). The response is weak for models trained to the early-date infection ratings when the tolerance is reduced to one point. The effect differs for subtraction and division depending on the timing and spectra aggregation function. Slight, but more consistent performance decreases are associated with median aggregation followed by subtraction and division combined with SNV, and also with the pre-processing adapted from Whetton et al. (2018b). The detrimental effects of these schemes are confirmed by the zero-error tolerance scenario, where the probabilities of a correct
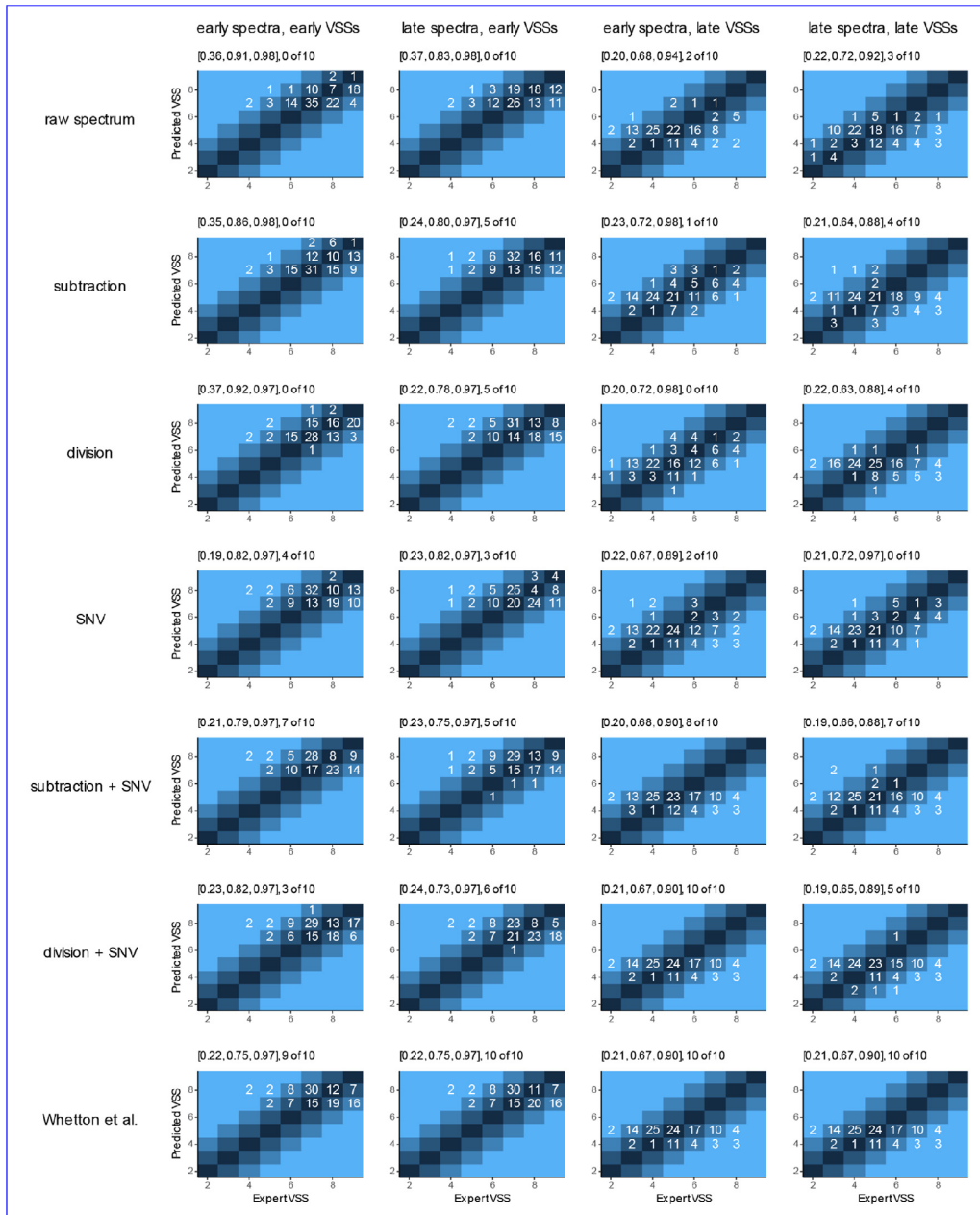
Fig. 6 − Aggregated confusion matrices according to spectral and ground-truth collection timing and spectra pre-processing scheme. The matrix labels explain mean accuracies corresponding to the prediction error tolerance levels (0−2 VSS points) and the partitioning count for which all model predictions had the same value. A lower rating denotes higher infection severity. Only SVM models trained to the spectra subjected to median aggregation were included in the analysis.

VSS assignment are reduced by (median estimates and 95% credibility intervals) 63% [32, 80], 56% [19, 76] and 58% [22, 78] when the milk-ripening phenological phase is involved. Both SNV and the remaining two transformations also exhibit this pattern, but the latter only for one timing combination. Mean spectra aggregation preserves the negative effects of subtraction followed by SNV and the scheme adapted from Whetton et al. (2018b). However, some evidence of accuracy increase with SNV (40% [−9, 125]) was obtained for early data collection scenario, as suggested by Fig. 7.
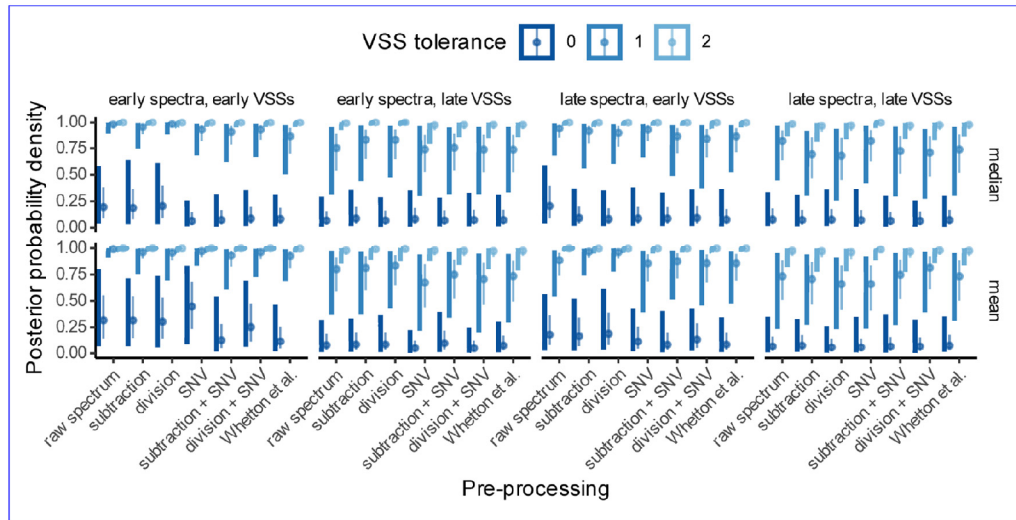
Fig. 7 — Posterior distributions, medians and 95% credibility intervals of SVM model correct VSS prediction mean probabilities according to the timing of spectral and ground-truth data collection, spectra pre-processing and allowed prediction error. Each panel corresponds to a separate linear model, with a distinct underlying spectra aggregation function.

Table 1 — Medians and 95% credibility intervals of the multiplicative effects on SVM model correct prediction mean probabilities according to median-aggregated spectra pre-processing and allowed prediction error, when a late spectral campaign is replaced by an early campaign and the VSS rating date remains fixed.

Contrast	Pre-processing	VSS tolerance		
		0	1	2
early spectra : late spectra | early VSSs	raw spectrum	0.97 [0.56, 1.66]	1.04 [1.00, 1.13]	1.00 [0.99, 1.01]
	subtraction	2.00 [1.10, 3.62]	1.04 [0.99, 1.16]	1.00 [0.99, 1.01]
	division	2.50 [1.41, 4.60]	1.09 [1.03, 1.26]	1.00 [0.99, 1.01]
	SNV	0.72 [0.37, 1.41]	1.00 [0.92, 1.08]	1.00 [0.99, 1.01]
	subtraction + SNV	0.82 [0.43, 1.59]	1.05 [0.96, 1.21]	1.00 [0.99, 1.01]
	division + SNV	0.94 [0.50, 1.78]	1.10 [1.01, 1.32]	1.00 [0.99, 1.01]
	Whetton et al.	1.07 [0.56, 2.13]	1.00 [0.88, 1.13]	1.00 [0.99, 1.01]
early spectra : late spectra | late VSSs	raw spectrum	0.87 [0.45, 1.71]	0.92 [0.73, 1.09]	1.01 [0.99, 1.04]
	subtraction	1.20 [0.63, 2.35]	1.18 [1.00, 1.58]	1.03 [1.01, 1.10]
	division	0.88 [0.44, 1.72]	1.21 [1.01, 1.64]	1.03 [1.01, 1.09]
	SNV	1.13 [0.59, 2.24]	0.91 [0.72, 1.07]	0.98 [0.93, 0.99]
	subtraction + SNV	1.07 [0.53, 2.15]	1.04 [0.84, 1.30]	1.01 [0.98, 1.05]
	division + SNV	1.15 [0.58, 2.24]	1.04 [0.83, 1.33]	1.00 [0.97, 1.04]
	Whetton et al.	1.00 [0.51, 1.92]	1.00 [0.80, 1.26]	1.00 [0.97, 1.03]

4.   Discussion

4.1.   Spectral patterns associated with VSS values, and the effect of pre-processing

The increase in visible-region reflectance when the spike infection becomes more severe (Fig. 3) indicates pigment breakdown and photosynthetic apparatus damage in the presence of fungi (West et al., 2003; Morin et al., 2017; Saccon et al., 2017). This agrees with the FHB 'bleaching' symptoms (McMullen et al., 2012; Bauriegel & Herppich, 2014; Ma et al., 2020). The observed red-edge displacement towards the shorter wavelengths is also well-described (Martinelli et al., 2015; West et al., 2017), as is the NIR reflectance decrease (Hamid Muhammed, 2005; Alisaac et al., 2018; Mahlein et al.,

2019; Zhang et al., 2020c,d). The qualitative change in spectral signature shape in the greatest disease severities suggests the onset of a new biological process, where the increase in red reflectance may indicate symptoms of fungus conidia formation, when pathogen tissue has a pink hue (Qiu et al., 2019; Mielniczuk & Skwaryło-Bednarz, 2020). However, Bauriegel et al. (2011) linked propagation to a different spectral pattern, which was not observed in this study, and an alternative explanation may be that the spikes become dry (Hamid Muhammed, 2005).

VSSs differentiation is apparent in raw and pre-processed spectra. Although distinguishing features in the schemes differ, the rating separation is preserved at least to some degree. This may explain the limited consistency of pre-processing effects across the examined timing combinations (Fig. 5). If both spectra acquisition and VSSs ratings are

Table 2 – Medians and 95% credibility intervals of the multiplicative effects on SVM model correct prediction mean probabilities according to mean-aggregated spectra pre-processing and allowed prediction error, when a late spectral campaign is replaced by an early campaign and the VSS rating date remains fixed.

Contrast	Pre-processing	VSS tolerance		
		0	1	2
early spectra : late spectra | early VSSs	raw spectrum	1.70 [1.06, 3.03]	1.12 [1.04, 1.32]	1.00 [1.00, 1.00]
	subtraction	1.87 [1.13, 3.31]	1.00 [0.95, 1.04]	1.00 [1.00, 1.01]
	division	1.57 [0.94, 2.66]	1.00 [0.94, 1.04]	1.00 [0.99, 1.01]
	SNV	3.84 [2.19, 7.23]	1.13 [1.04, 1.37]	1.00 [1.00, 1.01]
	subtraction + SNV	1.50 [0.80, 2.80]	1.06 [0.98, 1.22]	1.00 [1.00, 1.01]
	division + SNV	1.87 [1.07, 3.31]	1.12 [1.04, 1.36]	1.00 [1.00, 1.01]
	Whetton et al.	1.35 [0.72, 2.53]	1.08 [0.99, 1.27]	1.00 [0.99, 1.00]
early spectra : late spectra | late VSSs	raw spectrum	1.21 [0.60, 2.39]	1.08 [0.90, 1.38]	1.00 [0.97, 1.04]
	subtraction	1.14 [0.59, 2.24]	1.14 [0.96, 1.50]	1.02 [1.00, 1.07]
	division	1.48 [0.75, 3.06]	1.26 [1.05, 1.76]	1.03 [1.01, 1.09]
	SNV	0.85 [0.41, 1.78]	1.02 [0.79, 1.34]	0.99 [0.96, 1.01]
	subtraction + SNV	1.47 [0.75, 2.85]	1.00 [0.80, 1.24]	1.01 [0.97, 1.06]
	division + SNV	0.75 [0.36, 1.59]	0.87 [0.66, 1.04]	0.99 [0.95, 1.02]
	Whetton et al.	1.00 [0.51, 1.98]	1.00 [0.80, 1.24]	1.00 [0.97, 1.04]

Table 3 – Medians and 95% credibility intervals of the multiplicative effects of the median-aggregated spectra pre-processing schemes on SVM model correct prediction mean probabilities according to the spectral campaign and VSS assessment timing and allowed prediction error.

Timing	Pre-processing	VSS tolerance		
		0	1	2
early spectra, early VSSs	subtraction	0.95 [0.54, 1.62]	0.98 [0.92, 1.01]	1.00 [0.99, 1.01]
	division	1.05 [0.61, 1.76]	1.00 [0.98, 1.03]	1.00 [0.99, 1.01]
	SNV	0.33 [0.17, 0.60]	0.95 [0.86, 0.99]	1.00 [0.99, 1.01]
	subtraction + SNV	0.37 [0.20, 0.68]	0.93 [0.82, 0.98]	1.00 [0.99, 1.00]
	division + SNV	0.44 [0.24, 0.81]	0.95 [0.86, 0.99]	1.00 [0.99, 1.01]
	Whetton et al.	0.42 [0.22, 0.78]	0.88 [0.74, 0.96]	1.00 [0.99, 1.01]
early spectra, late VSSs	subtraction	1.29 [0.67, 2.57]	1.10 [0.93, 1.38]	1.01 [1.00, 1.03]
	division	1.01 [0.51, 1.97]	1.10 [0.94, 1.38]	1.01 [1.00, 1.03]
	SNV	1.21 [0.62, 2.37]	0.98 [0.79, 1.21]	0.98 [0.94, 1.00]
	subtraction + SNV	1.02 [0.51, 2.01]	1.00 [0.81, 1.23]	0.99 [0.95, 1.00]
	division + SNV	1.07 [0.56, 2.04]	0.98 [0.80, 1.20]	0.99 [0.95, 1.00]
	Whetton et al.	1.07 [0.56, 2.07]	0.98 [0.78, 1.20]	0.99 [0.95, 1.00]
late spectra, early VSSs	subtraction	0.45 [0.25, 0.85]	0.97 [0.88, 1.04]	1.00 [0.99, 1.01]
	division	0.40 [0.21, 0.73]	0.96 [0.85, 1.02]	1.00 [0.99, 1.01]
	SNV	0.43 [0.23, 0.80]	0.99 [0.91, 1.05]	1.00 [0.99, 1.01]
	subtraction + SNV	0.43 [0.23, 0.78]	0.92 [0.78, 0.99]	1.00 [0.99, 1.01]
	division + SNV	0.46 [0.25, 0.83]	0.90 [0.74, 0.98]	1.00 [0.99, 1.01]
	Whetton et al.	0.38 [0.20, 0.70]	0.92 [0.79, 1.00]	1.00 [0.99, 1.01]
late spectra, late VSSs	subtraction	0.93 [0.46, 1.85]	0.85 [0.65, 1.02]	0.98 [0.93, 1.01]
	division	0.99 [0.51, 1.89]	0.84 [0.61, 1.01]	0.99 [0.94, 1.01]
	SNV	0.93 [0.47, 1.79]	1.00 [0.84, 1.18]	1.01 [1.00, 1.05]
	subtraction + SNV	0.81 [0.42, 1.61]	0.89 [0.69, 1.05]	0.99 [0.95, 1.01]
	division + SNV	0.82 [0.40, 1.65]	0.87 [0.67, 1.03]	0.99 [0.95, 1.02]
	Whetton et al.	0.93 [0.48, 1.83]	0.91 [0.71, 1.06]	0.99 [0.96, 1.02]

performed at the milk-ripening stage, SNV can be recommended, but only if the spectra are aggregated using the mean function (Table 4). This transformation can reduce noise resulting from light beam scattering and it was employed in some studies on Fusarium-affected kernels (Femenias et al., 2020). With median aggregation no pre-processing scheme appears superior to raw spectra for SVM model training (Table 3). There is some advantage in subtraction or division by an uninfected spike spectrum when one point error-difference is accepted, but this is only in certain combinations of data acquisition timings and spectra aggregation schemes. The low

probability of these scenarios and the modest gain in prediction performance do not justify doubling the spectral campaign effort required to measure the healthy-spike spectra.

The negative effect of following these two transformations with SNV is likely due to the 'noisy' spectra resulting from these combinations (Fig. 3). A similar effect is obtained for the pre-processing scheme adapted from Whetton et al. (2018b), which yielded a tight VSS gradient. Although the ratings are consistently ordered, the signatures cross and overlap because of order-inversion along the

Table 4 − Medians and 95% credibility intervals of the multiplicative effects of the mean-aggregated spectra pre-processing schemes on SVM model correct prediction mean probabilities according to the spectral campaign and VSS assessment timing and allowed prediction error.

Timing	Pre-processing	VSS tolerance		
		0	1	2
early spectra, early VSSs	subtraction	1.00 [0.63, 1.60]	0.97 [0.91, 1.00]	1.00 [1.00, 1.00]
	division	0.96 [0.59, 1.54]	0.97 [0.90, 0.99]	1.00 [0.99, 1.00]
	SNV	1.40 [0.91, 2.25]	0.98 [0.93, 1.00]	1.00 [1.00, 1.00]
	subtraction + SNV	0.41 [0.23, 0.70]	0.94 [0.84, 0.98]	1.00 [1.00, 1.00]
	division + SNV	0.81 [0.49, 1.33]	0.97 [0.91, 1.00]	1.00 [1.00, 1.00]
	Whetton et al.	0.37 [0.20, 0.64]	0.93 [0.83, 0.98]	1.00 [0.99, 1.00]
early spectra, late VSSs	subtraction	1.07 [0.56, 2.07]	1.01 [0.86, 1.22]	1.01 [0.99, 1.05]
	division	1.06 [0.55, 2.07]	1.04 [0.90, 1.26]	1.01 [1.00, 1.05]
	SNV	0.62 [0.31, 1.23]	0.85 [0.63, 1.02]	1.00 [0.97, 1.03]
	subtraction + SNV	1.20 [0.63, 2.32]	0.94 [0.75, 1.12]	1.00 [0.96, 1.03]
	division + SNV	0.62 [0.31, 1.26]	0.89 [0.68, 1.07]	0.99 [0.95, 1.02]
	Whetton et al.	0.94 [0.47, 1.87]	0.92 [0.72, 1.10]	1.00 [0.96, 1.03]
late spectra, early VSSs	subtraction	0.92 [0.52, 1.58]	1.09 [1.02, 1.27]	1.00 [0.99, 1.00]
	division	1.05 [0.61, 1.81]	1.08 [1.02, 1.25]	1.00 [0.99, 1.00]
	SNV	0.63 [0.35, 1.13]	0.97 [0.84, 1.08]	1.00 [0.99, 1.00]
	subtraction + SNV	0.47 [0.25, 0.85]	0.99 [0.87, 1.11]	1.00 [0.99, 1.00]
	division + SNV	0.74 [0.43, 1.33]	0.97 [0.84, 1.08]	1.00 [0.99, 1.00]
	Whetton et al.	0.47 [0.25, 0.85]	0.97 [0.84, 1.08]	1.00 [0.99, 1.00]
late spectra, late VSSs	subtraction	1.13 [0.58, 2.28]	0.96 [0.75, 1.20]	1.00 [0.96, 1.03]
	division	0.87 [0.42, 1.79]	0.91 [0.67, 1.14]	0.99 [0.94, 1.02]
	SNV	0.87 [0.42, 1.77]	0.90 [0.67, 1.13]	1.01 [0.99, 1.05]
	subtraction + SNV	0.99 [0.48, 2.03]	1.02 [0.82, 1.26]	0.99 [0.95, 1.03]
	division + SNV	1.00 [0.49, 2.05]	1.10 [0.93, 1.41]	1.00 [0.97, 1.04]
	Whetton et al.	1.13 [0.57, 2.28]	1.00 [0.80, 1.26]	1.00 [0.97, 1.04]

wavelength dimension. This effect and the gradient steepness may have weakened the disease severity signal, and this could also explain the lack of SVM model sensitivity to variations in input spectra (Fig. 6).

4.2. SVM model accuracies, and their relationship to field phenotyping requirements

On most occasions, the models failed to predict the test data VSSs when no error was allowed (Figs. 5 and 7). This unsatisfactory performance can be linked to the high number of rating levels. The accuracy of the early visual assessment was acceptable only with one point permitted error, and not for models trained to late expert ratings. This result does not confirm the hypothesis for accurate late-stage disease severity estimation with reflectance spectra and SVM.

One VSS point difference in this study indicates disease symptom difference of approximately ten percentage points. Although this error allowance needed to obtain satisfactory predictions is much greater than errors reported in other FHB proximal sensing experiments, some authors made no attempt to discriminate between disease severity levels. These researchers performed binary classification experiments, where high success rates are expected with both hyperspectral and multispectral images (Dammer et al., 2011; Hamid Muhammed, 2005; Jin et al., 2018). Other work collected data on infected spikelet counts within spikes, but transformed the calculated disease severities to binary scale prior to model training, and reported around 90% prediction accuracy (Ma et al., 2020). Liu et al. (2020a) performed similar dichotomy in their UAV study.

Moreover, the test dataset employed by these authors was not independent of training and validation datasets, and this may explain the 90−98% accuracy obtained with their six multivariate models.

Combining deep learning and feature selection was studied extensively by Zhang et al. (2020d). While these authors reported close to 100% accuracy using R-squared statistics, they used a relatively simple threshold algorithm rather than employing a ground-truthing expert. An earlier publication by Zhang et al. (2019) records a six-point ground-truth scale, which is a typical scoring resolution employed by human rating (Bock et al., 2010). The reported accuracies of the deep learning models are lower there, but still high enough to justify replicating the experiment and testing the proposed approach in a phenotyping setting.

A rare field phenotyping study of three cultivars by Qiu et al. (2019) used a region-growing algorithm to estimate individual spike surface area percentages with visible FHB symptoms, a process which closely matches the expert rating procedure. The very high spatial resolution of the analysed imagery compensated for the low spectral resolution, an inverse relationship as in our work. The authors mention the 'optimum performance' of their model but provide no numerical measures to enable comparison with other studies. This was followed by Su et al. (2021) proposal of a refined approach for application to 55 wheat genotypes. This research was based on individual spike segmentation, and the post-binary accuracy was also near 100%. Although the authors reported a negative prediction bias when disease severity was quantified on a 14-point scale, some performance measures exceeded 70%. This suggests that RGB imagery combined with

Wiktor R. Żelazny, Jana Chrpová, and Pavel Hamouz: Fusarium head blight...

BIOSYSTEMS ENGINEERING 211 (2021) 97−113

109

advanced processing is sufficient to phenotype FHB resistance in the field; at least when the disease is visible to the naked eye. However, it is important to note that the ground truth in this experiment was obtained by manually labelling the images, rather than based on field-rating by an expert, who would examine all sides of the spikes.

The low performance of the SVM models was attributed to low prediction precision rather than bias. Although precision can be improved by increasing the size of the training dataset, the high number of expert ratings required negates the advantage of spectroscopy over the traditional approach. The workload here could be reduced by expending extensive effort only once and building a spectral signature library for multiple trials. The additional collected ratings could then be limited for validation.

Research by Huang et al. (2019) attempted to generalise FHB severity predictions to various acquisition dates and localities. The authors divided the reflectance of a healthy-spike spectrum in a newly located site by those acquired at the reference site, and then applied the obtained ratios to all target-site measurements. This approach supports our pre-processing division, which had a slightly positive effect on SVM model performance for some spectral aggregation and timing combinations (Table 3).

The predictions might also be improved by testing other classifications model families in addition to SVM. For example, Huang et al. (2019) compared the performance of SVM with both Fisher's linear discriminant analysis outcomes and combined SVM-and-Fisher effect, and found the resultant combination superior to the individual models. An alternate option to improve the model performance would be feature selection (Jin et al., 2018; West et al., 2017; Zhang et al., 2020b).

Although the VSSs were provided by an expert with 20 years' experience in FHB field phenotyping, incorrect labelling of some spike clusters cannot be discounted (Bock et al., 2010; Dammer et al., 2011; Whetton et al., 2018b). Better predictions may also have been expected if the actual infection severity values had been known. Multiple VSS ground-truthing by several technicians, combined with a photo interpretation approach, would have been essential to mitigate this study weakness (Bock et al., 2010; Whetton et al., 2018b), but the additional human and computational workload did not justify this inclusion considering the early stages of FHB field sensing research.

The prediction quality improved dramatically when some rating error was allowed. The accuracy approached 100% when two VSS-levels difference was permitted, and this was sufficient to reliably distinguish between low, medium and heavily-infected cultivars. This result is similar to the four-level classification tested by Huang et al. (2019), where the authors reported their greatest accuracies in the 85−90% range. Despite its coarseness, this differentiation is sufficient for certain field phenotyping tasks; including the preliminary screening of disease-intolerant breeding lines (Bock et al., 2010).

4.3. Significance for commercial large-scale farming, and the timing effects

The dates corresponding to the VSSs for SVM training always agree with those of the VSSs test subsets. Herein, we focused on phenotyping, where plants are intentionally exposed to stress without management adjustment (Mahlein, 2016). In contrast, commercial farming relies on continuous plant status monitoring, and it is in the farmer's interest to predict the infection risk and its severity in advance (Liu et al., 2020b). A question arises whether an SVM model trained to an early spectral and training VSS dataset can be employed for predicting late-date VSSs. Figure 8 shows that the bulk of the infected hill plots had 2 to 3 points rating deterioration between the two data collection campaigns. This is surprising consistency, because the presence of multiple cultivars could be expected to have a higher variety of infection responses. This pattern should therefore enable the spectral campaign at the milk-ripening phase to provide data for satisfactory prediction of infection severity at the wax-ripening phase.

Our study does not corroborate the hypothesis of model performance deterioration associated with spectral data collection early in wheat spike development. Evidence of the contrary was found especially for pre-processing schemes involving a subtraction or division by a signature of healthy spikes (Tables 1 and 2). The early date corresponded to the milk-ripening stage, which Whetton et al. (2018a) identified as being optimal for Fusarium infection severity prediction using partial least squares regression in laboratory. However, the diseased spikes have similar colour to spikes unaffected by FHB as the crop matures, as depicted in temporal difference between spectral curves in Fig. 3. This ripening adversely affects spike detection in images (Fernandez-Gallego et al., 2018), and is detrimental to disease severity estimation accuracy (Bauriegel et al., 2011; Dammer et al., 2011; Qiu et al., 2019).

FHB detection at the milk-ripening developmental phase is still not early enough for fungicide treatment, which typically occurs not later than during flowering (McMullen et al., 2012). Preliminary evidence favouring disease prediction feasibility with late-anthesis data was obtained in an indoor experiment (Zhang et al., 2020c) and recent field work, by Su et al. (2021), included late anthesis as one of the developmental phases covered by the imaging campaigns. The latter authors,
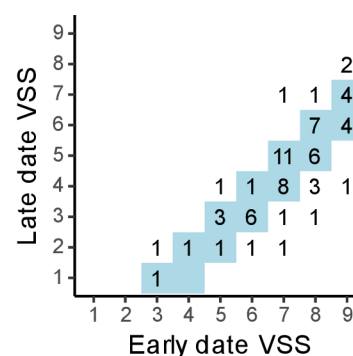

Fig. 8 − Infected hill plots counts according to their assigned early- and late-date VSSs. The highlighted background area corresponds to 2-or-3 point rating difference. A lower rating denotes higher infection severity.

however, did not examine the influence of timing factor in their experiment. Bauriegel et al. (2011) considered that successful diagnosis cannot be expected when the hyperspectral FHB assessment is performed at early flowering, and fluorescence imaging may be required to obtain the diagnosis at this phase (reviewed in Bauriegel & Herppich, 2014). However, fluorescence diagnosis has greater requirements to external conditions during data acquisition (Kuckenberg et al., 2009; Bauriegel & Herppich, 2014; Mahlein, 2016). West et al. (2017) then added that only infected spikes can be detected with this method.

Substantial mycotoxin production occurs in early infection, and this contributes to Fusarium aggression (Khaledi et al., 2017). The rate of pathogen spike penetration in high (Kang & Buchenauer, 2000), which narrows the fungicide treatment time window. Consequently, it is insufficient to just detect initiated infection when scheduling treatment or planning variable fungicide application rate. Instead, there must be identification of fields and field zones affected by other stress factors and susceptible to developing FHB in given weather conditions (McMullen et al., 2012). Symptom severity information obtained after infection took place enables efficacy assessment of prior fungicide intervention (Qiu et al., 2019). Moreover, areas for separate harvest of the mycotoxin-contaminated grain can be identified (Dammer et al., 2011; Bauriegel & Herppich, 2014; West et al., 2017), and contamination-level assays can be then limited to those latter grains (Bauriegel et al., 2011; Dammer et al., 2011; Saccon et al., 2017; but see Shen et al., 2022).

4.4. Deviations from the pre-registration

Nosek et al. (2018) contend that it is rare for a study to follow an unadjusted pre-registration protocol for data collection and analysis. They recommend a discussion of deviations and their consequences to preserve study transparency and demonstrate the validity of results.

Herein, several changes to the methods declared in the study pre-registration form (Żelazny et al., 2020) were necessary: Hyperspectral data cubes were collected, but it was not possible to analyse them so the hypothesis for comparison of the hyperspectral camera and point spectrometer could not be addressed. Shading of the area measured by spectroradiometer was abandoned for organisation reasons. However, shadows cast by the operator and contact probe were arguably sufficient to block direct illumination. Operator fatigue from equipment load and frequent bending to reach the spikes was the primary factor limiting the number of collected spectra, which was not envisaged as a possible stoping criterion in pre-registration. In addition, measurements from several plots could not be included because of undetected errors during spike-bundling, which would have provided incomplete spectral data. The declared order of the scans was also slightly adjusted to avoid acquiring spectral signatures of infected spikes without matching healthy-spike spectra. The development of disease symptoms on non-inoculated spikes was not anticipated, and the rationale for discarding the affected plot pairs was to avoid introducing artefacts during spectra subtraction and division. The order of spectra aggregation and pre-processing was then exchanged because paired signatures were required for the subtraction and division operations, and any pairing involving original spectra would have been arbitrary. Performance of the predictive model based on aggregated accuracy values could not be compared, because the 0% and 100% values in the dataset precluded logit-link application. This was solved by adding a further grouping level to the models, and treating the individual VSS prediction hits and misses as data points. This revised approach better reflected the data generation process, and potentially contributed to linear model validity.

There were also two extensions to the original plan: The use mean for spectra aggregation was considered prior to study pre-registration, but it was decided to employ the median. That decision was reconsidered during the spectra acquisition based on the patterns seen in the data. Although both approaches appear in the study, the mean spectra analysis results should be considered exploratory. Finally, the Fig. 6 confusion matrices and Fig. 8 contingency table were computed post-hoc to facilitate result discussion.

5. Conclusions

The patterns in the acquired spectral signatures can be explained by plant stress biology and linked to FHB visual symptoms. These agree with the spectral patterns obtained in other FHB studies, and this supports the suitability of a spectroradiometer with a contact probe for FHB phenotyping applications.

Ordinal SVM models applied to the raw spectra yielded predictions with confident distinction of low, moderate and high disease severity at the milk-ripening developmental phase. The proposed approach enables screening unpromising breeding lines, but other predictive models require testing for applications that demand finer FHB severity ratings. This is especially true for models trained to spectra collected at the wax-ripening developmental stage due to spike-drying discolour masking the disease signal. While deep learning may be suitable, the current published results require replication with multiple cultivar datasets and eliminated oversimplification of ground-truth data.

Predictions can in some situations be improved by subtracting healthy-spike reflectance measurements from the analysed signatures or by division. However, gains from these pre-processing schemes must be compared to the costs of additional scans. Further investigation is also required to determine if a predictive trained model can be applied year-to-year to reduce this effort. SNV pre-processing can be beneficial for milk-ripening phase predictions based on mean-aggregated spectra acquired at the same crop developmental stage.

The disease severity scores provided by ordinal SVM can help commercial farmers identify field zones for separate grain harvesting and evaluate the effects of fungicide applied in a given year using these models. However, they are unsuitable for scheduling protective measures because of Fusarium spp. specific mode of infection. This applies equally to all approaches that require spectral signatures from already diseased plants.

111

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biosystemseng.2021.08.019.

REFERENCES

Alisaac, E., Behmann, J., Kuska, M., Dehne, H.-W., & Mahlein, A.-K. (2018). Hyperspectral quantification of wheat resistance to *Fusarium* head blight: Comparison of two *Fusarium* species. *European Journal of Plant Pathology, 152*, 869—884.

Barmeier, G., & Schmidhalter, U. (2016). High-throughput phenotyping of wheat and barley plants grown in single or few rows in small plots using active and passive spectral proximal sensing. *Sensors, 16*(11), 1860.

Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U., & Herppich, W. (2011). Early detection of *Fusarium* infection in wheat using hyper-spectral imaging. *Computers and Electronics in Agriculture, 75*(2), 304—312.

Bauriegel, E., & Herppich, W. B. (2014). Hyperspectral and chlorophyll fluorescence imaging for early detection of plant diseases, with special reference to *Fusarium* spec. infections on wheat. *Agriculture, 4*(1), 32—57.

Behmann, J., Acebron, K., Emin, D., Bennertz, S., Matsubara, S., Thomas, S., Bohnenkamp, D., Kuska, M. T., Jussila, J., Salo, H., Mahlein, A.-K., & Rascher, U. (2018). Specim IQ: Evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection. *Sensors, 18*(2), 441.

Behmann, J., Schmitter, P., Steinrücken, J., & Plümer, L. (2014a). Ordinal classification for efficient plant stress prediction in hyperspectral data. *ISPRS - International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XL*(7), 29—36.

Behmann, J., Steinrücken, J., & Plümer, L. (2014b). Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing, 93*, 98—111.

Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., Casalicchio, G., & Jones, Z. M. (2016). mlr: Machine learning in R. *Journal of Machine Learning Research, 17*(170), 1—5.

Bischl, B., Richter, J., Bossek, J., Horn, D., Thomas, J., & Lang, M. (2018). *mlrMBO: A modular framework for model-based optimization of expensive black-box functions*. URL http://arxiv.org/abs/1703.03373.

Bock, C., Poole, G., Parker, P., & Gottwald, T. (2010). Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Critical Reviews in Plant Sciences, 29*(2), 59—107.

Buerstmayr, M., Steiner, B., & Buerstmayr, H. (2020). Breeding for Fusarium head blight resistance in wheat—Progress and challenges. *Plant Breeding, 139*(3), 429—454.

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), 395—411.

Cambaza, E., Koseki, S., & Kawamura, S. (2019). Why RGB imaging should be used to analyze *Fusarium Graminearum* growth and estimate deoxynivalenol contamination. *Methods and Protocols, 2*(1), 25.

Cao, X., Luo, Y., Zhou, Y., Duan, X., & Cheng, D. (2013). Detection of powdery mildew in two winter wheat cultivars using canopy hyperspectral reflectance. *Crop Protection, 45*, 124—131.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*(1), 1—32.

Courtès, L., & Wurmus, R. (2015). Reproducible and user-controlled software environments in HPC with Guix. In S. Hunold, A. Costan, D. Giménez, I. Alexandru, L. Ricci, M. E. Gómez Requena, V. Scarano, A. L. Verbanescu, S. L. Scott, S. Lankes, J. Weidendorfer, & M. Alexander (Eds.), *Euro-par 2015: Parallel processing workshops* (pp. 579—591). Vienna, Austria: Vienna University of Technology.

Dammer, K.-H., Möller, B., Rodemann, B., & Heppner, D. (2011). Detection of head blight (*Fusarium* ssp.) in winter wheat by color and multispectral image analyses. *Crop Protection, 30*(4), 420—428.

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., & Furbank, R. (2014). Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy, 4*(3), 349—379.

Femenias, A., Gatius, F., Ramos, A. J., Sanchis, V., & Marín, S. (2020). Use of hyperspectral imaging as a tool for Fusarium and deoxynivalenol risk management in cereals: A review. *Food Control, 108*, 106819.

Feng, W., Shen, W., He, L., Duan, J., Guo, B., Li, Y., Wang, C., & Guo, T. (2016). Improved remote sensing detection of wheat powdery mildew using dual-green vegetation indices. *Precision Agriculture, 17*(5), 608—627.

Fernandez-Gallego, J. A., Kefauver, S. C., Gutiérrez, N. A., Nieto-Taladriz, M. T., & Araus, J. L. (2018). Wheat ear counting in-field conditions: High throughput and low-cost approach using RGB images. *Plant Methods, 14*, 22.

Franke, J., & Menz, G. (2007). Multi-temporal wheat disease detection by multi-spectral remote sensing. *Precision Agriculture, 8*, 161—172.

Gilbert, J., & Haber, S. (2013). Overview of some recent research developments in fusarium head blight of wheat. *Canadian Journal of Plant Pathology, 35*(2), 149—174.

Hamid Muhammed, H. (2005). Hyperspectral crop reflectance data for characterising and estimating fungal disease severity in wheat. *Biosystems Engineering, 91*(1), 9—20.

Huang, W., Guan, Q., Luo, J., Zhang, J., Zhao, J., Liang, D., Huang, L., & Zhang, D. (2014). New optimized spectral indices for identifying and monitoring winter wheat diseases. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*(6), 2516—2524.

Huang, W., Lamb, D. W., Niu, Z., Zhang, Y., Liu, L., & Wang, J. (2007). Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Precision Agriculture, 8*, 187—197.

Huang, L., Li, T., Ding, C., Zhao, J., Zhang, D., & Yang, G. (2020). Diagnosis of the severity of *Fusarium* head blight of wheat ears on the basis of image and spectral feature fusion. *Sensors, 20*(10), 2887.

Huang, L., Wu, Z., Huang, W., Du, X., & Zhao, J. (2019). Identification of Fusarium head blight in winter wheat ears based on Fisher's linear discriminant analysis and a support vector machine. *Applied Sciences, 9*(18), 3894.

Hueni, A., & Bialek, A. (2017). Cause, effect, and correction of field spectroradiometer interchannel radiometric steps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10*(4), 1542–1551.

Jaillais, B., Roumet, P., Pinson-Gadais, L., & Bertrand, D. (2015). Detection of Fusarium head blight contamination in wheat kernels by multivariate imaging. *Food Control, 54*, 250–258.

Jin, X., Jie, L., Wang, S., Qi, H. J., & Li, S. W. (2018). Classifying wheat hyperspectral pixels of healthy heads and *Fusarium* head blight disease using a deep neural network in the wild field. *Remote Sensing, 10*(3), 395.

Kang, Z., & Buchenauer, H. (2000). Cytology and ultrastructure of the infection of wheat spikes by Fusarium culmorum. *Mycological Research, 104*(9), 1083–1093.

Kay, M. (2020). *tidybayes: Tidy data and geoms for bayesian models.* URL http://mjskay.github.io/tidybayes/.

Khaledi, N., Taheri, P., & Rastegar, M. F. (2017). Identification, virulence factors characterization, pathogenicity and aggressiveness analysis of *Fusarium* spp., causing wheat head blight in Iran. *European Journal of Plant Pathology, 147*, 897–918.

Kruschke, J. K. (2015). *Doing bayesian data analysis : A tutorial with R, JAGS, and stan* (pp. 621–648). London–San Diego: Academic Press. Waltham–Kidlington, Ch. Dichotomous Predicted Variable.

Kruschke, J. K., & Liddell, T. M. (2017). The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review*, 1–29.

Kuckenberg, J., Tartachnyk, I., & Noga, G. (2009). Temporal and spatial changes of chlorophyll fluorescence as a basis for early and precise detection of leaf rust and powdery mildew infections in wheat leaves. *Precision Agriculture, 10*, 34–44.

Liu, L., Dong, Y., Huang, W., Du, X., & Ma, H. (2020a). Monitoring wheat Fusarium head blight using unmanned aerial vehicle hyperspectral imagery. *Remote Sensing, 12*(22), 3811.

Liu, L., Dong, Y., Huang, W., Du, X., Ren, B., Huang, L., Zheng, Q., & Ma, H. (2020b). A disease index for efficiently detecting wheat *Fusarium* head blight using Sentinel-2 multispectral imagery. *IEEE Access, 8*, 52181–52191.

Ma, H., Huang, W., Jing, Y., Pignatti, S., Laneve, G., Dong, Y., Ye, H., Liu, L., Guo, A., & Jiang, J. (2020). Identification of *Fusarium* head blight in winter wheat ears using continuous wavelet analysis. *Sensors, 20*(1), 20.

Mahlein, A.-K. (2016). Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease, 100*(2), 241–251.

Mahlein, A.-K., Alisaac, E., Al Masri, A., Behmann, J., Dehne, H.-W., & Oerke, E.-C. (2019). Comparison and combination of thermal, fluorescence, and hyperspectral imaging for monitoring *Fusarium* head blight of wheat on spikelet scale. *Sensors, 19*(10), 2281.

Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L. R., Davis, C. E., & Dandekar, A. M. (2015). Advanced methods of plant disease detection. a review. *Agronomy for Sustainable Development, 35*(1), 1–25.

McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G., & Van Sanford, D. (2012). A unified effort to fight an enemy of wheat and barley: Fusarium head blight. *Plant Disease, 96*(12), 1712–1728.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2020). *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071).* TU Wien. URL https://CRAN.R-project.org/package=e1071.

Mielniczuk, E., & Skwaryło-Bednarz, B. (2020). Fusarium head blight, mycotoxins and strategies for their reduction. *Agronomy, 10*(4), 509.

Morin, M., Lawrence, R., Repasky, K. S., Sterling, T. M., McCann, C., & Powell, S. (2017). Agreement analysis and spatial sensitivity of multispectral and hyperspectral sensors in detecting vegetation stress at management scales. *Journal of Applied Remote Sensing, 11*(4), Article 046025.

Nosek, B. A., Ebersole, C. R., DeHaven, A., & Mellor, D. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences, 115*(11), 2600–2606.

Pinter, P. J., Jr., Jackson, R. D., Elaine Ezra, C., & Gausman, H. W. (1985). Sun-angle and canopy-architecture effects on the spectral reflectance of six wheat cultivars. *International Journal of Remote Sensing, 6*(12), 1813–1825.

Qiu, R., Yang, C., Moghimi, A., Zhang, M., Steffenson, B. J., & Hirsch, C. D. (2019). Detection of *Fusarium* head blight in wheat using a deep neural network and color imaging. *Remote Sensing, 11*(22), 2658.

R Core Team. (2020). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org/.

Rebetzke, G. J., Fischer, R. T. A., van Herwaarden, A. F., Bonnett, D. G., Chenu, K., Rattey, A. R., & Fettell, N. A. (2014). Plot size matters: Interference from intergenotypic competition in plant phenotyping studies. *Functional Plant Biology, 41*(2), 107–118.

Saccon, F. A., Parcey, D., Paliwal, J., & Sherif, S. S. (2017). Assessment of *Fusarium* and deoxynivalenol using optical methods. *Food and Bioprocess Technology, 10*, 34–50.

Shen, G., Cao, Y., Yin, X., Dong, F., Xu, J., Shi, J., & Lee, Y.-W. (2022). Rapid and nondestructive quantification of deoxynivalenol in individual wheat kernels using near-infrared hyperspectral imaging and chemometrics. *Food Control, 131*, 108420.

Söderström, M., & Börjesson, T. (2013). Within-field variation in deoxynivalenol (DON) contents in oats. In J. V. Stafford (Ed.), *Precision agriculture '13* (pp. 329–334). Lleida, Spain: Wageningen Academic Publishers.

Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., & Buerstmayr, H. (2017). Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Tropical Plant Pathology, 42*, 165–174.

Stevens, A., & Ramirez-Lopez, L. (2013). *An introduction to the prospectr package.*

Su, W.-H., Zhang, J., Yang, C., Page, R., Szinyei, T., Hirsch, C. D., & Steffenson, B. J. (2021). Automatic evaluation of wheat resistance to Fusarium head blight using dual mask-RCNN deep learning frameworks in computer vision. *Remote Sensing, 13*(1), 26.

Thomas, S., Kuska, M. T., Bohnenkamp, D., Brugger, A., Alisaac, E., Wahabzada, M., Behmann, J., & Mahlein, A.-K. (2018). Benefits of hyperspectral imaging for plant disease detection and plant protection: A technical perspective. *Journal of Plant Diseases and Protection, 125*(1), 5–20.

Westad, F., & Marini, F. (2015). Validation of chemometric models – a tutorial. *Analytica Chimica Acta, 893*, 14–24.

117

113

Vaughan, M., Backhouse, D., & Del Ponte, E. M. (2016). Climate change impacts on the ecology of *Fusarium graminearum* species complex and susceptibility of wheat to *Fusarium* head blight: A review. *World Mycotoxin Journal, 9*(5), 685−700.

West, J. S., Bravo, C., Oberti, R., Lemaire, D., Moshou, D., & McCartney, H. A. (2003). The potential of optical canopy measurement for targeted control of field crop diseases. *Annual Review of Phytopathology, 41*, 593−614.

West, J. S., Canning, G. G., Perryman, S. A., & King, K. (2017). Novel Technologies for the detection of Fusarium head blight disease and airborne inoculum. *Tropical Plant Pathology, 42*, 203−209.

Whetton, R. L., Hassall, K. L., Waine, T. W., & Mouazen, A. M. (2018a). Hyperspectral measurements of yellow rust and fusarium head blight in cereal crops: Part 1: Laboratory study. *Biosystems Engineering, 166*, 101−115.

Whetton, R. L., Waine, T. W., & Mouazen, A. M. (2018b). Hyperspectral measurements of yellow rust and fusarium head blight in cereal crops: Part 2: On-line field measurement. *Biosystems Engineering, 167*, 144−158.

Xiao, Y., Dong, Y., Huang, W., Liu, L., Ma, H., Ye, H., & Wang, K. (2020). Dynamic remote sensing prediction for wheat Fusarium head blight by combining host and habitat conditions. *Remote Sensing, 12*(18), 3046.

Żelazny, W. R., Chrpová, J., Lukáš, J., & Hamouz, P. (2020). *Fusarium head blight detection from spectral measurements in a field phenotyping setting*. URL https://osf.io/xsfyp.

Zhang, D.-Y., Chen, G., Yin, X., Hu, R.-J., Gu, C.-Y., Pan, Z.-G., Zhou, X.-G., & Chen, Y. (2020d). Integrating spectral and image data to detect Fusarium head blight of wheat. *Computers and Electronics in Agriculture, 175*, 105588.

Zhang, D., Chen, G., Zhang, H., Jin, N., Gu, C., Weng, S., Wang, Q., & Chen, Y. (2020a). Integration of spectroscopy and image for identifying fusarium damage in wheat kernels. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 236*, 118344.

Zhang, D., Wang, D., Gu, C., Jin, N., Zhao, H., Chen, G., Liang, H., & Liang, D. (2019). Using neural network to identify the severity of wheat *Fusarium* head blight in the field environment. *Remote Sensing, 11*(20), 2375.

Zhang, D., Wang, Q., Lin, F., Weng, S., Lei, Y., Chen, G., Gu, C., & Zheng, L. (2020b). New spectral classification index for rapid identification of Fusarium infection in wheat kernel. *Food Analytical Methods, 13*, 2165−2175.

Zhang, D., Wang, Q., Lin, F., Yin, X., Gu, C., & Qiao, H. (2020c). Development and evaluation of a new spectral disease index to detect wheat *Fusarium* head blight using hyperspectral imaging. *Sensors, 20*(8), 2260.

Zhang, J., Wang, N., Yuan, L., Chen, F., & Wu, K. (2017). Discrimination of winter wheat disease and insect stresses using continuous wavelet features extracted from foliar spectral measurements. *Biosystems Engineering, 162*, 20−29.

Wiktor R. Żelazny: Application of feature selection for predicting leaf chlorophyll content in oats (*Avena sativa* L.) from hyperspectral imagery

Application of feature selection for predicting leaf chlorophyll content in oats (*Avena sativa* L.) from hyperspectral imagery

W.R. Żelazny[1,2,*]

[1]Crop Research Institute, Division of Crop Management Systems, Drnovská 507/73, CZ161 06 Praha 6 Ruzyně, Czech Republic
[2]Czech University of Life Sciences Prague, Faculty of Engineering, Department of Agricultural Machines, Kamýcká 129, CZ165 00 Praha 6 Suchdol, Czech Republic
[*]Correspondence: wzelazny@vurv.cz

Abstract. Feature selection can improve predictions generated by *partial least squares* models. In the context of hyperspectral imaging, it can also enable the development of affordable devices with specialized applications. The feasibility of feature selection for oat leaf chlorophyll estimation from hyperspectral imagery was assessed using a public domain dataset. A wrapper approach resulted in a simplistic model with poor predictive performance. The number of model inputs decreased from 94 to 3 bands when a filter approach based on the *minimum redundancy, maximum relevance* criterion was attempted. The filtering led to improved prediction quality, with the *root mean square error* decreasing from 0.17 to 0.16 g m^{-2} and R^2 increasing from 0.57 to 0.62. Accurate predictions were obtained especially for low chlorophyll levels. The obtained model estimated leaf chlorophyll concentration from near infra-red reflectance, canopy darkness, and its blueness. The prediction robustness needs to be investigated, which can be done by employing an ensemble methodology and testing the model on a new dataset with improved ground-truth measurements and additional crop species.

Key words: remote sensing, imaging spectroscopy, unmanned aerial vehicles, partial least squares, reproducibility.

INTRODUCTION

The indispensability of chlorophyll for plant photosynthesis (Sims & Gamon, 2002; Main et al., 2011) and its contribution to crop optical properties (Ollinger, 2011) make the estimation of leaf chlorophyll concentration an important remote sensing application. In large-scale assessments, leaf chlorophyll remote sensing is useful for yield prediction (Moharana & Dutta, 2016). At finer spatial scales, it can be used for the delineation of management zones for precision agriculture (Miao et al., 2009). As chlorophyll breaks down under stress, its monitoring provides information about the crop status, and enables a timely intervention to prevent the yield loss (Peñuelas et al., 1995; Sims & Gamon, 2002).

Traditional broad-band optical remote sensing relies on vegetation indexes for assessing crop status (e.g., Basso et al., 2016; Domínguez et al., 2017; Barbosa et al., 2019). Consequently, it is of limited use for estimating the concentrations of individual pigments, such as leaf chlorophyll. Many of these indexes have been adapted for use

2665

with hyperspectral imaging products (Miao et al., 2009; Verrelst et al., 2019, often leading to improved results (e.g., Miao et al., 2009). Moharana & Dutta (2016) evaluated ten indexes in terms of rice chlorophyll prediction from proximal spectroradiometric data. Some of the band combinations gave unsatisfactory estimates despite their high performance in other experimental settings, which is a common problem for vegetation indexes. On the other hand, the formulations that excelled during the screening provided realistic maps of rice chlorophyll concentration when applied to EO-1 Hyperion imagery. The limited index transferability across crops can be in part related to differences between plant architectures (Ollinger, 2011). A study involving six crop species evaluated the robustness of relationships between vegetation indexes and leaf chlorophyll with respect to canopy structural parameters. A total of 58 formulations were tested; of this number, only 2 were considered truly robust when applied to both measured and simulated spectra (Zou et al., 2015). Corti et al. (2018) published a meta-analysis intended to identify factors that foster accurate estimation of maize biochemical parameters from optical measurements. Their results suggest that satisfactory predictions can be obtained by avoiding certain families of vegetation indexes-regardless of sensor type, acquisition model, and crop developmental stage. The article indicates that only statistically significant relationships were included in the study, which means that this finding needs to be approached with caution. A recent review by Hatfield et al. (2019) cites additional studies devoted to vegetation indexes suitable for chlorophyll estimation. According to the authors, vegetation indexes should be a first choice in remote sensing applications, as they avoid computational challenges of more sophisticated approaches.

Yet, the main advantage of hyperspectral imagery lies in the possibility of applying 'full-spectrum' methods borrowed from chemometrics and machine learning (Corti et al., 2018; Verrelst et al., 2019). *Partial least squares (PLS) regression* was employed to diagnose chlorophyll levels in winter wheat leaf laboratory samples (Zhang et al., 2012). Scanning of single leaves under controlled illumination allowed the authors to evade the challenges inherent to canopy-level imaging in outdoor conditions, and without doubt contributed to extremely accurate (R statistics up to 0.99) predictions. Unfortunately, unclear study design description undermines the trustworthiness of the findings. Kanning et al. (2018) tested a pushbroom system as a way to overcome some limitations of 2D frame hyperspectral cameras. An experimental winter wheat field was scanned using a UAV, and the measurements subjected to *PLS* modelling. When the model was applied to the pixels of the field orthoimage, the individual nitrogen fertilization treatment levels could be discerned. The estimation quality was sufficient to fit a model for predicting grain yield from the obtained values. Meij et al. (2017) employed *PLS* to predict chlorophyll content in oats from unmanned aerial vehicle (UAV) campaign data. The study also included 25 published vegetation indexes. The *PLS* approach yielded validation predictions inferior to the estimates obtained by using the best of the indexes. Still, according to Verrelst et al. (2019), chemometric methods are in principle more powerful than vegetation indexes for estimating canopy biophysical parameters. The chemometric approach tends, in turn, to be surpassed by machine learning methods, capable of modelling non-linear relationships. A comparison of selected algorithms from both groups demonstrated substantial performance variability within the machine learning family. Robust leaf chlorophyll content predictions for multiple crops were obtained with *kernel ridge* and *Gaussian process regression*. On the other hand, *artificial neural networks*, an approach with a comparable level of sophistication, failed to provide

2666

consistently reliable estimates (Caicedo et al., 2014). By applying *support vector machines (SVM)* to maize hyperspectra, Karimi et al. (2008) obtained very good validation estimates for the tasseling stage. The prediction quality was worse, but still satisfactory, for the early growth stage, which the authors attributed to the soil showing through the crop canopy.

Despite its potential, the adoption of imaging spectroscopy remains hindered, in part by the high investment costs involved (Corti et al., 2018). Scene acquisition using a modern 2D camera tends to be slow due to sequential capture of a large number of bands. As a consequence, the speeds of airborne platforms become constrained (Honkavaara et al., 2017) and band registration needs to be performed during the imagery post-processing (Jakob et al., 2017). The voluminous data contained in hyperspectral data cubes require substantial computational capacities and specialized knowledge to process (Yang et al., 2017; Aasen et al., 2018). In the realm of field point spectrometry, similar challenges have been overcome by the development and commercialization of specialized proximity sensors, such as chlorophyll meters (Govender et al., 2009; Miao et al., 2009). These sensors exploit information from limited numbers of pre-selected bands, and have a predictive model embedded in the firmware to perform the computations. A similar route could be taken for imaging spectrometers in order to make the technology more accessible (Govender et al., 2009). One can envision an affordable specialized device capable of capturing narrow-band imagery, as hyperspectral cameras do, comprising bands that were pre-selected to optimize for accurate remote chlorophyll content estimation.

Feature selection methods have proven to be useful for the screening of spectral bands for a variety of applications. In addition to reducing the number of required model inputs, they were shown to improve the prediction accuracy (Ding & Peng, 2005; Mehmood et al., 2012). Fewer computations are required to process data subjected to feature selection, and model interpretation is facilitated (Ding & Peng, 2005). Band pre-selection prior to data acquisition can also address the problem of slow operation of hyperspectral cameras (Yang et al., 2013; Zhang & He, 2013). As demonstrated by the Zhang & He (2013) oilseed rape yield study, substantial reduction of data volume can be attained without impairing model performance. Discarding of 98% of hyperspectral bands had a minimal effect on the quality of nitrogen content prediction in pepper plants, while significantly simplifying the obtained model (Yu et al., 2014). Behmann et al. (2014) proposed an *SVM* model for detecting water stress in barley. The model inputs comprised vegetation indexes, the combinations of which were determined using wrapper feature selection. Increased detection sensitivity was obtained, allowing for earlier drought detection relative to the raw indexes. The aim of the present study is to investigate the effect of two feature selection approaches on the prediction of leaf chlorophyll concentration in oats from hyperspectral imaging data.

MATERIALS AND METHODS

Experimental data

The present study partially replicates and extends the results of Meij et al. (2017), using the same experimental data. Their experiment evaluated the soil-mediated carry-over effects of preceding and cover crops on crop-of-interest status. The data collection took place in summer 2015, which was the second year of the study, and was focused on experimental plots with oats in the grain-filling developmental stage.

2667

The dataset includes narrow-band reflectance spectra of the experimental plots (one averaged spectrum per plot) obtained from UAV imagery. The spectra cover the range of wavelengths from 450 to 915 nm, i.e., between visible blue and near infra-red. The spectral resolution is 5 nm, thus yielding 94 bands. The spectra are accompanied by ground-truth measurements describing the crop's physiological status. They include, among others, SPAD-estimates of leaf chlorophyll concentrations (one averaged estimate per plot), which are the focus of the present study. There are 56 data points in total, labelled as either calibration or validation data in 1:1 proportion. The dataset is in the public domain, and for the purpose of this study, it was downloaded from the Dryad repository (Meij et al., 2018).

Reproduction of the original analysis

In order to obtain a baseline for the assessment of feature selection performance, a reproduction of the Meij et al. (2017) result was prepared. The original study employed vegetation indexes and *PLS* modelling for predicting leaf chlorophyll from the imaging spectra. This paper focuses on the latter approach.

The data partitioning from the original dataset was preserved, and a *PLS regression* model was fitted to the calibration subset. Leaf chlorophyll concentration was modelled as the dependent variable, and the reflectance values for the whole range of the wavelengths as the independent variables. The number of latent variables was tuned using *leave-one-out cross-validation* by calculating the *cross-validation root mean square error (RMSE)* for each value from between 1 and 20. The validation spectra were then fed to the model exhibiting the lowest error, and the generated predictions compared with the SPAD chlorophyll estimates to obtain *validation RMSE, normalized RMSE (NRMSE),* and the R^2 statistics. To reproduce the original validation results, *RMSE* had to be normalized by dividing it by the mean chlorophyll concentration, rather than the standard deviation or range. Likewise, R^2 had to be calculated as the square of the *correlation* coefficient between the predicted and observed values, rather than derived from the sums of squares.

Application of feature selection

Next, the fitting of the *PLS* model to the calibration dataset was repeated, but in addition to the tuning of the latent variable number, feature selection was performed. Two approaches to feature selection were tested: a filter method based on the *minimum redundancy, maximum relevance (MRMR)* criterion, and a forward selection wrapper method.

Under the filtering approach, variables are evaluated independently of model fitting, according to a measure the value of which determines which of them will be discarded (Mehmood et al., 2012). In the *MRMR* method, this measure is the mutual information shared by the candidate feature and the predicted variable, reduced by the average mutual information shared by the candidate feature and the features already accepted for inclusion into the model. The mutual information is a function of the *correlation* coefficient (De Jay et al., 2013).

With wrapping, models are fitted to multiple pre-selected feature subsets, and the fit quality itself serves as the selection performance criterion, making it a computationally more demanding approach (Mehmood et al., 2012). The wrapper forward selection method is analogous to the forward selection in the *stepwise regression:* candidate

2668

features are picked one by one from the feature pool, and their influence on the performance of the refitted model is assessed. The variable associated with the highest performance increase is kept in the model, and the process continues iteratively, until there is no further improvement.

For each method, the present study aimed to obtain a series of models with the input feature number ranging from 2 to all 94 bands (i.e., no selection). In this way, the influence of feature selection intensity on the prediction quality could be investigated.

Computational reproducibility

The analysis was prepared with reproducibility in mind (Piccolo & Frampton, 2016). It was programmed in the R language (R Core Team, 2019), using the packages pls (Mevik et al., 2019) for model fitting, mRMRe (De Jay et al., 2013) for assessing the *MRMR* criterion, and mlr (Bischl et al., 2016) for model tuning. GNU Make (Stallman et al., 2016) was used as the build tool, and GNU Guix enabled isolation and reproducibility of the software environment for performing the analysis (Courtès & Wurmus, 2015). The computational scripts are available from a Zenodo repository (Żelazny, 2020). On an IA-64 machine, the analysis took approximately 100 minutes without parallelization and excluding the time needed to set up the environment. The latter can last hours on the first run, depending on the state of a the Guix store (Courtès & Wurmus, 2015) and availability of pre-compiled package substitutes. It is reduced to minutes on subsequent runs.

RESULTS AND DISCUSSION

Visual data assessment
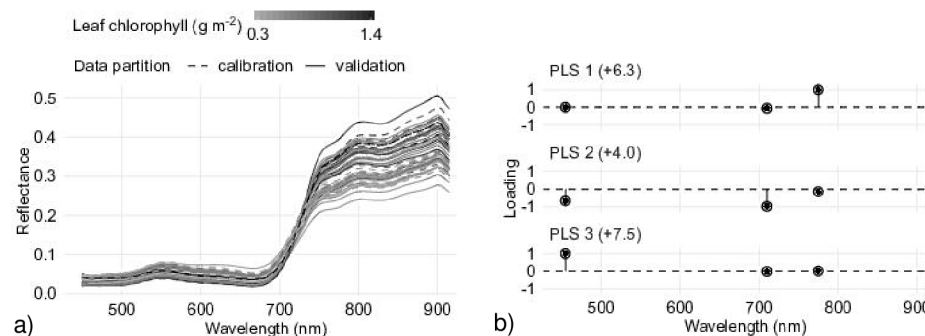
Figure 1. a) Narrow-band spectra of experimental oat plots in the calibration and validation data subsets acquired using an unmanned aerial vehicle. Line hues reflect the differences in SPAD-estimated leaf chlorophyll concentrations. The figure can be rendered in color by running the computational scripts that accompany the article; b) Loadings in the *partial least squares* model for predicting leaf chlorophyll concentrations from the narrow-band spectra. The model is based on three bands obtained from *minimum redundancy, maximum relevance* filtering. Latent variable loadings are given in the parentheses, wavelength loadings are given on the *y* axis.

Fig. 1, a depicts the experimental plot spectra matched to the ground-truth data, analogously to Fig. 4 in Meij et al. (2017). High leaf chlorophyll concentration appears

2669

to be associated with increased near infra-red reflectance and a steep red edge–both regions repeatedly considered important for chlorophyll prediction by earlier studies (Govender et al., 2009; Main et al., 2011). On the other hand, contrary to expectation, no apparent red-edge shift can be discerned. The calibration and validation spectra are well mixed in terms of the chlorophyll measurements, as can be expected from the stratified random partitioning, employed by the original study. Regarding the reflectance, the validation subset seems to cover a wider range of values than the calibration subset, but the difference is too small to raise concerns about a mismatch between the partitions.

Reproduction of Meij (Meij et al. (2017)

Despite the variety of existing *PLS* flavours and implementations, the attempt to reproduce the validation results of the Meij et al. (2017) paper turned out to be successful, with only *NRMSE* showing a slight deviation (Table 1, row 'Reproduction'). However, as discussed above, the high number of bands contributing to the model make the 'full-spectrum' approach infeasible for practical application - at least until hyperspectral imagers become affordable (Aasen et al., 2018). In addition, the result of model tuning, which set the number of the latent variables to five, makes an insight into its workings challenging.

Table 1. Tuning parameters and validation statistics of the *partial least squares* models. Each model was calibrated using 28 spectra and validated using another set of 28 spectra

Study	Input bands	Latent variables	*RMSE* (g m^{-2})	*NRMSE* (%)	R^2
Meij et al. (2017)	94	5	0.17	23.82	0.57
Reproduction	94	5	0.17	23.75	0.57
Filter feature selection	19	7	0.21	28.36	0.52
Filter feature selection (truncated)	3	3	0.16	21.84	0.62
Wrapper feature selection	1	1	0.20	28.23	0.43

RMSE = root mean square error, NRMSE = normalized root mean square error.

Feature selection

The *cross-validation* results of models employing filter feature selection exhibit two local error minima (Fig. 2). The absolute minimum corresponds to 19 input bands, a much lower number than for the reference model, but still too high for developing reasonably priced specialized device. What is more, the model shows higher validation error and involves even more latent variables (seven) than the reproduction model (Table 1, 'Filter feature selection').

Conversely, three wavelengths, as in the second minimum, seem a good middle-ground between technical feasibility and expected estimation error. The fact that the number of latent variables in *PLS regression* cannot exceed the number of inputs contributes to the model interpretability. Notable is the improvement of the validation statistics (Table 1, 'Filter feature selection (truncated)'), which corroborates the positive influence of feature selection on prediction accuracy (Mehmood et al., 2012). Although the obtained gains may seem modest, one should consider other advantages offered by feature selection, such as the reduced cost of a specialized imager (Govender et al., 2009), more efficient data acquisition (Yang et al., 2013; Zhang & He, 2013), and

2670

smaller volumes of the collected data (Zhang & He, 2013). On a closer examination, the model appears to give accurate predictions for low levels of chlorophyll, but its performance deteriorates above the level of about 0.75 g m^{-2} (Fig. 3). A similar pattern occurred in the Kanning et al. (2018) pushbroom imager study. An attempt to further improve the prediction quality could be made by log-transforming the chlorophyll content values prior to modelling.

Fig. 1, b depicts the band loadings for each latent variable and the latent variable loadings of this model. The chlorophyll content is, thus, predicted as $LCC = 6.3\,PLS_1 + 4.0\,PLS_2 + 7.5\,PLS_3$. The value of the first component $PLS_1 = 0.0\,r_{455} - 0.1\,r_{710} + 1.0\,r_{775}$ corresponds to the near infra-red reflectance, in accordance with the visual assessment, above. The second component $PLS_2 = -0.7\,r_{455} - 1.0\,r_{710} - 0.1\,r_{775}$ includes the bottom part of the red edge and, interestingly, a blue band., it can be interpreted as canopy darkness (low visible albedo), and linked to the absorbance in the photosynthetically-active spectral region. The third component value $PLS_3 = 1.0\,r_{455} + 0.0\,r_{710} + 0.0\,r_{775}$ is determined by canopy blueness (blue hue intensity).

Figure 2. *Cross-validation* prediction performance and tuning results of the oat leaf chlorophyll prediction models according to the number of features selected using the *minimum redundancy, maximum relevance* filter. *CV RMSE = cross-validation root mean square error.*

Wavelength combinations similar to the one picked by the filtering algorithm seldom occur in vegetation index formulations. They can be found in the Enhanced Vegetation Index (Gao et al., 2000), the Structure Insensitive Pigment Index (Peñuelas et al., 1995), the Modified Simple Ratio, and the Modified Normalized Difference (mND$_{705}$) (Sims & Gamon, 2002). No such index was investigated by Meij et al. (2017). In the study by Main et al. (2011), the first three indexes fared poorly when used for predicting chlorophyll content in maize leaves at various developmental stages. The authors attribute this to the weak relationship between the blue spectral region and the leaf chlorophyll concentration.

Regarding mND$_{705}$, it was among the best-performing indexes in Main et al. (2011), and in Miao et al. (2009) - also a maize study. On the other hand, it occurred to be a poor predictor of chlorophyll content in rice (Moharana & Dutta, 2016). The mND$_{705}$ index formula includes blue reflectance as a way to account for specular reflectance (Sims & Gamon, 2002). The third latent variable of the discussed *PLS* model may play the same role.

Alternatively, it may adjust for Rayleigh scattering. According to Beisl et al. (2008), atmospheric effects occur even in low-altitude airborne remote sensing applications. Although the analysed dataset has been subjected to atmospheric correction, it was based on a single reference panel measurement (Meij et al., 2017). The weakness of this approach is the assumption of constant illumination conditions as

2671

individual images are acquired. The blue band information may account for the residual error that still remained after the correction.

The forward selection within the wrapper approach stopped after picking one band (775 nm), thus reducing the *PLS* model to a classical *regression* model with a single independent variable. The selected wavelength lies in the near infra-red spectral region, which agrees with the observation from the visual assessment, above. According to the validation statistics (Table 1 'Wrapper feature selection'), despite its extreme simplicity, the model performs surprisingly well in terms of *RMSE*. However, the low R^2 value puts in question the feasibility of its practical use. Moreover, like the preceding model, it exhibits uneven prediction quality for various levels of chlorophyll (Fig. 3).
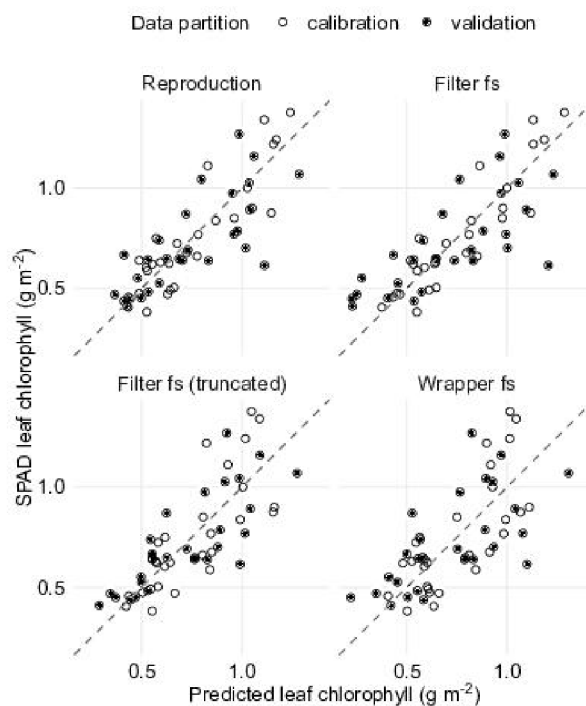
Figure 3. Prediction error patterns of the studied models with respect to the ground-truth data.

In the light of this finding, it can be recommended to avoid wrapper selection for chlorophyll content prediction, especially considering the substantial computational demands of this approach (Ding & Peng, 2005; Mehmood et al., 2012). Conversely, the encouraging results attained with *MRMR* suggest high potential of the filter strategy towards picking highly predictive spectral bands. The *MRMR* criterion seems particularly well-suited to data acquired using optical remote sensing methods. As reflectance measurements exhibit substantial spectral autocorrelation (Karimi et al., 2008; Verrelst et al., 2019), a naive algorithm could pick a set of neighbouring bands, with information content barely exceeding that of a single band. The 'minimum redundancy' aspect of *MRMR* avoids this issue by taking correlations between features

2672

into consideration (Ding & Peng, 2005). Still future research might consider examination of feature selection methods from the filter family. The performance of the three classes of methods reviewed by Mehmood et al. (2012): based on loading weights, *regression coefficients*, and variable importance in projection; could be compared, for instance.

Possibilities of assessing and improving study generalizability

The present study illustrates the application of feature selection for obtaining a parsimonious predictive model with high interpretability. Just as omitting model *cross-validation* can lead to over-fitting, a model that performs well on a single validation dataset does not necessarily generalize to new circumstances. This is especially true for unstable models, whose parameters change radically in response to even slight modification of the training data.

In the present study, an improvement of validation statistics was obtained after filtering the spectral bands using the *MRMR* algorithm. As highlighted by De Jay et al. (2013), the algorithm in its original form produces results that are unstable with respect to data modifications. The cited authors proposed an ensemble extension of the filter to stabilize its output.

Ensemble modelling has been shown to improve prediction accuracy, as exemplified by *random forests* (Breiman, 2001), and enable interval estimation, as exemplified by *bootstrap* methods (Wood, 2005). Its obvious application in the discussed study would be to abandon the fixed data partitioning, which was inherited from Meij et al. (2017), in favour of multiple analyses, each based on a different assignment of the data points to the calibration and validation subsets. By the subsequent aggregation of the obtained partial results, the stability of the best performing models could be assessed - not only with respect to the selected wavelengths, but also to their loadings and validation statistics.

Two candidate models fitted to filtered bands were elected by hand for further evaluation based on *CV RMSE* and feature selection intensity as an auxiliary criterion. Repeated data partitioning would result in proliferation of models, making the manual approach unfeasible. Replacing it with an algorithm would necessitate taking both optimization criteria into account, which can be accomplished with aid of *model-based multi-objective optimization* (Horn et al., 2015).

These avenues could not have been taken due to high computational complexity involved, especially if wrapper feature selection were also included. In the future, an adaptation of the analysis for an execution in a high-performance computing environment might be attempted. At that point, an extension of the study to include ensemble modelling would become feasible.

An evident weakness of both the present and the original Meij et al. (2017) study is the fact that the ground-truth data were obtained using a SPAD chlorophyll meter, and thus include spectroscopic estimation errors (Uddling et al., 2007). It is possible that similar errors present in the discussed *PLS* results become masked in the consequence, leading to overoptimistic validation statistics. Therefore, it would be desirable to replicate the study using laboratory analyses for the ground truth, instead.

Spectral responses of leaf pigments differ across plant genotypes. Although the chlorophyll signal is readily discernible in a leaf or canopy spectrum (Ollinger, 2011), the reflectance is modified by additional factors. They include leaf and canopy anatomy and morphology (Asner, 1998; Jacquemoud & Ustin, 2001; Ollinger, 2011) and spectral

2673

properties of additional foliar pigments present in the tissues (Jacquemoud & Ustin, 2001; Ollinger, 2011). Research is needed to establish whether feature selection can yield a set of bands that enable calibration of models for chlorophyll content estimation in multiple crops, and how big this set needs to be for the models to be accurate.

CONCLUSIONS

Filtering of bands according to the *minimum redundancy, maximum relevance* criterion can improve the performance of a *partial least squares* model aimed at oat leaf chlorophyll prediction from airborne hyperspectral imagery. Chlorophyll concentration can be estimated from near infra-red reflectance, canopy darkness, and its blueness. The obtained size of the feature space (three bands in the present study) is sufficiently small for the development of affordable single-purpose imagers. Although a wrapper approach based on forward feature selection can yield an even more parsimonious model, the resulting prediction quality is not satisfactory. The robustness of the findings remains to be investigated using an ensemble of dataset partitionings and ground truth obtained from laboratory analyses based on samples collected from multiple crops.

REFERENCES

Aasen, H., Honkavaara, E., Lucieer, A. & Zarco-Tejada, P. 2018. Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: A review of sensor technology, measurement procedures, and data correction workflows. *Remote Sensing* **10**, 1091.

Asner, G.P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment* **64**, 234–253.

Barbosa, B., Ferraz, G., Gonçalves, L., Marin, D., Maciel, D., Ferraz, P. & Rossi, G. 2019. RGB vegetation indices applied to grass monitoring: A qualitative analysis. *Agronomy Research* **17**, 349–357. https://doi.org/10.15159/AR.19.119

Basso, B., Fiorentino, C., Cammarano, D. & Schulthess, U. 2016. Variable rate nitrogen fertilizer response in wheat using remote sensing. *Precision Agriculture* **17**, 168–182.

Behmann, J., Steinrücken, J. & Plümer, L. 2014. Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing* **93**, 98–111.

Beisl, U., Telaar, J. & Schönermark, M. 2008. Atmospheric correction, reflectance calibration and BRDF correction for ADS40 image data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **37**, 7–12.

Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., Casalicchio, G. & Jones, Z.M. 2016. mlr: Machine learning in R. *Journal of Machine Learning Research* **17**, 1–5.

Breiman, L. 2001. Random forests. *Machine Learning* **45**, 5–32.

Caicedo, J.P.R., Verrelst, J., Muñoz-Marí, J., Moreno, J. & Camps-Valls, G. 2014. Toward a semiautomatic machine learning retrieval of biophysical parameters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7**, 1249–1259.

2674

Corti, M., Cavalli, D., Cabassi, G., Gallina, P.M. & Bechini, L. 2018. Does remote and proximal optical sensing successfully estimate maize variables? A review. *European Journal of Agronomy* **99**, 37–50.

Courtès, L. & Wurmus, R. 2015. Reproducible and user-controlled software environments in HPC with Guix. In: *Euro-Par 2015: Parallel Processing Workshops*. Vienna University of Technology, Vienna, pp. 579–591.

De Jay, N., Papillon-Cavanagh, S., Olsen, C., El-Hachem, N., Bontempi, G. & Haibe-Kains, B. 2013. MRMRe: An R package for parallelized mRMR ensemble feature selection. *Bioinformatics* **29**, 2365–2368.

Ding, C. & Peng, H. 2005. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology* **3**, 185–205.

Domínguez, J., Kumhálová, J. & Novák, P. 2017. Assessment of the relationship between spectral indices from satellite remote sensing and winter oilseed rape yield. *Agronomy Research* **15**, 55–68.

Gao, X., Huete, A.R., Ni, W. & Miura, T. 2000. Optical–biophysical relationships of vegetation spectra without background contamination. *Remote Sensing of Environment* **74**, 609–620.

Govender, M., Govender, P.J., Weiersbye, I.M., Witkowski, E.T.F. & Ahmed, F. 2009. Review of commonly used remote sensing and ground-based technologies to measure plant water stress. *Water SA* **35**, 741–752.

Hatfield, J.L., Prueger, J.H., Sauer, T.J., Dold, C., O'Brien, P. & Wacha, K. 2019. Applications of vegetative indices from remote sensing to agriculture: Past and future. *Inventions* **4**, 71.

Honkavaara, E., Rosnell, T., Oliveira, R. & Tommaselli, A. 2017. Band registration of tuneable frame format hyperspectral UAV imagers in complex scenes. *ISPRS Journal of Photogrammetry and Remote Sensing* **134**, 96–109.

Horn, D., Wagner, T., Biermann, D., Weihs, C. & Bischl, B. 2015. Model-based multi-objective optimization: Taxonomy, multi-point proposal, toolbox and benchmark. In: *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, Guimarães, Portugal, pp. 64–78.

Jacquemoud, S. & Ustin, S.L. 2001. Leaf optical properties: A state of the art. In: *8th International Symposium of Physical Measurements & Signatures in Remote Sensing*. CNES, Aussois, pp. 223–332.

Jakob, S., Zimmermann, R. & Gloaguen, R. 2017. The need for accurate geometric and radiometric corrections of drone-borne hyperspectral data for mineral exploration: MEPHySTo – a toolbox for pre-processing drone-borne hyperspectral data. *Remote Sensing* **9**, 88.

Kanning, M., Kühling, I., Trautz, D. & Jarmer, T. 2018. High-resolution UAV-based hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction. *Remote Sensing* **10**, 2000.

Karimi, Y., Prasher, S., Madani, A. & Kim, S. 2008. Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations. *Canadian Biosystems Engineering* **50**, 13–20.

Main, R., Cho, M.A., Mathieu, R., O'Kennedy, M.M., Ramoelo, A. & Koch, S. 2011. An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS Journal of Photogrammetry and Remote Sensing* **66**, 751–761.

Mehmood, T., Liland, K.H., Snipen, L. & Sæbø, S. 2012. A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* **118**, 62–69.

Meij, B. van der, Kooistra, L., Suomalainen, J., Barel, J.M. & Deyn, G.B.D. 2017. Remote sensing of plant trait responses to field-based plant–soil feedback using UAV-based optical sensors. *Biogeosciences* **14**, 733–749.

Meij, B. van der, Kooistra, L., Suomalainen, J., Barel, J.M. & De Deyn, G.B. 2018. Data from: Remote sensing of plant trait responses to field-based plant–soil feedback using UAV-based optical sensors. doi:10.5061/dryad.75k1d

2675

Mevik, B.-H., Wehrens, R. & Liland, K.H. 2019. *pls: Partial least squares and principal component regression*. https://CRAN.R-project.org/package=pls

Miao, Y., Mulla, D.J., Randall, G.W., Vetsch, J.A. & Vintila, R. 2009. Combining chlorophyll meter readings and high spatial resolution remote sensing images for in-season site-specific nitrogen management of corn. *Precision Agriculture* **10**, 45–62.

Moharana, S. & Dutta, S. 2016. Spatial variability of chlorophyll and nitrogen content of rice from hyperspectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* **122**, 17–29.

Ollinger, S.V. 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist* **189**, 375–394.

Peñuelas, J., Baret, F. & Filella, I. 1995. Semi-empirical indices to assess carotenoids/chlorophyll *a* ratio from leaf spectral reflectance. *Photosynthetica* **31**, 221–230.

Piccolo, S.R. & Frampton, M.B. 2016. Tools and techniques for computational reproducibility. *GigaScience* **5**, 30.

R Core Team. 2019. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Sims, D.A. & Gamon, J.A. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment* **81**, 337–354.

Stallman, R.M., McGrath, R. & Smith, P.D. 2016. *GNU Make. A program for directing recompilation*. Free Software Foundation, Boston. 220 pp.

Uddling, J., Gelang-Alfredsson, J., Piikki, K. & Pleijel, H. 2007. Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynthesis Research* **91**, 37–46.

Verrelst, J., Malenovský, Z., Van der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J.-P., Lewis, P., North, P. & Moreno, J. 2019. Quantifying vegetation biophysical variables from imaging spectroscopy data: A review on retrieval methods. *Surveys in Geophysics* **40**, 589–629.

Wood, M. 2005. Bootstrapped confidence intervals as an approach to statistical inference. *Organizational Research Methods* **8**, 454–470.

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., Xu, B., Yang, X., Zhu, D., Zhang, X., Zhang, R., Feng, H., Zhao, X., Li, Z., Li, H. & Yang, H. 2017. Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. *Frontiers in Plant Science* **8**, 1–26.

Yang, W., Duan, L., Chen, G., Xiong, L. & Liu, Q. 2013. Plant phenomics and high-throughput phenotyping: Accelerating rice functional genomics using multidisciplinary technologies. *Current Opinion in Plant Biology* **16**, 180–187.

Yu, K.-Q., Zhao, Y.-R., Li, X.-L., Shao, Y.-N., Liu, F. & He, Y. 2014. Hyperspectral imaging for mapping of total nitrogen spatial distribution in pepper plant. *PLoS ONE* **9**, e116205.

Zhang, D., Wang, X., Ma, W. & Zhao, C. 2012. Research vertical distribution of chlorophyll content of wheat leaves using imaging hyperspectra. *Intelligent Automation & Soft Computing* **18**, 1111–1120.

Zhang, X. & He, Y. 2013. Rapid estimation of seed yield using hyperspectral images of oilseed rape leaves. *Industrial Crops Products* **42**, 416–420.

Zou, X., Hernández-Clemente, R., Tammeorg, P., Lizarazo Torres, C., Stoddard, F.L., Mäkelä, P., Pellikka, P. & Mõttus, M. 2015. Retrieval of leaf chlorophyll content in field crops using narrow-band indices: Effects of leaf area index and leaf mean tilt angle. *International Journal of Remote Sensing* **36**, 6031–6055.

Żelazny, W.R. 2020. Application of feature selection for predicting leaf chlorophyll content in oats (Avena sativa L.) from hyperspectral imagery (computational scripts). doi:10.5281/zenodo.3573320

2676

Appendix D

Wiktor R. Żelazny and Tomáš Šimon: Calibration spiking of MIR-DRIFTS soil spectra for carbon predictions using CPLSR and log-ratio transformations

agriculture

MDPI

Calibration Spiking of MIR-DRIFTS Soil Spectra for Carbon Predictions Using PLSR Extensions and Log-Ratio Transformations

Wiktor R. Żelazny [1,2,*] and **Tomáš Šimon** [1]

1 Division of Crop Management Systems, Crop Research Institute, Drnovská 507/73, CZ161 06 Praha, Czech Republic; cropscience@vurv.cz

2 Faculty of Engineering, Czech University of Life Sciences Prague, Kamýcká 129, CZ165 00 Praha, Czech Republic

* Correspondence: wzelazny@vurv.cz; Tel.: +420-702-087-676

Abstract: There is a need to minimize the usage of traditional laboratory reference methods in favor of spectroscopy for routine soil carbon monitoring, with potential cost savings existing especially for labile pools. Mid-infrared spectroscopy has been associated with accurate soil carbon predictions, but the method has not been researched extensively in connection to C lability. More studies are also needed on reducing the numbers of samples and on how to account for the compositional nature of C pools. This study compares performance of two classes of partial least squares regression models to predict soil carbon in a global (models trained to data from a spectral library), local (models trained to data from a target area), and calibration-spiking (spectral library augmented with target-area spectra) scheme. Topsoil samples were+ scanned with a Fourier-transform infrared spectrometer, total and hot-water extractable carbon determined, and isometric log-ratio coordinates derived from the latter measurements. The best RMSEP was estimated as 0.38 and 0.23 percentage points TC for the district and field scale, respectively—values sufficiently low to make only qualitative predictions according to the RPD and RPIQ criteria. Models estimating soil carbon lability performed unsatisfactorily, presumably due to low labile pool concentration. Traditional weighing of spiking samples by including multiple copies thereof in training data yielded better results than canonical partial least squares regression modeling with embedded weighing. Although local modeling was associated with the most accurate predictions, calibration spiking addressed better the trade-off between data acquisition costs and model quality. Calibration spiking with compositional data analysis is, therefore, recommended for routine monitoring.

Keywords: soil organic matter; MIR-DRIFTS; chemometrics; compositional data analysis; reproducibility

1. Introduction

SC is a primary indicator of soil quality [1,2], and in recent years, estimation of atmospheric CO_2 sequestration has boosted interest in SC monitoring [3–6]. In addition to SC quantity, its fractional composition can be of interest in evaluating soil status. Research has been devoted to the labile fraction, which can give insight into SC turnover processes [7]. Labile C determines the rate of nitrogen release from soil organic matter, a factor to be accounted for while fertilizing the soil [8,9], and it can also inform about the long-term stability of sequestered carbon [10].

Changes in SC content occur over long time frames [5,11]—in certain conditions also on arable land despite higher risk of depletion by mineralization [1,12]. Although it suffices to sample soil every ten years for monitoring [3,5], SC can exhibit high spatial variability [3,4], which increases the necessary sampling effort [13]. Additional collection campaigns are needed to capture the dynamics of SC labile pools, which, on arable land, are readily influenced by fertilizer and soil amendment inputs, crop residue management, and soil tillage [11,14,15]. Traditional analysis of samples collected for this purpose is

costly and time consuming due to the laboriousness of laboratory SC fractionation [16–19]. Environmental concerns have also been raised [20,21].

Higher throughput and economical viability can be attained with soil spectroscopy [4,22]. Here, MIR-DRIFTS is one of the methods considered suitable for chemical soil analysis [13,20,23] owing to fundamental vibrations of soil molecules arising in the MIR spectral region [6,13,24]. In particular, it can give accurate estimates of SC content [13,22,25,26], and according to Reeves III [25], this high performance may extend to SC fraction assessments. However, the modest number of publications devoted to SC lability [27] is in contrast with the extensive literature on total C (TC) or the large organic C (OC) pool estimation with MIR-DRIFTS.

Quantitative assessment of soil properties from spectral measurements requires a predictive model trained to a reference dataset, in which spectra are paired with reference laboratory data [4,28]. Bellon-Maurel and McBratney [26] and Gholizadeh et al. [29] stress an importance of a large calibration library for satisfactory accuracy. In particular, the number of samples corresponding to soil properties similar to those in the target area should be sufficient to avoid a prediction bias with the trained model [13,30,31]. Applications of libraries have been limited in MIR spectroscopy [32], and although large collections are increasingly available [16,32–34], many regions remain not represented. An important prerequisite is to follow the sample collection and analysis methodology that was employed for building the library [25,28,35]. This is problematic given the fact that even different units of one spectrometer model can yield MIR scans that do not match [29].

For scenarios with an insufficient library size or coverage, calibration spiking can be employed [6,36]. The library is augmented with a limited number of samples collected at the target site prior to the training of a predictive model [37,38]. Samples for calibration spiking can be picked according to leverage selection to minimize their number or spiking intensity [28]. This process preserves the representativeness of the resulting subset by taking into account spectral similarities of the samples in the available pool [37]. According to Guerrero et al. [39], a reference library does not need to be large to obtain satisfactory predictions with calibration spiking. However, even with a modestly sized reference dataset, there is going to be a disproportion between the number of spiking and library samples. One way of addressing this problem is to use a subset of the latter [38]. As an alternative, which does not incur information loss, local samples can be given bigger weight relative to the samples in the library. Such weighing is typically performed by multiplying the local sample occurrences in a model training dataset [36,39,40]. However, another approach is also possible, where a model allowing for specification of case weights is employed instead [41].

Partial least squares regression (PLSR) continues to be the most common approach for analyzing soil spectra and predictive model calibration [3,13], including MIR-DRIFTS SC studies [22,26]. When estimating multiple properties, accuracy can be improved by accounting for their correlations [42], and utility of multiresponse PLSR (PLSR2) models in pedology has been demonstrated before [43–46]. Indahl et al. [47] proposed combining PLSR with canonical correlation analysis and developed the canonical PLSR (CPLSR) class of models. Like PLSR2, this method permits a multivariate response variable, but in addition to that, it offers a possibility to weigh the individual observations.

Baumann et al. [34] hypothesized that library samples "would stabilize and reduce the errors" associated with spike samples. However, spiking a reference library that does not match the target calibration domain can lead to less satisfactory results than the training of a model to local samples only [37,41]. Guerrero et al. [38] and Wetterlind and Stenberg [48] questioned the necessity of a reference library at all by pointing to superior model calibrations obtained with samples from the vicinity of a target area, exclusively.

The aim of this study is to investigate the influence of calibration spiking and local modeling on SC content and lability prediction performance of PLSR2 and CPLSR models trained to MIR-DRIFTS spectra corresponding to crop farming localities with different soil and climatic conditions. We hypothesized that the spiking of a library with observations from several long-term experiments would reduce the number of samples subjected to

traditional laboratory analysis compared to relying only on target-site spectra. Furthermore, CPLSR models with embedded sample weighing were expected to perform better than weighing by multiplication followed with PLSR2 modeling. The study also explores the influence of spectra pre-processing schemes and leverage sampling algorithms on the model predictions.

2. Materials and Methods

2.1. Site Description and Data Collection

Two groups of soil samples were collected at the territory of the Czech Republic: (1) time series of archived samples obtained from long-term crop trials, which served as a reference library, and (2) samples from two commercial sites, Ústí nad Orlicí and Janovice, as prediction targets of interest (Figure 1). The long-term experiments were maintained by the Crop Research Institute Praha-Ruzyně (CRI) and the Central Institute for Supervising and Testing in Agriculture; their primary focus was fertilization. A brief description can be found in Table 1. As seen in Table S1, the library was unbalanced with respect to the sample, year, and experimental treatment counts. Topsoil samples from the upper 20 cm were collected using a field shovel following a uniform protocol. The soil was collected from three spots of each plot, and the partial samples were combined into approximately 2 kg lots and homogenized.

Ústí nad Orlicí comprises multiple localities scattered over one district (Figure 1), making it a heterogeneous site. The fields were managed with conventional tillage and sown with winter wheat, winter and spring barley, silage maize, and oilseed rape. The heterogeneity was additionally augmented by an extended timing of the soil sample collection, which took place every spring and fall between 2012 and 2015. About 40 topsoil samples from fields with winter wheat and winter barley were collected by the farmers or their designated persons during each campaign, yielding a total of 335 samples. The commercial site Janovice denotes a single conventionally tilled field, with a crop rotation of silage maize, winter wheat, potatoes, and clover–grass mixture. It contributed 45 topsoil (0–20 cm) samples collected by CRI employees in fall 2017, after the sowing of winter wheat. The sampling points were delimited every 120 m in a way to obtain roughly uniform coverage of the field. There were six partial samples per composite sample of approximately 0.5 kg, which was then homogenized.
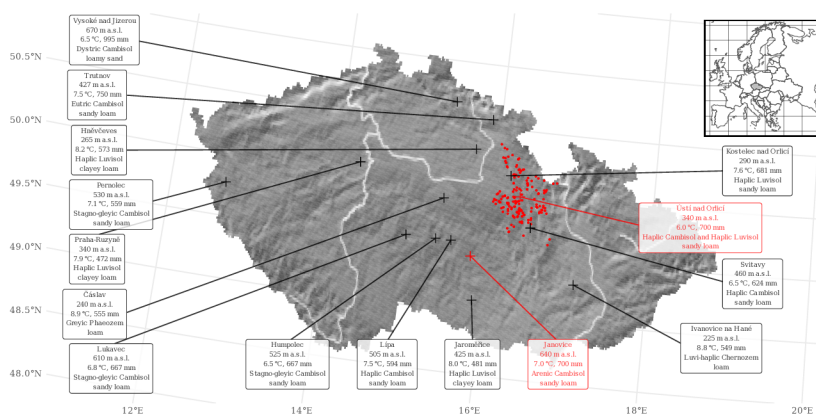

Figure 1. Locations, altitudes, mean annual temperatures, and precipitation sums in the years of data collection, and soil types and textures at the experimental sites. The target sites are marked with red color. For Ústí nad Orlicí, individual soil sampling locations are displayed, and their mean altitude is provided.

Table 1. Characteristics of the long-term field experiments.

Experiment [a]	Est.	Layout [b]	Crop Rotation [c]	Reference
CRE	1956	1b × 3t	various (25%)–(WW or TR)–(POT or SB or SM)–(SBA or WW)	[49]
CRT	1984	5b × 1t	WW and SBA (50–100%) complemented with CL, O, PEA, SB, SM	unpublished
FE	1958	1b × 7t	fallow	[50]
FFFE	1979	1b × 6t	(AL or CL)–WW–SM–WW–SBA–(SB or POT)–SBA	[51]
IOSDV	1983	1b × 4t	(SB or POT)–SBA–WBA	[52]
OaMNFE(dc)	2011	1b × 5t	POT–WW–SM–SBA–OSR–WW	[53]
OaMNFE(sf)	1965	1b × 6t	WW–POT–SBA–LCM–WW–POT–O–CL	[54]
RFE	1955	2b × 8t	SW–SB or AL–AL–WW–SB–SBA–POT–WW–SB–SBA	[55]

[a] CRE—Crop Rotation Experiment, CRT—Crop Rotation Trial, FE—Fallow Experiment, FFFE—Fraction Factorial Fertilization Experiment, IOSDV—International Long-Term Organic Nitrogen Nutrition Experiments, OaMNFE(dc)—Organic (digestate, compost) and Mineral N Fertilization Experiment, OaMNFE(sf)—Organic (straw, farmyard manure) and Mineral N Fertilization Experiment, RFE—Ruzyně Fertilizer Experiment; [b] The number of blocks and treatments per block at each site; [c] AL—alfalfa (*Medicago sativa* L.), CL—red clover (*Trifolium pratense* L.), LCM—legume–cereal mixture, O—oat (*Avena sativa* L.), OSR—winter oilseed rape (*Brassica napus* L.), PEA—cultivated pea (*Pisum sativum* subsp. *sativum* L.), POT—potato (*Solanum tuberosum* L.), SB—sugar beet (*Beta vulgaris* subsp. *vulgaris* L.), SBA—spring barley (*Hordeum vulgare* conv. *distichon* (L.) Alef.), SM—maize for silage (*Zea mays* subsp. *mays* L.), SW—spring wheat (*Triticum aestivum* L.), TR—triticale (× *Triticosecale* Wittm. ex A. Camus.), WBA—winter barley (*Hordeum vulgare* conv. *vulgare* L.), WW—winter wheat (*Triticum aestivum* L.).

The soil samples were dried, sieved through 2 mm mesh, and milled. MIR-DRIFTS spectra were measured using a Thermo Nicolet Avatar 320 FTIR spectrometer with a Ge beam splitter and a TGS detector, equipped with a Smart Diffuse Reflectance accessory (Nicolet, Madison, WI, USA) in a homogeneous mixture of 300 mg bulk soil and 900 mg FTIR grade KBr (Sigma-Aldrich, Darmstadt, Germany) prepared by hand in an agate mortar. Each sample was transferred to a 12 mm diameter diffuse reflectance cup and levelled with a microscope glass slide in a way to avoid compressing mechanically the mixture. Three scans comprising 1869 equidistant bands in the 4002–399 cm^{-1} wavenumber range were performed, each spectrum was corrected against pure KBr as a background spectrum, and the obtained apparent absorbance (hereafter, absorbance) values averaged to obtain a spectrum with reduced noise [35]. TC content was determined by dry combustion using Vario/CNS analyzer (Elementar Analysensysteme GmbH, Langenselbold, Germany), and hot-water extractable carbon (HWC) content was determined according to Körschens et al. [8] as a measure of SC lability [27,56].

2.2. Data Partitioning and Pre-Processing of MIR-DRIFTS Spectra

The collected data were subjected to a number of pre-processing and subsetting operations, the character of which was differentiated according to the study questions; depending on the scenario, one or more operations could also be omitted. PLSR models for predicting TC and HWC contents from MIR-DRIFTS spectra were then trained, tuned, and validated using the derived datasets. Figure 2 depicts the data processing workflow.

The samples in the library part of the dataset served as the calibration samples in the global (library only) modeling scenario (Figure 3a), equivalent to removal of the "raw non-test target pool"–"sample weighing by multiplication" workflow branch in Figure 2. For each commercial site, 10 independent sets of 12 samples were picked randomly for testing of predictive model quality. The target-site spectra not included in a testing partition made a pool from which samples were picked for model training in other scenarios (Figure 3b,c). The order of samples within these pools was randomized.
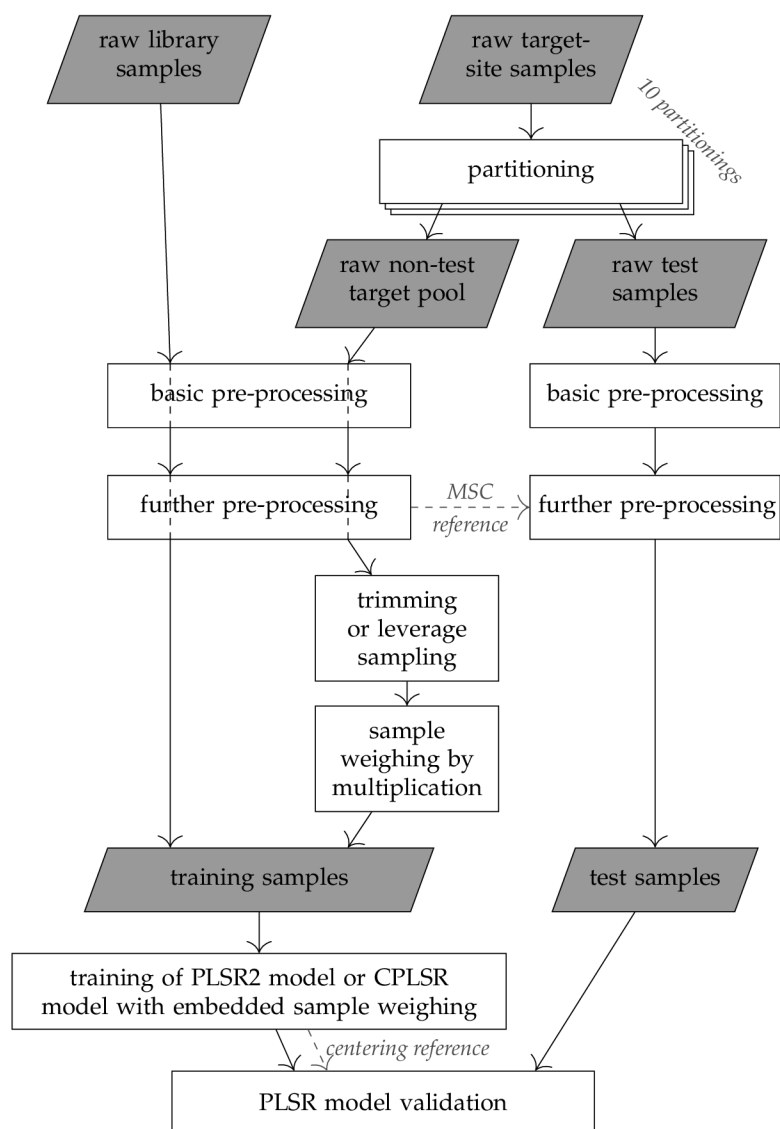
Figure 2. Data processing workflow.

Spectral pre-processing was performed before the selection of target-site training samples from the training pools. Noisy bands up to $600\,\mathrm{cm}^{-1}$ [17] and CO_2-affected measurements in the 2268–$2389\,\mathrm{cm}^{-1}$ wavenumber range [32] were discarded. For additional signal recovery, the spectra were processed using a moving-average filter with an 11-band window.

In addition to analyzing the resulting spectra, hereafter "raw spectra", we tested five further pre-processing schemes [57], with each scheme comprising two phases. In the first phase, the moving-average smoothing was either followed with multiplicative scatter correction (MSC) or left unchanged. In the second phase, (1) standard normal variate (SNV), (2) derivative transformation using the Savitzky–Golay filter with third-order polynomial smoothing applied over a moving window of 11 bands, or (3) no transformations were applied to the result. No change to the spectra in both phases was equivalent to removal of the "further pre-processing" box in Figure 2. Initially, continuum removal by dividing the spectrum by its convex hull was also attempted, but it had to be abandoned as extreme

outliers were generated. Unlike the remaining transformations, MSC employs information from multiple spectra to derive a common reference spectrum. We were careful to perform this operation using the data in the training spectra pools, exclusively [6,58].
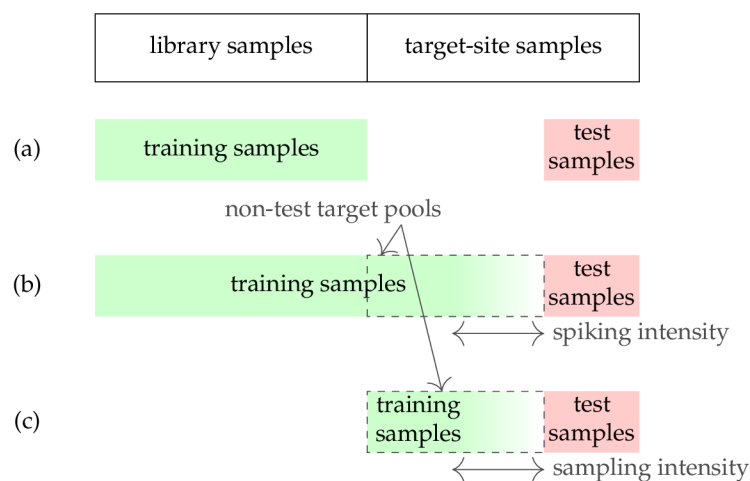
Figure 3. The possibilities of data subsetting compared in the study: (**a**) Library-only partitioning without calibration spiking. (**b**) The library data are augmented with target samples from a training pool, the number of which is given by the spiking intensity. (**c**) Local-only models trained exclusively to target-site samples, the number of which is given by the sampling intensity.

2.3. Calibration Spiking

Calibration spiking was introduced, based on increasing spiking sample counts to the level of 16 samples with a step of 4 samples (Figure 3b). The pre-randomized calibration sample pools were trimmed while preserving the sample orders. In addition to this random scheme, two leverage sampling approaches were assessed: the Kennard–Stone algorithm [59] and conditioned Latin hypercube [60]. The spectra were subjected to PCA prior to the Kennard–Stone algorithm application to reduce the number of dimensions below the sample pool size level.

In order to test for the possibility of a local modeling superiority with respect to models trained both to global and spiked datasets, additional scenarios mirroring the calibration-spiking scenarios but without samples from the long-term experiments were included (Figure 3c). This was equivalent to omitting the "library samples" branch in Figure 2. The training sample selection followed the same three schemes as for calibration spiking, with the same sampling intensity levels.

2.4. Reference Laboratory Data Pre-Processing

TC content cannot exceed a certain level of SC saturation [61,62], whereas HWC cannot be larger than TC. While applying statistical methods to measurements of sample constituents' concentrations, such as TC and HWC, it is recommended to follow the principles of compositional data analysis. Otherwise, models can yield nonphysical predictions, such as those of negative concentrations, a problem encountered by Baldock et al. [16] and Janik et al. [63], or component sizes the sum of which exceeds 100 %.

Classical statistical tools can be employed to compositional data after subjecting them to log-ratio transformations. Accordingly, three components summing up to the whole soil sample were derived from the TC and HWC measurements: (1) HWC, (2) the part of TC resistant to hot-water extraction (nHWC), and (3) the non-TC part of a sample $(1 - TC)$.

In the next step, the component values were transformed into two isometric log-ratio (ilr) coordinates according to the formulas [64]:

$$\text{ilr}_{\text{TC}} = \sqrt{\frac{2}{3}} \log \frac{(\text{HWC} \cdot \text{nHWC})^{\frac{1}{2}}}{1 - \text{TC}}, \tag{1}$$

$$\text{ilr}_{\text{HWC}} = \sqrt{\frac{1}{2}} \log \frac{\text{HWC}}{\text{nHWC}}. \tag{2}$$

The ilr_{TC} coordinate is closely related to TC but accounts for the finite size of a sample, while ilr_{HWC} can be interpreted as transformed C lability [65]. The latter formulation not only respects the compositional character of the reference data but also avoids confounding lability with TC, thus facilitating their independent analysis. This is unlike raw HWC, the value of which can be affected by both factors [9].

2.5. PLSR Modeling with Unweighed and Weighed Training Samples

The relationship between ilr values and MIR-DRIFTS spectral patterns was modeled using PLSR. Two multiresponse PLSR extensions were trained to both coordinates to account for multivariate character of compositional data [66]. For data partitionings that included both reference-library and target-area samples, the influence of spiking sample weighing was examined by introducing models with 5-fold and 25-fold weighted local observations, in addition to unweighted models. The weighing was performed either in the standard way by data row multiplication—in which case a PLSR2 model [42] was used—or by exploiting the internal weighing capability of the CPLSR model family [47] as a proposed approach. The latter case detoured the "sample weighing by multiplication" Figure 2 workflow step. Obviously, the weighing was restricted to the calibration-spiking scenarios, as the remainder, that is library-only and local-only scenarios, involved only single sources of samples.

Centered values of ilr coordinates were the dependent variables (responses) and centered MIR-DRIFTS intensity values were the independent variables (features) in these models. Like for MSC, the centering was based on information in the training data only. The numbers of PLSR components were tuned using leave-one-out cross-validation with values between 1 and 12 considered. The number of components to keep was determined using one standard error heuristics [67] applied separately to ilr_{TC} and ilr_{HWC} RMSECV. In this way, 12 240 bivariate models were calibrated and twice as many tuned models obtained.

The performance of each model was evaluated using test data partitions in terms of R^2, prediction bias, and RMSEP, followed with RPD_{P} and RPIQ_{P} statistics:

$$R^2 = \frac{V_{Res}(0) - \text{RMSEP}^2}{V_{Res}(0)}, \tag{3}$$

$$\text{bias} = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n}, \tag{4}$$

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}, \tag{5}$$

$$\text{RPD}_{\text{P}} = \frac{s_{\text{P}}}{\text{SEP}}, \tag{6}$$

$$\text{RPIQ}_{\text{P}} = \frac{\text{IQR}_{\text{P}}}{\text{SEP}}, \tag{7}$$

where $V_{Res}(0)$—mean square ground truth value, $\hat{y}_i$—predicted ith value, y_i—ith ground truth value, n—test sample count, s_{P}—standard deviation of ground-truth values, IQR_{P}—

interquartile range of ground-truth values, and SEP—standard error of prediction, which was defined as:

$$\text{SEP} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i - \text{bias})^2}{n-1}}. \tag{8}$$

These were summarized, and the relative influence of the experimental factors on the model performance measures was also examined visually after plotting the relationships.

2.6. Reproducing the Study

The analysis was coded using the R language and executed in the 3.6.2 version of the interpreter [68]. The package `vegan` (version 2.5.6) [69] was used for PCA, `prospectr` (0.1.3) [70] and `pls` (2.7.2) [71] for spectra pre-processing, `prospectr` and `clhs` (0.7.2) [72] for leverage sampling, `compositions` (1.40.3) [73] for ilr transformations, and `pls` for PLSR modeling. GNU Make [74] was employed for workflow control, and GNU Guix functional package management and containerization capabilities [75] were exploited to obtain reproducible results. The data and code are available from a Zenodo repository (doi:10.5281/zenodo.6337394). Reproduction of the study is going to require the availability of HPC infrastructure. It took approximately three weeks of operation of a 16-CPU virtual machine to complete a full computation cycle and obtain the results.

3. Results

3.1. Patterns in the Raw and Pre-Processed Data

Ústí nad Orlicí spectral signatures were highly varied and, in certain regions, extended beyond the envelope of the library samples regardless of pre-processing (Figures 4 and S1). The scans were subjected to PCA to obtain more insight into the spectral dissimilarity [39]. According to the first two principal component scores, there is substantial overlap between the reference library spectra and Ústí nad Orlicí soil samples, but a significant fraction of the observations occupy the area of the PCA space devoid of library data points due to high PC2 scores (Figure 5). As could be expected, the bulk of high-PC2 library observations represent experimental stations located close to the discussed district, namely Hněvčeves, Svitavy, Čáslav, and Kostelec nad Orlicí (Figure 1). Notable are the large ranges of Ústí nad Orlicí PCA scores, comparable to those of the long-term experiments. In contrast to that pattern, Janovice spectra were enveloped by the library spectra (Figure 4), and the data points form a compact cluster in Figure 5, similar in extent to several individual library sites, as shown using convex-hull polygons.

In addition, the C measurement variation was high in Ústí nad Orlicí and not much smaller than that of the library samples despite the different geographical scales (Table 2 and left-hand plot in Figure 6). Both TC and HWC are somewhat shifted upwards relative to the bulk of the reference library. Unlike the PCA scores, the mismatch between target-site C measurements and reference library measurements is more apparent for Janovice. Both TC and HWC are high here, and the only library samples with similar characteristics are a group of Praha-Ruzyně Fallow Experiment experimental plots. A closer examination revealed that those had been assigned to compost fertilization treatments.

Regardless of the data subset, the raw measurements were skewed towards lower values (left-hand plot in Figure 6). The skew, and to a degree high kurtosis, were reduced after the ilr transformations (right-hand plot in Figure 6 and Table S2). Figure 7 depicts the relationships between the raw component values and ilr coordinates. While the TC–ilr_{TC} relationship is smooth and close to linear, a broken stick pattern was obtained for HWC–ilr_{HWC}. The outlying samples with HWC in excess of $1.2 \, \text{mg} \, \text{g}^{-1}$ all came from Praha-Ruzyně Fallow Experiment plots where compost was applied. Although ilr_{TC} and ilr_{HWC} are not simple transformations of, respectively, TC and HWC, as additional components were accounted for in their derivation (Equations (1) and (2)), the relationships are strong enough to permit comparing our results with those reported by authors who had not considered the compositional nature of SC pools.
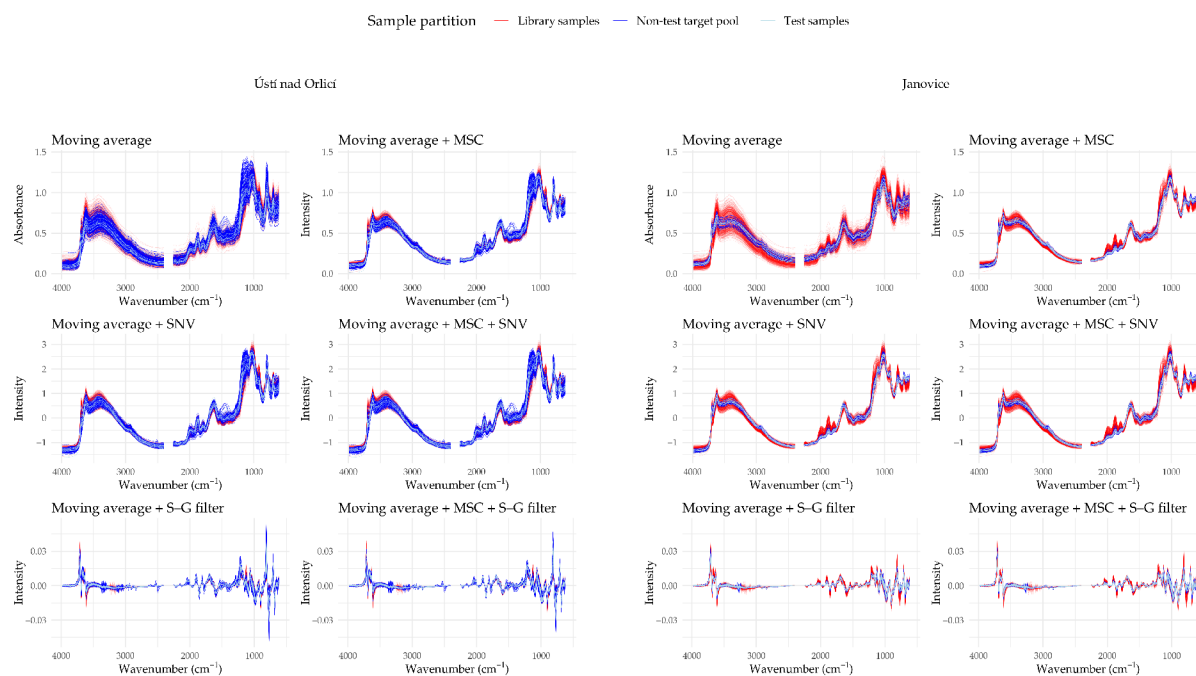
141

Figure 4. Pre-processed library and target-site spectra. The reference spectrum for the multiplicative scatter correction (MSC) transformation is based on the first training-pool–test partitioning.
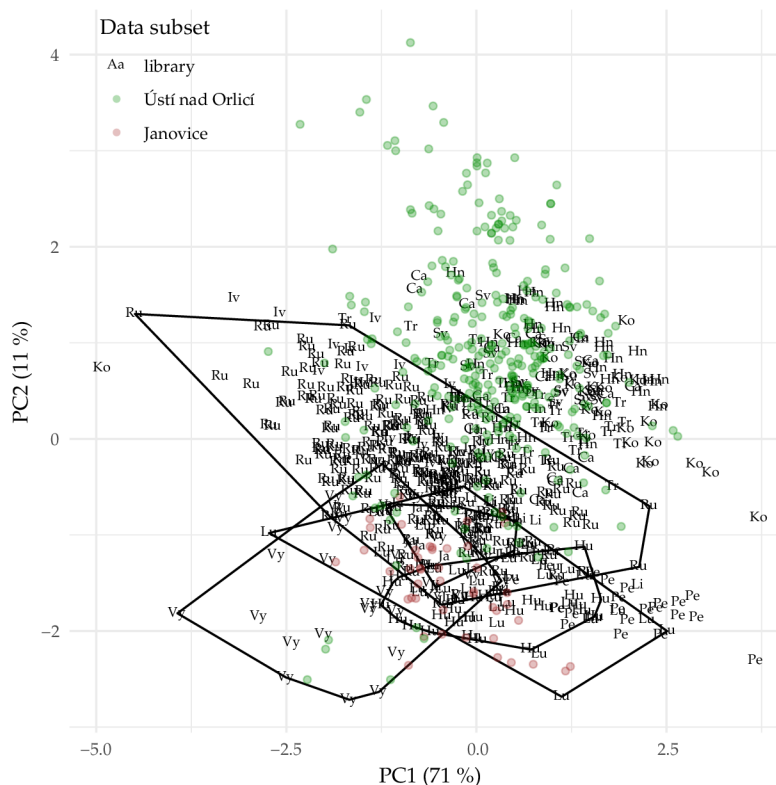
Figure 5. Projection of the principal component space derived from MIR-DRIFTS spectra after basic pre-processing. Convex hulls of the library sites similar to Janovice are displayed. The smallest, central polygon represents Jaroměřice. Ca—Čáslav, Hn—Hněvčeves, Hu—Humpolec, Iv—Ivanovice na Hané, Ja—Jaroměřice, Ko—Kostelec nad Orlicí, Li—Lípa, Lu—Lukavec, Pe—Pernolec, Ru—Praha-Ruzyně, Sv—Svitavy, Tr—Trutnov, Vy—Vysoké nad Jizerou.

Table 2. Location and scale statistics describing the distributions of soil carbon (SC) measurements before and after isometric log-ratio (ilr) transformations.

Statistics	Sample Partition	C Measurement			
		Raw		ilr-Transformed	
		TC	HWC	ilr_{TC}	ilr_{HWC}
range		(%)	(mg g^{-1})		
	library	0.73–4.45	0.13–2.55	−5.63−−3.70	−3.25−−1.97
	Ústí nad Orlicí	0.94–3.68	0.27–1.09	−5.22−−4.19	−2.83−−2.04
	Janovice	1.35–3.04	0.46–1.16	−4.89−−4.21	−2.47−−2.15
median		(%)	(mg g^{-1})		
	library	1.41	0.38	−4.98	−2.54
	Ústí nad Orlicí	1.65	0.51	−4.77	−2.43
	Janovice	2.10	0.77	−4.51	−2.31
IQR		(pp)	(mg g^{-1})		
	library	0.45	0.16	0.27	0.29
	Ústí nad Orlicí	0.46	0.17	0.23	0.17
	Janovice	0.32	0.16	0.14	0.09
n	library	603			
	Ústí nad Orlicí	335			
	Janovice	45			
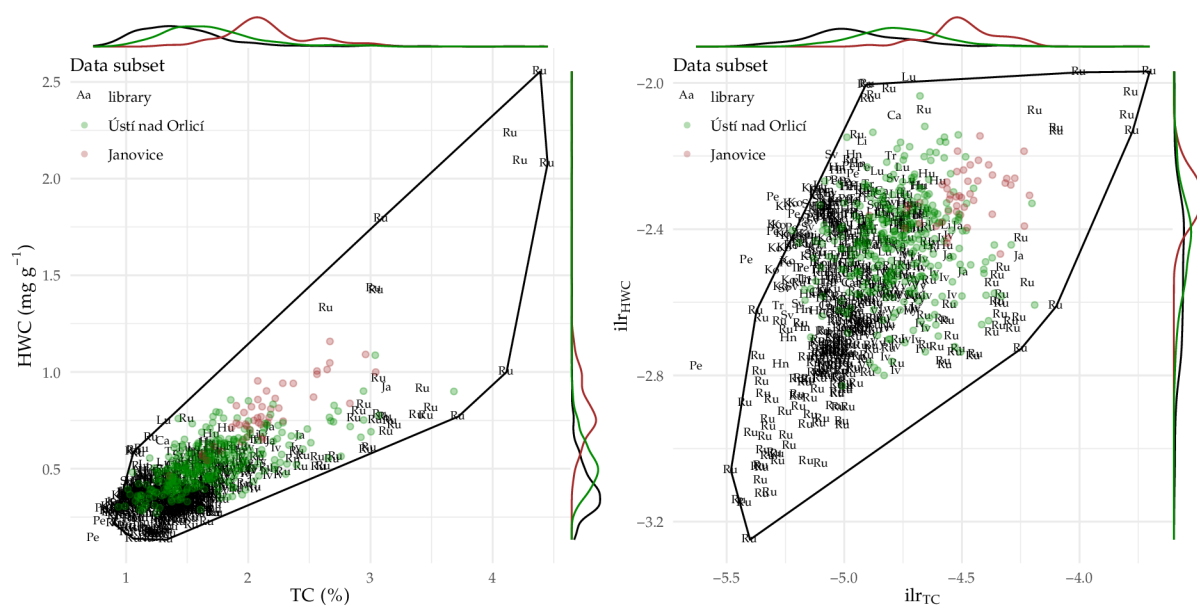
Figure 6. Joint and marginal distributions of SC raw and ilr-transformed measurements. The convex hull depicts the extent of Praha-Ruzyně observations. Ca—Čáslav, Hn—Hněvčeves, Hu—Humpolec, Iv—Ivanovice na Hané, Ja—Jaroměřice, Ko—Kostelec nad Orlicí, Li—Lípa, Lu—Lukavec, Pe—Pernolec, Ru—Praha-Ruzyně, Sv—Svitavy, Tr—Trutnov, Vy—Vysoké nad Jizerou.
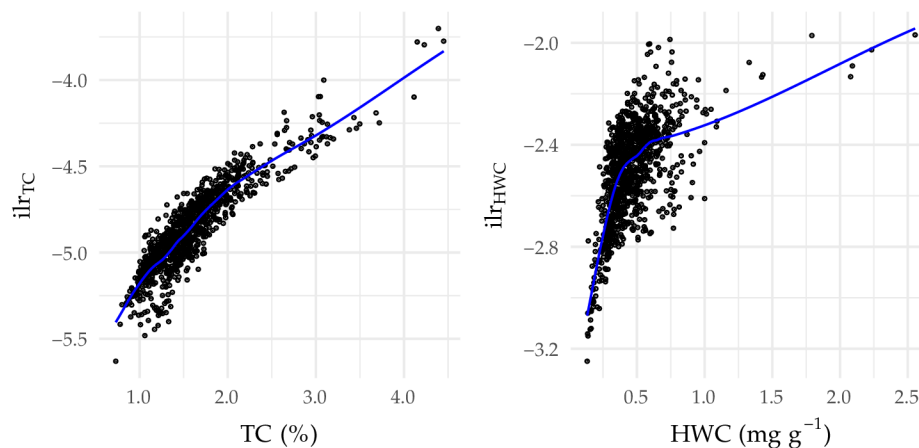
Figure 7. The relationships between raw SC reference measurements and ilr-transformed values, with overlaid loess smoothers.

3.2. Accuracy and Precision of the PLSR Models

The predictive performance of the PLSR models varied substantially, as illustrated by the R^2 statistics (Table 3). Although negative values were obtained for the worst models, models corresponding to R^2 in excess of 0.80 could be found for each ilr coordinate and target site combination, which is a high quality result according to Janik et al. [20]. However, after aggregating the values across all data partitionings, R^2 exceeded 0.50, still an unsatisfactory value, only for Janovice while predicting ilr_{TC}, whereas both ilr_{HWC} and Ústí nad Orlicí scenarios gave poor results.

Table 3. Ranges of PLSR model performance measures according to the dependent variable and the target site. The values outside and inside the brackets correspond to performances obtained for individual data partitionings and performances that were median aggregated across the partitionings, respectively.

Performance Measure	ilr_{TC}		ilr_{HWC}	
	Ústí nad Orlicí	Janovice	Ústí nad Orlicí	Janovice
R^2	−9.10 [−3.97, 0.33] 0.81	−18.79 [−8.76, 0.57] 0.88	−6.90 [−1.36, 0.18] 0.85	−37.43 [−18.98, 0.35] 0.82
bias	−0.42 [−0.30, 0.16] 0.32	−0.49 [−0.47, 0.07] 0.21	−0.19 [−0.13, 0.07] 0.14	−0.28 [−0.24, 0.04] 0.09
RMSEP	0.07 [0.13, 0.35] 0.51	0.04 [0.08, 0.48] 0.51	0.05 [0.11, 0.19] 0.24	0.03 [0.04, 0.26] 0.29
RPD	0.33 [0.47, 1.27] 2.42	0.23 [0.33, 1.60] 3.01	0.37 [0.68, 1.15] 2.73	0.17 [0.23, 1.29] 2.45
RPIQ	0.30 [0.62, 1.70] 3.09	0.13 [0.26, 1.45] 2.52	0.38 [0.69, 1.26] 2.84	0.18 [0.30, 1.59] 3.79

The worst negative biases and RMSEP values were comparable, amounting to 0.4–0.5 for ilr_{TC} and 0.2–0.3 for ilr_{HWC}. In terms of raw component values, these correspond to approximately 1.30 TC percentage points and 0.09–2.79 mg g^{-1} HWC, depending on the baseline HWC value (Figure 7). The best models had RMSEP of only 0.04 for ilr_{TC} (approximately 0.12 pp TC) and 0.03 for ilr_{HWC} (0.34 mg g^{-1} HWC for high value range and less for low value range). More conservative estimates, based on partitioning medians, suggested a possibility of predicting ilr_{TC} with an error of 0.13 (0.38 pp TC) and 0.08 (0.23 pp TC) in Ústí nad Orlicí and Janovice, respectively, while for ilr_{HWC}, the corresponding values were 0.11 and 0.04 (0.04–1.23 and 0.01–0.45 mg g^{-1} HWC).

Models with RPD_P or $RPIQ_P$ above 2.5 or even 3.0 were obtained in some scenarios and test data partitions, described in literature as good and excellent predictions [76]. However, typically one should not expect the performance to be higher than 1.7, that is, barely sufficient to estimate the values even as high or low. Unlike for the other measures, Janovice models did not yield consistently superior RPD_P and $RPIQ_P$ relative to Ústí nad Orlicí.

There is an agreement between PLSR regression coefficients of the best Janovice models for predicting ilr_{TC} regardless of the performance measure in which a model excelled (Figure 8). The pattern is similar to that presented for Baldock et al. [16] square-root transformed TC model, including the presence of aliphatic $C-H$ (at approximately $2890\,cm^{-1}$), $C=O$ ($1740\,cm^{-1}$), and negative carbonate ($1810\,cm^{-1}$) peaks. In contrast, the coefficients for Ústí nad Orlicí disagree and the pattern is malformed, which may suggest model overfitting. Regression coefficient values are comparable among two of the best-performing Janovice ilr_{HWC} models. Their patterns do not resemble those published by Zimmermann et al. [17] for labile OC, but these authors modeled raw component sizes, rather than lability, and presented individual PLSR loadings, rather than regression coefficients. There is a major negative peak in the $3700-3600\,cm^{-1}$ wavenumber range, which corresponds to $O-H$ stretching of clay minerals [77,78]. Other peaks occur at approximately $1000\,cm^{-1}$ and below. Here, notable is the positive $1050\,cm^{-1}$ peak, assigned to quartz reflectance [19]. However, according to Nocita et al. [28], the interpretation for the $<1000\,cm^{-1}$ region is challenging due to mineral species vibrations interfering with those of organic molecules. These include iron compounds [13] and carbonates [79]. The peaks do not include $2930\,cm^{-1}$ and $1620\,cm^{-1}$ wavenumbers proposed by Demyan et al. [80] for lability assessment. The model minimizing bias behaved differently, and for Ústí nad Orlicí, the smallest-bias model happened to be insensitive to input data variation, which indicates that models should not be selected according to the bias criterion. As with ilr_{TC}, the pattern is unstable for this latter target site.

3.3. Factors Affecting PLSR Model Performance

The relationships between the modeling approaches and performance measure values were visualized to identify factors contributing to prediction quality. We present a selection that illustrates the most clear patterns, which, with the exception of the final comparison, is restricted to the models trained to the raw spectra, as the effect of spectra pre-processing was limited. The complete set of visualizations along with input data points can be found in Figure S2.

PLSR models trained to the spectral library, that is, with zero target-site samples, performed poorly, especially for Janovice, as can be seen at the left edge of all plots in Figure 9. Note that this and subsequent figures for legibility depict confidence intervals, whereas ranges are referred to in this section. The R^2 statistic was negative with the exception of Ústí nad Orlicí ilr_{HWC} models, in the case of which it ranged between -6.24 and 0.47. The generated predictions were negatively biased, while their imprecision measured by RMSEP exceeded 0.17 units for ilr_{TC} (about 0.49 pp TC) and 0.08 units for ilr_{HWC} ($0.03-0.89\,mg\,g^{-1}$ HWC).

Training of PLSR models to a selection of target-site samples only, while excluding the spectral library, had a clearly positive effect on all measures even with only four training samples, as illustrated by the black lines in Figure 9. However, R^2 was still negative at this sampling intensity level. Here, predictions for Janovice appear superior to those obtained for Ústí nad Orlicí, especially in terms of RMSEP. Further additions of samples led to more accurate ilr_{TC} predictions in Janovice, as depicted in more detail in Figure 10. In particular, R^2 exhibited an increasing trend, with positive values up to 0.88, obtained in a number of scenarios with 16 samples. Prediction improvement of ilr_{TC} with higher sampling intensity is not so clear for Ústí nad Orlicí. Instead, a pattern of Kennard–Stone leverage sampling inferiority could be discerned, especially in terms of high bias, up to 0.32 units (0.92 pp TC).
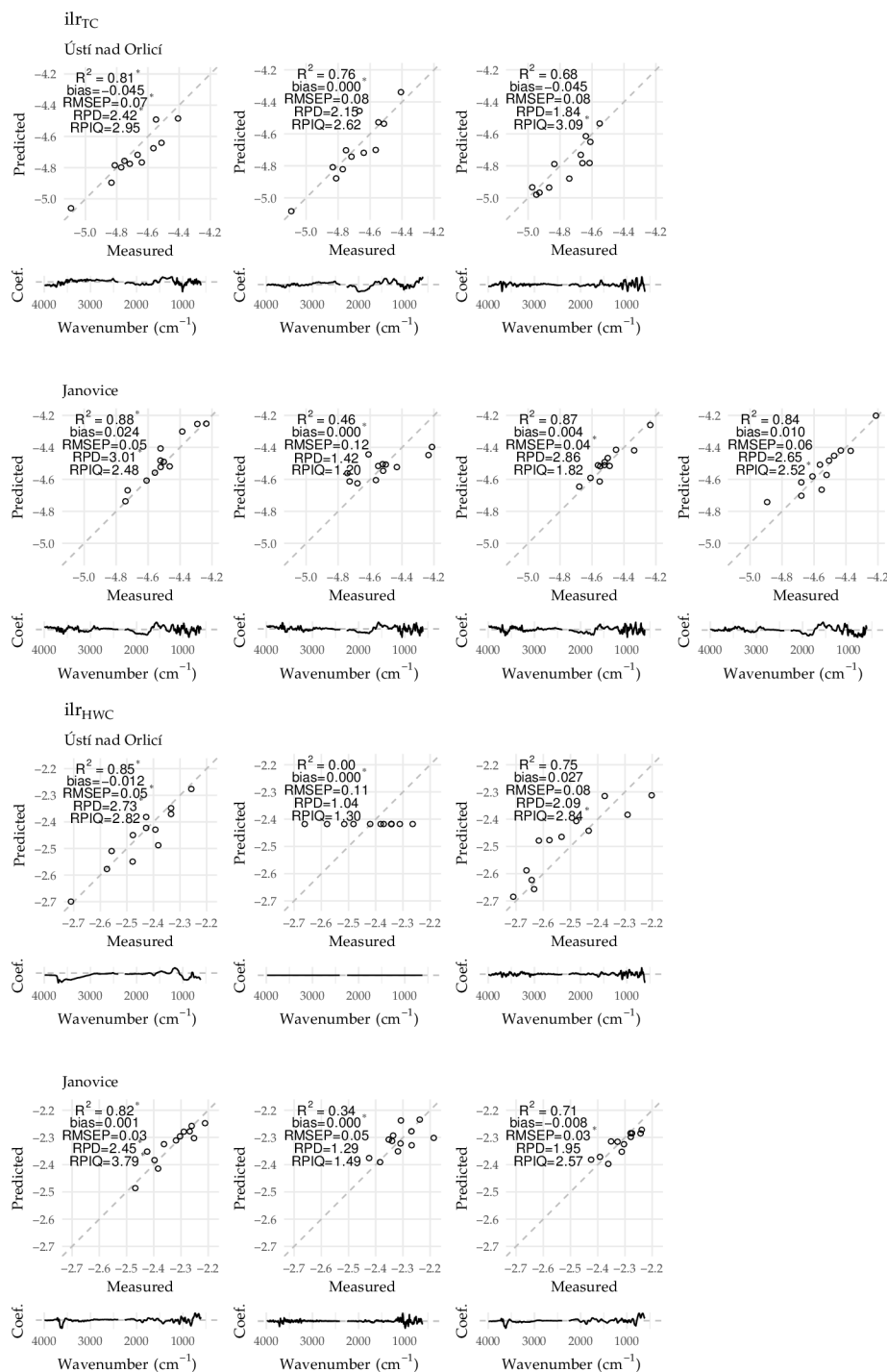
146

Figure 8. SC predictions corresponding to the top-performing models. The performance measures according to which individual formulations performed best are marked with asterisks. PLSR regression coefficients are shown for each model on a relative scale due to the coefficient ranges differing by orders of magnitude between the models.
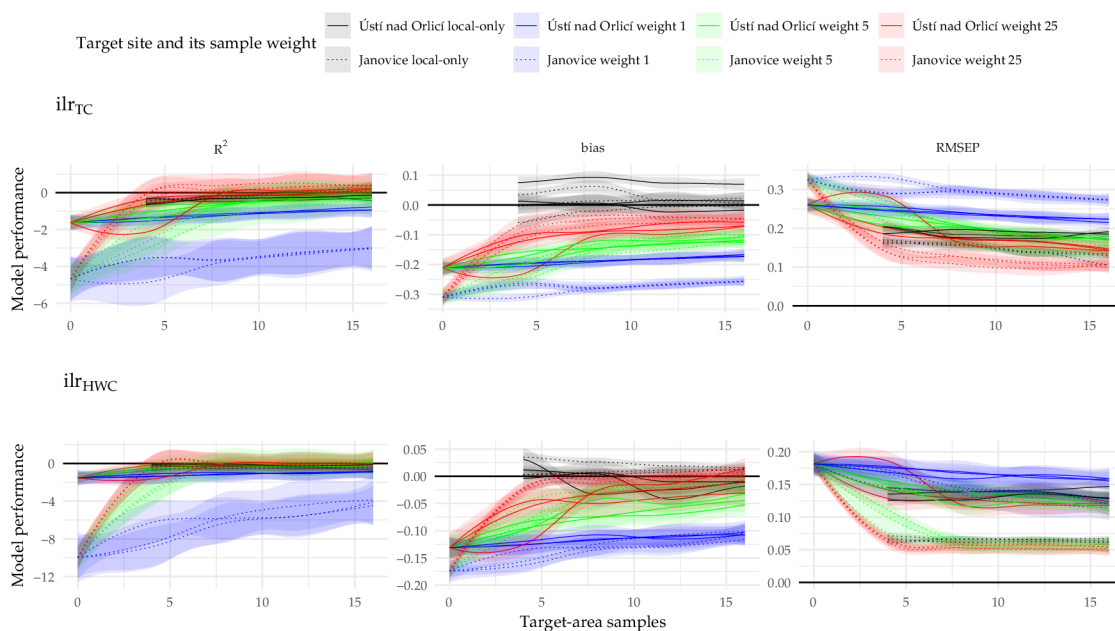
Figure 9. The influence of calibration spiking, weighing of the spiking samples, and removal of library spectra from the training dataset on partial least squares 2 regression (PLSR2) model performances. Only scenarios with basic and no further spectra pre-processing are included. Each line represents one combination of levels of the remaining experimental variables: leverage sampling strategy and predictive model family. A mean across 10 test datasets is drawn along with its 95% confidence interval.

148
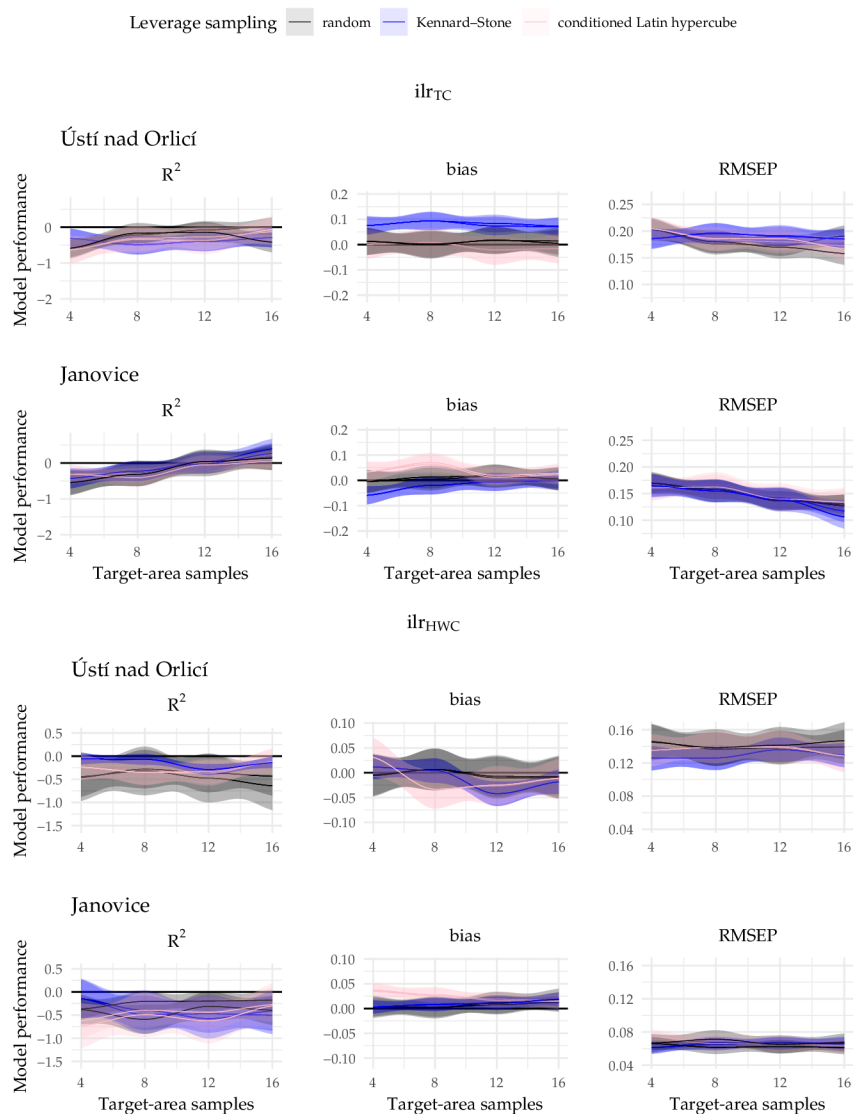
Figure 10. The influence of sampling intensity and leverage sampling on predictive local-only model performances. Only scenarios with basic and no further spectra pre-processing are included. Each line represents an ensemble of either partial least squares 2 regression (PLSR2) models or canonical partial least squares regression (CPLSR) models with the same level of spike sample weights. A mean across 10 test datasets is drawn along with its 95% confidence interval.

RMSEP of ilr_{HWC} was hardly affected by increasing sampling intensity. On the other hand, a trend towards increased bias can be discerned for Janovice under the random sampling and Kennard–Stone leverage sampling scenarios, but these strategies still do not appear consistently inferior to conditioned Latin hypercube. Positive R^2 was attained by few and apparently random Janovice models and almost no Ústí nad Orlicí models even at maximum sampling intensity, suggesting a general unsuitability of the local approach to estimating this ilr coordinate.

In Janovice scenarios with PLSR2 models, augmenting the library samples with spike samples yielded results competitive with the local approach when the target-site training

Wiktor R. Żelazny and Tomáš Šimon: Calibration spiking of MIR-DRIFTS soil...

Agriculture **2022**, *12*, 682   17 of 26

samples were given a weight of 25, as shown using red lines in Figure 9. R^2 up to 0.71 could be attained with only four spiking samples for ilr_{TC}—in contrast to R^2 of corresponding local-only models, which was always negative. A notable exception was prediction bias, in the case of which about 85 % of the models still underestimated the value of this coordinate. Models with the weight of five (green lines) were competitive with local-only models only in predicting ilr_{HWC} and only in terms of R^2 and RMSEP. More spiking samples were required to obtain a desirable effect than with 25-fold spiking sample weighing. The superiority of global Ústí nad Orlicí models relative to Janovice vanished or became inversed as spike samples were added to training datasets. The performance remained better only in scenarios without spike sample weighing (blue lines), but here the prediction quality was poor for both target sites, making this class of scenarios not interesting.

Leverage sampling had little effect on the quality of models that involved spiked library spectra, but the performance measures responded to the choice between PLSR2 and CPLSR family (Figure S3). The application of the CPLSR method was clearly detrimental for the prediction quality of both ilr_{TC} and ilr_{HWC} in Janovice samples compared to the standard approach. In the case of Ústí nad Orlicí, the effect of replacing PLSR2 with CPLSR was not so strong, but it still appears negative. The limited sensitivity of model performance to spectra pre-processing can be illustrated by two favorable combinations of spectra selection and weighing strategies. As depicted in Figure S3, systematic prediction quality differences are hard to discern except for the uninteresting library-only scenario, where all models failed.

4. Discussion

4.1. Distributional Data Properties and the Effect of Log-Ratio Transformation

The high scatter of observations in PCA (Figure 5) and SC (Figure 6) measurements, comparable in extents to those of long-term experiments, indicates high spatial heterogeneity of Ústí nad Orlicí district soils. This pattern corroborates the need for dense soil sampling to map and monitor SC in the conditions of the Czech Republic and, arguably, beyond [3,4], from which the need to develop cost-effective assessment methods follows [4]. However, in addition to the variability of soil properties, non-uniform sampling techniques might have also been a contributing factor, as unlike in the remaining campaigns, the task was relegated to farmers. In contrast to that, the relative compactness of the Janovice PCA cluster corresponds to the fact that the data collection was constrained to a single field. The high TC and HWC contents encountered at this locality might have been related to long-term organic fertilization of this field.

High performance of a PLSR model can be attained when the predicted variable has a Gaussian distribution, and in chemometric studies, it is common to transform target measurements [13]. Stenberg et al. [6] highlighted skewness of organic matter concentrations in cropland soil samples towards low values, a common pattern that can contribute to prediction bias [34]. Normalization of such data can be attained by applying a square-root [16,20,39] or a logarithmic [81–83] transformation. However, while these bound the predictions to be above zero [16], the maximum values remain unbounded.

A log-ratio affects the shape of data distribution like the above transformations, but in addition to that, back-transformed predictions correspond to physical reality for compositional components [64]. The present study demonstrates improved skewness and kurtosis of ilr coordinates relative to raw component concentrations (Figure 6) and provides evidence of compatibility of log-ratios with PLSR predictive modeling. The proposed data analysis approach could be refined in the future by accounting for carbon saturation limits [61,62] in the ilr transformation. Another potential extension would be to consider also the spectral measurements as compositional [84].

4.2. Absolute Performance of the Predictive Models

The top R^2 conservative estimate of only 0.57 when predicting ilr_{TC} and low RPD_P and $RPIQ_P$ evaluations (Table 3) do not corroborate the purported potential of MIR-DRIFTS

to become a cost-effective yet reliable laboratory method for SC assessment [13,25,35]. The agreement between the PLSR regression coefficient patterns obtained in the present study (Figure 8) and reported in literature [16,33,81] rules out major errors during both reference data collection and sample scanning and subsequent data analysis. Barra et al. [22] and Bellon-Maurel and McBratney [26] summarized model quality estimates for predicting OC and TC from MIR spectra. Although high-performing models prevail in reported research, a number of SC studies suffer from methodological issues that arguably bias the results towards higher accuracy. For example, Zimmermann et al. [17] employed a systematic rather than random validation sample and, moreover, included the validation data in PLSR model tuning dataset. More recently, Zhang et al. [18] erroneously [23] considered optimistic bias of model cross-validation results as an advantage and did not present the obtained independent validation statistics. It can be presumed that the models performed not so satisfactorily on the test datasets. Deiss et al. [31] contrasted the performance of PLSR and support vector machine models to predict OC in soil samples from two sites. Despite testing multiple combinations of spectral pre-processing and modeling scenarios, the authors presented only the performance measures of their best models. Those happened to be comparable to our top-rated results. In addition, their selection was based on full-validation statistics, which draws an over-optimistic picture of MIR-DRIFTS potential for real-life applications, where only few or even no validation samples would be available.

Methodological issues aside, not all models have been reported to perform well. The Bellon-Maurel and McBratney [26] review includes formulations that resulted in modest RPD_P values, similar to those obtained in the present study. In the more recent Page et al. [10] work, MIR-DRIFTS substantially underestimated OC loss over time in a long-term experiment, similar to our negative ilr_{TC} biases. Moreover, the estimated effect of evaluated management treatments contradicted that which was inferred using traditional OC determination. Calderón et al. [85] predicted OC in several crop experiments using PLSR and obtained RMSEP of 0.67–0.80 pp; that is beyond our upper RMSEP conservative bracket for TC. More research, preferably based on cooperation between multiple spectroscopy laboratories, is needed to determine to what degree different prediction performance results across studies can be attributed to the training samples at hand [29], sample preparation and scanning process differences [13,25,29], reference laboratory effect [13,29], or predictive model family and calibration workflow [13,25].

The fragility of MIR-DRIFTS to assess SC is further illustrated by C lability prediction performance. The negative 3650–3600 cm^{-1} and positive 1050 cm^{-1} Janovice PLSR regression coefficient peaks (Figure 8) can be related to the protective function of clay minerals with respect to soil organic matter [7,62]. However, with the majority of the remaining major peaks located in the <1000 cm^{-1} region, the predictions are prone to noise introduced by variation in soil mineralogy [28]. Also in the area of lability assessment, studies with over-optimistic results can be found. Our best ilr_{HWC} calibrations performed similarly in terms of R^2 and RPD_P to the PLSR models developed by Zhang et al. [27] for predicting raw HWC. Like Deiss et al. [31], these authors presented only their top-performing models for each investigated scenario, and in addition to that, they did not employ an independent test dataset, reporting only cross-validation statistics. Yang et al. [86] adopted a similar approach for the prediction of particulate organic carbon (POC), with comparable outcomes. Zimmermann et al. [17] attempted to predict two labile pools and reported RPD_P of only 2.0 for dissolved OC. Although the correlation between predicted and measured values was satisfactory and particulate organic matter was predicted with high accuracy, there was an information leak from the validation dataset while training of their models. A similar error was made by Calderón et al. [85] while tuning PLSR models for permanganate oxidizable carbon (POXC) predictions in a study that reported a high R^2 of about 0.8.

One factor contributing to prediction performance deterioration of all of the present study's models was probably the noise introduced to the spectra by grinding the soil samples by hand. Stumpe et al. [87] demonstrated that long grinding can reduce undesirable MIR spectra random variability. However, uniform grinding, a condition not attainable

with a manual operation, turned out to be even more important for OC prediction quality. The importance of controlled grinding in a MIR spectroscopy workflow is acknowledged also by other authors [13,16,33]. Particle size differences, a problem related to soil sample grinding [26,87], generate undesirable baseline shifts [88]. Many workers [80,85,86], including those reporting highly accurate predictions [16,27,32,63], routinely apply baseline correction to their measurements. Although we tested several combinations of spectral pre-processing workflows, this step was not included in the present study, which might have contributed to scanning artifacts remaining in the data. However, methods such as MSC and Savitzky–Golay derivation also address baseline variations [58], yet we were unable to associate them with systematic prediction improvement (Figure S4). According to Du and Zhou [24], moving average can diminish information in absorbance features, so perhaps we should have avoided it as a routine pre-processing step to remove noise.

The attempt to predict total, rather than organic, C probably also impaired the obtained results. In addition to OC, TC includes carbonates as a major C source, which have a different spectral profile, potentially interfering with the OC signal [17,25,88]. In the present study, the Praha-Ruzyně is a site with moderate carbonate content. Although average topsoil pH does not exceed seven, carbonates are visible by eye in a deeper soil layer. Moreover, the locality included experimental plots with compost amendments, which were associated with atypical C patterns (Figure 6). A compost fertilization experiment disrupted PLSR prediction quality also in the Calderón et al. [85] study. The authors reported an improvement after removing the problematic site from the dataset, and it is possible that a similar effect would be obtained in the present study. Perhaps, with OC being modeled instead of TC, PLSR regression coefficient peaks would have avoided the $<1000\,\mathrm{cm}^{-1}$ region, hypothesized to interfere with $\mathrm{ilr_{HWC}}$ predictions (Figure 8).

Some errors might have been related to insufficient sample dilution with KBr [89], especially for Ústí nad Orlicí spectra, which lied outside of the long-term experiments envelope primarily in the high-absorbance zone (Figure 4). This region coincided with the 1280–1070 cm^{-1} wavelength range associated with the silicate inversion feature that can interfere with carbonates signal below a certain dilution level [88]. However, Demyan et al. [80] did not confirm this effect and, instead, associated strong dilutions with the absence of certain absorption features. The traditional view on the need to mix soil samples with KBr for MIR-DRIFTS has been put into question also by Reeves III [25], and according to Tinti et al. [89] and Reeves [90], it can even have a negative effect. Perhaps, then, it would have been preferable to use neat samples in the present experiment.

Inferior $\mathrm{ilr_{HWC}}$ fit relative to $\mathrm{ilr_{TC}}$ might have been related to low HWC concentrations in the soil samples. Measurements of such minute pools tend to be more affected by external conditions than those of major components [17,27]. Although HWC appears in both ilr formulas, one can argue that a ratio, as employed for $\mathrm{ilr_{HWC}}$ (Equation (2)), is more sensitive to error than a geometric mean in the $\mathrm{ilr_{TC}}$ formula (Equation (1)).

4.3. Model Performance with Individual Training Data Subsets

In addition to the Praha-Ruzyně issue, the obtained poor performance of global scenarios can be attributed to the calibration domain mismatch between the library samples collected from long-term experiments and those collected at the target sites (Figures 5 and 6). Especially in the case of Janovice, notable are the high TC and HWC contents, which explain the strong negative bias in the predictions [26]. The negative influence of OC mismatch across datasets on its predictions was demonstrated by Seidel et al. [30] with VisNIR and by Guerrero et al. [39] with NIR spectroscopy.

The reference spectra in the experiment comprise long time series of observations but represent a limited number of locations. Similarly, Zhang et al. [27] obtained their samples from a limited number of long-term experiments, and their reported results are similar to ours. Various authors stress an importance of long-term experiments for studying SC, especially in the context of the low rates of its quantitative changes [3,5,11]. Nevertheless, maximizing the geographical extent of the reference data should apparently

be prioritized for predicting a factor with a high spatial variability, as it is the case for SC and its fractions [3–5]. A number of studies that adopted this strategy [16,20,33,35,82] demonstrate that high-quality predictive models can be developed in this way.

These issues do not apply to the local-only models, which do not involve any library spectra and a possibility of calibration domain mismatch is largely eliminated. Superior predictions characterizing locally calibrated PLSR models in the present study can be in part linked to the absence of Praha-Ruzyně samples in the training dataset, analogously to the effect observed by Calderón et al. [85] after training a model without an atypical site found in their data. This strategy largely removed ilr_{TC} prediction bias in our study (Figure 9), corroborating the calibration domain mismatch problem related to the reference library. However, the model quality was still unsatisfactory, especially for ilr_{HWC}, perhaps due to the limited sizes of the training data. The importance of a sufficient sample size was demonstrated by Guerrero et al. [39] in a NIR study and by Brown [91] in a VisNIR study, where the obtained performance approached that of calibration-spiking models only when large numbers of training samples were available. The costs and uncertain results involved in such a scenario make the advantage of spectroscopic estimation over standard oxidation methods questionable. According to Soriano-Disla et al. [13], local models are particularly suitable at small spatial scales with homogeneous sites. This condition may explain why the predictions for Janovice were superior and responded better to sampling intensity increase (Figure 10) relative to Ústí nad Orlicí. In particular, it might have been related to the smaller range of C measurements from this more homogeneous target site. After accounting for this effect, the prediction quality superiority was not apparent, anymore, as illustrated by the RPD_P and $RPIQ_P$ statistics.

Calibration spiking avoids an excessive reduction of training dataset sizes, and some of the best models in the present study could be associated with this strategy. A generally consistent positive relationship between the sampling and spiking intensity and PLSR model performance was obtained across the scenarios. It is similar to the OC prediction pattern with NIR spectroscopy obtained by Guerrero et al. [39] while increasing the number of target samples from 8 to 16 and 32. Analogously to the ilr_{HWC} pattern in the present study, Janik et al. [20] reported improved POC prediction quality with both calibration-spiking and local post-hoc models relative to unsatisfactory library-only predictions. The weaker effect of spiking on the performance of Ústí nad Orlicí models than for Janovice can, again, be explained by the high spectral variation of the geographically scattered samples, a situation described by Cezar et al. [36] in an experiment with ASD Fieldspec measurements.

An interest in calibration spiking is motivated by economical and environmental reasons [36]. Accordingly, satisfactory results should be expected even with a modest number of spiking samples [38]. The prediction improvement equivalent to maximizing the spiking intensity, but obtained by mere introduction of additional copies of the target-site data points, as observed for Janovice, is encouraging in this regard. It is also in line with our hypothesis on the potential of calibration spiking to reduce the number of samples for which laboratory reference data need to be obtained. Similarly, Guerrero et al. [39] reported that, for some target sites and a baseline spiking intensity of 8 samples, 25-fold weighing had a stronger positive effect on OC prediction quality than increasing the spiking sample number to 16 or 32. Perhaps further improvement would have been obtained with even heavier weights. However, Stork and Kowalski [40] tested weights up to 70 and determined an optimal number of spike sample copies as 24 in one scenario and less in the remainder, according to the Hotelling's T^2 statistics. In recent years, possibilities of predicting SC from MIR spectra collected in field rather than laboratory conditions without sample pretreatment have been explored [92]. Studies are needed to find out whether the positive influence of calibration spiking replicates in this more challenging setting.

4.4. The Effect of Leverage Sampling and Evidence against the CPLSR Internal Weighing Superiority Hypothesis

Clairotte et al. [33] and D'acqui et al. [81] reported OC prediction improvement with MIR-DRIFTS spectroscopy when leverage sampling was employed. In the present study, no apparent systematic differences were obtained with respect to the prediction performance among the random spiking and the spiking spectra selection based on conditioned Latin hypercube. The Kennard–Stone algorithm, on the other hand, was associated with biased ilr_{TC} predictions in Ústí nad Orlicí scenarios. This leverage sampling scheme tends to pick distant observations, located at the edges of a hyperspace (Figure S5). It also operates incrementally, as opposed to conditioned Latin hypercube, in the case of which the spectra are picked at once and can be more representative of a dataset [83]. Kennard–Stone application to the heterogeneous Ústí nad Orlicí dataset might have yielded outlier spiking samples, perhaps corresponding to soils with atypical textures [87] or mineralogy [85]. Ng et al. [83] obtained unstable calibrations involving this scheme except for large training samples. This apparent unreliability of the Kennard–Stone algorithm for small sample sizes relative to the size and heterogeneity of a target area puts in question its usability in campaigns aimed at minimizing reference data collection effort to obtain cost-effective predictions.

Internal weighing capability of the CPLSR extension of PLSR [47] was tested as an alternative to the spiking set augmentation by data point copies. Contrary to our hypothesis, the obtained models performed poorly, especially for Janovice. Sankey et al. [41] attempted to predict SC from VisNIR spectral data using boosted regression trees for different levels of local sample weights relative to the weights of the samples in the reference library. The authors expressed skepticism with respect to their results, in which the model performance decreased substantially for one target site, and while a positive relationship was observed for another, the obtained improvement was modest. Still, given the limited number of studies devoted to the topic so far, it seems worthwhile to further explore effects of embedded weighing with other data and other classes of predictive models [31,63].

5. Conclusions

Log-ratio transformation of laboratory reference measurements is recommended to avoid non-physical predictions, separate confounding factors, and improve data distributional properties. Accounting for carbon saturation limits and treating spectral measurements as compositional are potential further refinements of this approach.

Conservative estimates of PLSR model performances were lower than the values typically reported for MIR-DRIFTS SC predictions. This discrepancy could be attributed to the noise in the data introduced by manual sample grinding, their inadequate dilution with KBr, presence of an atypical site with carbonate soil and compost fertilization in the spectral library, the library's insufficient geographical coverage, and calibration domain mismatch relative to the validation samples. It was also in part explained by optimistic bias encountered in the literature due to preference of cross-validation over independent model validation, information leaks from training to testing datasets, and presenting only top-performing validated models by certain authors. There is a need for international cooperation to identify leverage points that could improve reliability of MIR-DRIFTS SC assessments, standardize data collection and treatment workflows, harmonize spectral libraries, and facilitate their use.

Target-site comparison revealed differences in sample heterogeneity related to uneven geographical extents and, possibly, varied soil sampling protocols where farmers were involved. Not enough representative training data were available to satisfactorily predict soil C properties in the more geographically extensive district-scale dataset. Here, spectral and reference laboratory data variation was similar to that of the data from more scattered long-term experiments, corroborating a need for a dense sampling grid to monitor soil C and concerns about potential costs involved.

Predicting soil properties at a field scale removed the issues related to the reference library. Although some models performed very well, the quality was unstable with respect

to the choice of validation data even with an application of leverage selection algorithms. C lability predictions were especially fragile, presumably due to the small size of the hot-water extractable pool. The quality of field-scale models responded positively to increasing sampling intensity in local-only scenarios, but further additions of samples in an attempt to obtain more representative training datasets would have been incompatible with the aim of reducing reference laboratory analysis expenses.

Calibration spiking combined with PLSR2 modeling was associated with a steep increase of model quality as additional target-site calibration samples were added, especially in combination with heavy weighing. It, therefore, appears to be a promising cost-effective and environmentally friendly SC monitoring solution but only under the assumption that the available spectral library accounts to a sufficient degree for soil variability. A similar effect could not be obtained with CPLSR models and embedded weighing enabled by this PLSR extension. Although prediction performance was poor in the present study, the internal weighing approach may still be worth testing with other multivariate model families. A training-sample size constraint was encountered while applying Kennard–Stone leverage sampling to the heterogeneous district-scale dataset, and it appears that application of this algorithm is not compatible with the aim of reducing costs of SC assessments.

References

1. Reeves, D. The role of soil organic matter in maintaining soil quality in continuous cropping systems. *Soil Tillage Res.* **1997**, *43*, 131–167. [CrossRef]
2. Bünemann, E.K.; Bongiorno, G.; Bai, Z.; Creamer, R.E.; De Deyn, G.; De Goede, R.; Fleskens, L.; Geissen, V.; Kuyper, T.W.; Mäder, P.; et al. Soil quality—A critical review. *Soil Biol. Biochem.* **2018**, *120*, 105–125. [CrossRef]
3. Smith, P.; Soussana, J.F.; Angers, D.; Schipper, L.; Chenu, C.; Rasse, D.P.; Batjes, N.H.; Van Egmond, F.; McNeill, S.; Kuhnert, M.; et al. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Glob. Chang. Biol.* **2019**, *26*, 219–241. [CrossRef] [PubMed]
4. Paustian, K.; Collier, S.; Baldock, J.; Burgess, R.; Creque, J.; DeLonge, M.; Dungait, J.; Ellert, B.; Frank, S.; Goddard, T.; et al. Quantifying carbon for agricultural soil management: From the current status toward a global soil information system. *Carbon Manag.* **2019**, *10*, 567–587. [CrossRef]
5. Batjes, N.H.; Van Wesemael, B. Chapter Measuring and Monitoring Soil Carbon. In *Soil Carbon: Science, Management and Policy for Multiple Benefits*; Banwart, S.A., Noellemeyer, E., Milne, E., Eds.; CAB International: Wallingford, UK, 2015; Volume 71, pp. 188–201.
6. Stenberg, B.; Rossel, R.A.V.; Mouazen, A.M.; Wetterlind, J. Visible and Near Infrared Spectroscopy in Soil Science. In *Advances in Agronomy*; Elsevier: Amsterdam, The Netherlands, 2010; Volume 107, pp. 163–215. [CrossRef]
7. Kan, Z.R.; Liu, W.X.; Liu, W.S.; Lal, R.; Dang, Y.P.; Zhao, X.; Zhang, H.L. Mechanisms of soil organic carbon stability and its response to no-till: A global synthesis and perspective. *Glob. Chang. Biol.* **2021**, *28*, 693–710. [CrossRef]
8. Körschens, M.; Schulz, E.; Behm, R. Heißwasserlöslicher C und N im Boden als Kriterium für das N-Nachlieferungsvermögen. *Zentralblatt Für Mikrobiol.* **1990**, *145*, 305–311. [CrossRef]
9. Thomas, B.W.; Whalen, J.K.; Sharifi, M.; Chantigny, M.; Zebarth, B.J. Labile organic matter fractions as early-season nitrogen supply indicators in manure-amended soils. *J. Plant Nutr. Soil Sci.* **2016**, *179*, 94–103. [CrossRef]
10. Page, K.; Dalal, R.; Dang, Y. How useful are MIR predictions of total, particulate, humus, and resistant organic carbon for examining changes in soil carbon stocks in response to different crop management? A case study. *Soil Res.* **2013**, *51*, 719–725. [CrossRef]
11. Haynes, R. Labile Organic Matter Fractions as Central Components of the Quality of Agricultural soils: An Overview. *Adv. Agron.* **2005**, *85*, 221–268. [CrossRef]
12. Sanderman, J.; Hengl, T.; Fiske, G.J. Soil carbon debt of 12,000 years of human land use. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 9575–9580. [CrossRef]
13. Soriano-Disla, J.M.; Janik, L.J.; Viscarra Rossel, R.A.; Macdonald, L.M.; McLaughlin, M.J. The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. *Appl. Spectrosc. Rev.* **2014**, *49*, 139–186. [CrossRef]
14. Bongiorno, G.; Bünemann, E.K.; Oguejiofor, C.U.; Meier, J.; Gort, G.; Comans, R.; Mäder, P.; Brussaard, L.; De Goede, R. Sensitivity of labile carbon fractions to tillage and organic matter management and their potential as comprehensive soil quality indicators across pedoclimatic conditions in Europe. *Ecol. Indic.* **2019**, *99*, 38–50. [CrossRef]
15. Gregorich, E.; Carter, M.; Angers, D.; Monreal, C.; Ellert, B. Towards a minimum data set to assess soil organic matter quality in agricultural soils. *Can. J. Soil Sci.* **1994**, *74*, 367–385. [CrossRef]
16. Baldock, J.; Hawke, B.; Sanderman, J.; Macdonald, L. Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. *Soil Res.* **2013**, *51*, 577–595. [CrossRef]
17. Zimmermann, M.; Leifeld, J.; Fuhrer, J. Quantifying soil organic carbon fractions by infrared-spectroscopy. *Soil Biol. Biochem.* **2007**, *39*, 224–231. [CrossRef]
18. Zhang, L.; Yang, X.; Drury, C.; Chantigny, M.; Gregorich, E.; Miller, J.; Bittman, S.; Reynolds, D.; Yang, J. Infrared spectroscopy prediction of organic carbon and total nitrogen in soil and particulate organic matter from diverse Canadian agricultural regions. *Can. J. Soil Sci.* **2018**, *98*, 77–90. [CrossRef]
19. Calderón, F.J.; Reeves, J.B.; Collins, H.P.; Paul, E.A. Chemical Differences in Soil Organic Matter Fractions Determined by Diffuse-Reflectance Mid-Infrared Spectroscopy. *Soil Sci. Soc. Am. J.* **2011**, *75*, 568–579. [CrossRef]
20. Janik, L.J.; Skjemstad, J.; Shepherd, K.; Spouncer, L. The prediction of soil carbon fractions using mid-infrared-partial least square analysis. *Soil Res.* **2007**, *45*, 73–81. [CrossRef]
21. Gredilla, A.; de Vallejuelo, S.F.O.; Elejoste, N.; De Diego, A.; Madariaga, J.M. Non-destructive Spectroscopy combined with chemometrics as a tool for Green Chemical Analysis of environmental samples: A review. *TrAC Trends Anal. Chem.* **2016**, *76*, 30–39. [CrossRef]
22. Barra, I.; Haefele, S.M.; Sakrabani, R.; Kebede, F. Soil spectroscopy with the use of chemometrics, machine learning and pre-processing techniques in soil diagnosis: Recent advances–A review. *TrAC Trends Anal. Chem.* **2021**, *135*, 116166. [CrossRef]

23. Armenta, S.; De la Guardia, M. Vibrational spectroscopy in soil and sediment analysis. *Trends Environ. Anal. Chem.* **2014**, *2*, 43–52. [CrossRef]
24. Du, C.; Zhou, J. Evaluation of Soil Fertility Using Infrared Spectroscopy—A Review. In *Climate Change, Intercropping, Pest Control and Beneficial Microorganisms*; Springer: Dordrecht, The Netherlands, 2009; pp. 453–483._16. [CrossRef]
25. Reeves III, J.B. Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done? *Geoderma* **2010**, *158*, 3–14. [CrossRef]
26. Bellon-Maurel, V.; McBratney, A. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils—Critical review and research perspectives. *Soil Biol. Biochem.* **2011**, *43*, 1398–1410. [CrossRef]
27. Zhang, L.; Yang, X.; Drury, C.; Chantigny, M.; Gregorich, E.; Miller, J.; Bittman, S.; Reynolds, W.D.; Yang, J. Infrared spectroscopy estimation methods for water-dissolved carbon and amino sugars in diverse Canadian agricultural soils. *Can. J. Soil Sci.* **2018**, *98*, 484–499. [CrossRef]
28. Nocita, M.; Stevens, A.; van Wesemael, B.; Aitkenhead, M.; Bachmann, M.; Barthès, B.; Dor, E.B.; Brown, D.J.; Clairotte, M.; Csorba, A.; et al. Chapter Four-Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring. *Adv. Agron.* **2015**, *132*, 139–159. [CrossRef]
29. Gholizadeh, A.; Borůvka, L.; Saberioon, M.; Vašát, R. Visible, Near-Infrared, and Mid-Infrared Spectroscopy Applications for Soil Assessment with Emphasis on Soil Organic Matter Content and Quality: State-of-the-Art and Key Issues. *Appl. Spectrosc.* **2013**, *67*, 1349–1362. [CrossRef]
30. Seidel, M.; Hutengs, C.; Ludwig, B.; Thiele-Bruhn, S.; Vohland, M. Strategies for the efficient estimation of soil organic carbon at the field scale with vis-NIR spectroscopy: Spectral libraries and spiking vs. local calibrations. *Geoderma* **2019**, *354*, 113856. [CrossRef]
31. Deiss, L.; Margenot, A.J.; Culman, S.W.; Demyan, M.S. Tuning support vector machines regression models improves prediction accuracy of soil properties in MIR spectroscopy. *Geoderma* **2020**, *365*, 114227. [CrossRef]
32. Dangal, S.R.; Sanderman, J.; Wills, S.; Ramirez-Lopez, L. Accurate and Precise Prediction of Soil Properties from a Large Mid-Infrared Spectral Library. *Soil Syst.* **2019**, *3*, 11. [CrossRef]
33. Clairotte, M.; Grinand, C.; Kouakoua, E.; Thébault, A.; Saby, N.P.; Bernoux, M.; Barthès, B.G. National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. *Geoderma* **2016**, *276*, 41–52. [CrossRef]
34. Baumann, P.; Helfenstein, A.; Gubler, A.; Keller, A.; Meuli, R.G.; Wachter, D.; Lee, J.; Viscarra Rossel, R.A.; Six, J. Developing the Swiss mid-infrared soil spectral library for local estimation and monitoring. *Soil* **2021**, *7*, 525–546. [CrossRef]
35. Seybold, C.A.; Ferguson, R.; Wysocki, D.; Bailey, S.; Anderson, J.; Nester, B.; Schoeneberger, P.; Wills, S.; Libohova, Z.; Hoover, D.; et al. Application of Mid-Infrared Spectroscopy in Soil Survey. *Soil Sci. Soc. Am. J.* **2019**, *83*, 1746–1759. [CrossRef]
36. Cezar, E.; Nanni, M.R.; Guerrero, C.; da Silva Junior, C.A.; Cruciol, L.G.T.; Chicati, M.L.; Silva, G.F.C. Organic matter and sand estimates by spectroradiometry: Strategies for the development of models with applicability at a local scale. *Geoderma* **2019**, *340*, 224–233. [CrossRef]
37. Capron, X.; Walczak, B.; De Noord, O.; Massart, D. Selection and weighting of samples in multivariate regression model updating. *Chemom. Intell. Lab. Syst.* **2005**, *76*, 205–214. [CrossRef]
38. Guerrero, C.; Zornoza, R.; Gómez, I.; Mataix-Beneyto, J. Spiking of NIR regional models using samples from target sites: Effect of model size on prediction accuracy. *Geoderma* **2010**, *158*, 66–77. [CrossRef]
39. Guerrero, C.; Stenberg, B.; Wetterlind, J.; Viscarra Rossel, R.; Maestre, F.; Mouazen, A.M.; Zornoza, R.; Ruiz-Sinoga, J.; Kuang, B. Assessment of soil organic carbon at local scale with spiked NIR calibrations: Effects of selection and extra-weighting on the spiking subset. *Eur. J. Soil Sci.* **2014**, *65*, 248–263. [CrossRef]
40. Stork, C.L.; Kowalski, B.R. Weighting schemes for updating regression models–a theoretical approach. *Chemom. Intell. Lab. Syst.* **1999**, *48*, 151–166. [CrossRef]
41. Sankey, J.B.; Brown, D.J.; Bernard, M.L.; Lawrence, R.L. Comparing local vs. global visible and near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. *Geoderma* **2008**, *148*, 149–158. [CrossRef]
42. Frank, I.E.; Friedman, J.H. A Statistical View of Some Chemometrics Regression Tools. *Technometrics* **1993**, *35*, 109–135. [CrossRef]
43. Brown, D.J.; Bricklemyer, R.S.; Miller, P.R. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma* **2005**, *129*, 251–267. [CrossRef]
44. Jean-Philippe, S.R.; Labbé, N.; Franklin, J.A.; Johnson, A. Detection of mercury and other metals in mercury contaminated soils using mid-infrared spectroscopy. *Proc. Int. Acad. Ecol. Environ. Sci.* **2012**, *2*, 139–149.
45. Stellacci, A.M.; Castellini, M.; Diacono, M.; Rossi, R.; Gattullo, C.E. Assessment of soil quality under different soil management strategies: Combined use of statistical approaches to select the most informative soil physico-chemical indicators. *Appl. Sci.* **2021**, *11*, 5099. [CrossRef]
46. Bricklemyer, R.S.; Brown, D.J.; Barefield, J.E.; Clegg, S.M. Intact Soil Core Total, Inorganic, and Organic Carbon measurement Using Laser-Induced Breakdown Spectroscopy. *Soil Sci. Soc. Am. J.* **2011**, *75*, 1006–1018. [CrossRef]
47. Indahl, U.G.; Liland, K.H.; Næs, T. Canonical partial least squares—A unified PLS approach to classification and regression problems. *J. Chemom.* **2009**, *23*, 495–504. [CrossRef]
48. Wetterlind, J.; Stenberg, B. Near-infrared spectroscopy for within-field soil characterization: Small local calibrations compared with national libraries spiked with local samples. *Eur. J. Soil Sci.* **2010**, *61*, 823–843. [CrossRef]

49. Kunzová, E.; Hejcman, M. Yield development of winter wheat over 50 years of FYM, N, P and K fertilizer application on black earth soil in the Czech Republic. *Field Crop. Res.* **2009**, *111*, 226–234. [CrossRef]

50. Šimon, T. Quantitative and qualitative characterization of soil organic matter in the long-term fallow experiment with different fertilization and tillage. *Arch. Agron. Soil Sci.* **2007**, *53*, 241–251. [CrossRef]

51. Madaras, M.; Koubova, M.; Lipavský, J. Stabilization of available potassium across soil and climatic conditions of the Czech Republic. *Arch. Agron. Soil Sci.* **2010**, *56*, 433–449. [CrossRef]

52. Stehlíková, I.; Madaras, M.; Lipavský, J.; Šimon, T. Study on some soil quality changes obtained from long-term experiments. *Plant Soil Environ.* **2016**, *62*, 74–79. [CrossRef]

53. Smatanová, M.; Vodáková, M. *Porovnání účinnosti Digestátů s Různými Typy Hnojiv při Hospodaření ve Zranitelné Oblasti.* Technical Report; Ústřední Kontrolní a Zkušební ústav Zemědělský: Brno, Czech Republic, 2020.

54. Lipavský, J.; Kubát, J.; Zobač, J. Long-term effects of straw and farmyard manure on crop yields and soil properties. *Arch. Agron. Soil Sci.* **2008**, *54*, 369–379. [CrossRef]

55. Hejcman, M.; Kunzová, E.; Šrek, P. Sustainability of winter wheat production over 50 years of crop rotation and N, P and K fertilizer application on illimerized luvisol in the Czech Republic. *Field Crop. Res.* **2012**, *139*, 30–38. [CrossRef]

56. Sparling, G.; Vojvodić-Vuković, M.; Schipper, L. Hot-water-soluble C as a simple measure of labile soil organic matter: The relationship with microbial biomass C. *Soil Biol. Biochem.* **1998**, *30*, 1469–1472. [CrossRef]

57. Rinnan, Å.; van den Berg, F.; Engelsen, S.B. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal. Chem.* **2009**, *28*, 1201–1222. [CrossRef]

58. Peris-Díaz, M.D.; Krężel, A. A guide to good practice in chemometric methods for vibrational spectroscopy, electrochemistry, and hyphenated mass spectrometry. *TrAC Trends Anal. Chem.* **2021**, *135*, 116157. [CrossRef]

59. Kennard, R.W.; Stone, L.A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11*, 137–148. [CrossRef]

60. Minasny, B.; McBratney, A.B. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput. Geosci.* **2006**, *32*, 1378–1388. [CrossRef]

61. Chen, S.; Arrouays, D.; Angers, D.A.; Martin, M.P.; Walter, C. Soil carbon stocks under different land uses and the applicability of the soil carbon saturation concept. *Soil Tillage Res.* **2019**, *188*, 53–58. [CrossRef]

62. Six, J.; Conant, R.T.; Paul, E.A.; Paustian, K. Stabilization mechanisms of soil organic matter: Implications for C-saturation of soils. *Plant Soil* **2002**, *241*, 155–176. [CrossRef]

63. Janik, L.; Forrester, S.; Rawson, A. The prediction of soil chemical and physical properties from mid-infrared spectroscopy and combined partial least-squares regression and neural networks (PLS-NN) analysis. *Chemom. Intell. Lab. Syst.* **2009**, *97*, 179–188. [CrossRef]

64. Kynčlová, P.; Filzmoser, P.; Hron, K. Modeling compositional time series with vector autoregressive models. *J. Forecast.* **2015**, *34*, 303–314. [CrossRef]

65. Blair, G.J.; Lefroy, R.D.B.; Lisle, L. Soil Carbon Fractions Based on their Degree of Oxidation, and the Development of a Carbon Management Index for Agricultural Systems. *Aust. J. Agric. Res.* **1995**, *46*, 1459–1466. [CrossRef]

66. Chen, J.; Zhang, X.; Hron, K. Partial least squares regression with compositional response variables and covariates. *J. Appl. Stat.* **2021**, *48*, 3130–3149. [CrossRef]

67. Hastie, T.; Tibshirani, R.; Friedman, J. Chapter Model Assessment and Selection. In *The Elements of Statistical Learning*; Springer: New York, NY, USA, 2009; pp. 219–260.

68. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.

69. Oksanen, J.; Blanchet, F.G.; Friendly, M.; Kindt, R.; Legendre, P.; McGlinn, D.; Minchin, P.R.; O'Hara, R.B.; Simpson, G.L.; Solymos, P.; et al. Vegan: Community Ecology Package. Available online: https://CRAN.R-project.org/package=vegan (accessed on 19 February 2020).

70. Stevens, A.; Ramirez-Lopez, L. An Introduction to the Prospectr Package. Available online: https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr.html (accessed on 9 May 2022).

71. Mevik, B.H.; Wehrens, R.; Liland, K.H. pls: Partial Least Squares and Principal Component Regression. Available online: https://CRAN.R-project.org/package=pls (accessed on 19 February 2020).

72. Roudier, P. clhs: A R Package for Conditioned Latin Hypercube Sampling. Available online: https://CRAN.R-project.org/package=clhs (accessed on 19 February 2020).

73. van den Boogaart, K.G.; Tolosana-Delgado, R. "compositions": A unified R package to analyze compositional data. *Comput. Geosci.* **2008**, *34*, 320–338. [CrossRef]

74. Stallman, R.M.; McGrath, R.; Smith, P.D. *GNU Make. A Program for Directing Recompilation*; Free Software Foundation: Boston, MA, USA, 2016.

75. Courtès, L.; Wurmus, R. Reproducible and User-Controlled Software Environments in HPC with Guix. In Proceedings of the Euro-Par 2015: Parallel Processing Workshops, Vienna, Austria, 24–25 August 2015; pp. 579–591. [CrossRef]

76. Saeys, W.; Mouazen, A.M.; Ramon, H. Potential for Onsite and Online Analysis of Pig Manure Using Visible and Near Infrared Reflectance Spectroscopy. *Biosyst. Eng.* **2005**, *91*, 393–402. [CrossRef]

77. Farmer, V. Transverse and longitudinal crystal modes associated with OH stretching vibrations in single crystals of kaolinite and dickite. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2000**, *56*, 927–930. [CrossRef]

78. Madejová, J.; Kečkéš, J.; Pálková, H.; Komadel, P. Identification of components in smectite/kaolinite mixtures. *Clay Miner.* **2002**, *37*, 377–388. [CrossRef]

79. Tatzber, M.; Stemmer, M.; Spiegel, H.; Katzlberger, C.; Haberhauer, G.; Gerzabek, M. An alternative method to measure carbonate in soils by FT-IR spectroscopy. *Environ. Chem. Lett.* **2007**, *5*, 9–12. [CrossRef]

80. Demyan, M.; Rasche, F.; Schulz, E.; Breulmann, M.; Müller, T.; Cadisch, G. Use of specific peaks obtained by diffuse reflectance Fourier transform mid-infrared spectroscopy to study the composition of organic matter in a Haplic Chernozem. *Eur. J. Soil Sci.* **2012**, *63*, 189–199. [CrossRef]

81. D'acqui, L.; Pucci, A.; Janik, L. Soil properties prediction of western Mediterranean islands with similar climatic environments by means of mid-infrared diffuse reflectance spectroscopy. *Eur. J. Soil Sci.* **2010**, *61*, 865–876. [CrossRef]

82. Knox, N.; Grunwald, S.; McDowell, M.; Bruland, G.; Myers, D.; Harris, W. Modelling soil carbon fractions with visible near-infrared (VNIR) and mid-infrared (MIR) spectroscopy. *Geoderma* **2015**, *239–240*, 229–239. [CrossRef]

83. Ng, W.; Minasny, B.; Malone, B.; Filippi, P. In search of an optimum sampling algorithm for prediction of soil properties from infrared spectra. *PeerJ* **2018**, *6*, e5722. [CrossRef] [PubMed]

84. Hinkle, J.; Rayens, W. Partial least squares and compositional data: Problems and alternatives. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 159–172. [CrossRef]

85. Calderón, F.J.; Culman, S.; Six, J.; Franzluebbers, A.J.; Schipanski, M.; Beniston, J.; Grandy, S.; Kong, A.Y. Quantification of Soil Permanganate Oxidizable C (POXC) Using Infrared Spectroscopy. *Soil Sci. Soc. Am. J.* **2017**, *81*. [CrossRef]

86. Yang, X.; Xie, H.; Drury, C.; Reynolds, W.; Yang, J.; Zhang, X. Determination of organic carbon and nitrogen in particulate organic matter and particle size fractions of Brookston clay loam soil using infrared spectroscopy. *Eur. J. Soil Sci.* **2012**, *63*, 177–188. [CrossRef]

87. Stumpe, B.; Weihermüller, L.; Marschner, B. Sample preparation and selection for qualitative and quantitative analyses of soil organic carbon with mid-infrared reflectance spectroscopy. *Eur. J. Soil Sci.* **2011**, *62*, 849–862. [CrossRef]

88. Parikh, S.J.; Goyne, K.W.; Margenot, A.J.; Mukome, F.N.; Calderón, F.J. Chapter One - Soil Chemical Insights Provided Through Vibrational Spectroscopy. *Adv. Agron.* **2014**, *126*, 1–148. [CrossRef]

89. Tinti, A.; Tugnoli, V.; Bonora, S.; Francioso, O. Recent applications of vibrational mid-Infrared (IR) spectroscopy for studying soil components: A review. *J. Cent. Eur. Agric.* **2015**, *16*, 1–22. [CrossRef]

90. Reeves, J.B. Mid-infrared diffuse reflectance spectroscopy: Is sample dilution with KBr necessary, and if so, when? *Am. Lab.* **2003**, *35*, 24–28.

91. Brown, D.J. Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed. *Geoderma* **2007**, *140*, 444–453. [CrossRef]

92. Vohland, M.; Ludwig, B.; Seidel, M.; Hutengs, C. Quantification of soil organic carbon at regional scale: Benefits of fusing vis-NIR and MIR diffuse reflectance data are greater for in situ than for laboratory-based modelling approaches. *Geoderma* **2022**, *405*, 115426. [CrossRef]