Europ. J. Agronomy 78 (2016) 44–59

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley

Pablo Rischbeck[a,*], Salah Elsayed[b], Bodo Mistele[c], Gero Barmeier[a], Kurt Heil[a], Urs Schmidhalter[a]

[a] Chair of Plant Nutrition, Technical University of Munich, Emil-Ramann-Straße 2, D-85350 Freising-Weihenstephan, Germany
[b] Evaluation of Natural Resources Department, Environmental Studies and Research Institute, Sadat City University, Egypt
[c] Department of Agricultural Engineering, Fachhochschule Südwestfalen, Soest, Germany

ARTICLE INFO

Article history:
Received 30 October 2015
Received in revised form 5 April 2016
Accepted 27 April 2016

Keywords:
Abiotic stress
Data fusion
High throughput phenotyping
Multi-annual
Multiple linear regression
Partial least square
Phenomics
Yield modelling

ABSTRACT

Yield modelling based on visible and near infrared spectral information is extensively used in proximal and remote sensing for yield prediction of crops. Distance and thermal information contain independent information on canopy growth, plant structure and the physiological status. In a four-years' study hyperspectral, distance and thermal high-throughput measurements were obtained from different sets of drought stressed spring barley cultivars. All possible binary, normalized spectral indices as well as thirteen spectral indices found by others to be related to biomass, tissue chlorophyll content, water status or chlorophyll fluorescence were calculated from hyperspectral data and tested for their correlation with grain yield. Data were analysed by multiple linear regression and partial least square regression models, that were calibrated and cross-validated for yield prediction. Overall partial least square models improved yield prediction ($R^2 = 0.57$; RMSEC = 0.63) compared to multiple linear regression models ($R^2 = 0.46$; RMSEC = 0.74) in the model calibration. In cross-validation, both methods yielded similar results (PLSR: $R^2 = 0.41$, RMSEV = 0.74; MLR: $R^2 = 0.40$, RMSEV = 0.78). The spectral indices R_{780}/R_{550}, R_{760}/R_{730}, R_{780}/R_{700}, the spectral water index R_{900}/R_{970} and laser and ultrasonic distance parameters contributed favourably to grain yield prediction, whereas the thermal based crop water stress index and the red edge inflection point contributed little to the improvement of yield models. Using only more uniform modern cultivars decreased the model performance compared to calibrations done with a set of more diverse cultivars. The partial least square models based on data fusion improved yield prediction ($R^2 = 0.62$; RMSEC = 0.59) compared to the partial least square models based only on hyperspectral data ($R^2 = 0.48$; RMSEC = 0.69) in the model calibration. This improvement was confirmed by cross-validation (data fusion: $R^2 = 0.39$, RMSEV = 0.76; hyperspectral data only: $R^2 = 0.32$, RMSEV = 0.79). Thus, a combination of spectral multiband and distance sensing improved the performance in yield prediction compared to using only hyperspectral sensing.

1. Introduction

In agricultural remote and proximal sensing the visible and near infrared spectral range is used for modelling crop yield (Hansen et al., 2002; Li et al., 2015). A number of yield related plant traits like green biomass (Mistele et al., 2012), canopy water mass (Winterhalter et al., 2011), leaf senescence and chlorophyll content (Kipp et al., 2014) can be assessed spectrally. The spectral reflectance in the water band at 970 nm (Peñuelas et al., 1993) and 1450 nm in addition to various normalised difference indices combining NIR and VIS wavelengths (Rischbeck et al., 2014) can be used to assess plant tissue dehydration. Mistele et al. (2004) tested several spectral indices that were obtained from an oligo view optic device connected to a spectrometer in winter wheat field trials. The NDVI $(R_{780} - R_{670})/(R_{780} + R_{670})$ and NIR/Red (R_{780}/R_{670}) indices were correlated with biomass. The indices Red Edge/NIR (R_{740}/R_{780}), NIR/Green (R_{780}/R_{550}) and NIR/Red (R_{780}/R_{700}) are useful for assessing biomass and the nitrogen status. Unlike the simple ratio or normalised indices that are sensitive to the pigment absorption in the visible range, the red edge inflection point (REIP) indicates the shift of the slope connecting the reflectance in the red and NIR spectral regions (Herrmann et al., 2010). The REIP is

* Corresponding author.
E-mail address: Pablo.rischbeck@wzw.tum.de (P. Rischbeck).

http://dx.doi.org/10.1016/j.eja.2016.04.013
1161-0301/

Fig. 1. PhenoTrac 4, measuring platform equipped with hyperspectral, thermal and laser distance sensors developed by the Chair of Plant Nutrition of the Technical University of Munich.

an indicator of the amount of chlorophyll in the field of view of the sensor. Mistele et al. (2004) found the REIP is useful for assessing both the biomass and the nitrogen status, because both parameters are related to the amount of absorbing pigments.

However, the information that can be obtained from spectrometry is limited to the interactions of light with plant or soil material. Transpiration is not directly related to spectral reflectance. Canopy height is related to passive spectral reflectance by the leaf number and leaf architecture. Sensors using other physical principles might add information to spectral assessments. The use of information from different sensors might also increase the flexibility of yield prediction under different weather conditions and at different growth stages of the crop.

For example, thermometry detects thermal infrared emission in the range of 8–20 μm and is used to measure surface temperature (Jones, 2004). Idso et al. (1981) developed a crop water stress index based on thermal measurements that ranges between 0 (full transpiration) and 1 (minimum stomatal conductance).

Laser distance sensors measure the traveling time of light to and from a reflecting surface. Ultrasonic distance sensors measure the time interval between sending a sound signal and receiving the echo to measure the distance to an object. Both sensors enable the measurement of canopy height and structure (Zhao et al., 2010).

In recent years unmanned aerial vehicles have been equipped with spectrometers (Agüera Vega et al., 2015) and stereo camera systems are used for distance measurements (Honkavaara et al., 2013) for applications in precision agriculture. These systems allow rapid image and data acquisition of neighbouring fields. A number of workgroups have used carrier vehicles equipped with spectral and other sensors to increase the throughput of trait assessments in breeding nurseries and field trials (Schmidhalter et al., 2001; Montes et al., 2011; White et al., 2012; Comar et al., 2012; Rebetzke et al., 2013). The Chair of Plant Nutrition from the Technical University of Munich developed a carrier vehicle (PhenoTrac 4, Fig. 1), equipped with hyperspectral, thermal and distance sensors (Mistele and Schmidhalter, 2010; Erdle et al., 2011; Winterhalter et al., 2013; Kipp et al., 2014). In this study statistical data fusion of sensor measurements was used and tested for improving yield modelling under different levels of drought stress and at different phenological stages.

Under temperate European climatic conditions the interannual precipitation can vary strongly, ranging from sufficient to deficient. In Germany since 2000 the yearly precipitation was below the longterm average (1881–2014) in the years 2003, 2006, 2011, 2012 and 2014 (Deutscher Wetterdienst, 2015). Drought stress due to low water availability in the soil and high potential evapotranspiration affects cereal crops by decreasing their growth and yield. The direct impacts of drought stress are reduced water uptake by roots, decreasing water status, reduced stomatal conductance, reduced CO_2 uptake, reduced transpiration and increased surface temperature (Bradford and Hsiao, 1982; Schmidhalter et al., 1998; Siddique et al., 2000; Wall et al., 2011). Reduced water uptake from dry soils also decreases nitrogen uptake and affects the nutritional status of the crop (Rostamza et al., 2011). For this study rain-out shelters were used to impose mild to severe drought stress on spring barley grown in field trials. Variable growth conditions in cereal production can be tested by including artificially created drought conditions into model calibrations.

Yield models can be adapted to a certain species, genotype and local soil and growth conditions (Laurent et al., 2015). However, statistical yield modelling based on remote sensing is most commonly used on a regional scale using sensors mounted on satellites (Zhao et al., 2015). Regional and multi-field applications require models, that can cope with different genotypes and soils and the spatial and interannual variability of weather and abiotic stresses that affect plant growth. Combining data across phenological stages, field trials and years in broad calibrations might lead to such broadly adapted yield models or improve yield prediction.

There has been progress in developing sophisticated statistical approaches for linking data gathered from sensors with agronomic or physiological traits. Thus, yield modelling approaches can be based on single spectral indices (Gutierrez-Rodriguez et al., 2004). Both linear regression models and multiple linear regression models can also be derived from spectral indices. Royo et al. (2003) found that 17.3% to 65.2% of durum wheat yield could be accounted for with multiple linear regression models based on varying spectral indices. However, multiple linear regression analysis is limited to a few predictor variables in cases of collinearity of the measured variables. Additionally, multiple linear regression analysis also requires homogeneity of variance (homoscedasticity) (Heil and Schmidhalter, 2012).

An alternative approach is to use partial least square regression (PLSR). In PLSR orthogonal components are unaffected by collinearity and are derived from all variables. Partial least square regression models of hyperspectral reflectance were used by Weber et al. (2012) at anthesis and during the milk-grain stage of maize (*Zea mays* L.). The models explained between 49% to 69% of the variation in grain yield in calibration studies and 23% to 40% after model validation. Hyperspectral data has also been used for yield modelling in partial least square regressions (Sharabian et al., 2014). Fu et al. (2013) combined optimal binary, normalized indices retrieved in a spectral range of 350 nm to 2500 nm with band depth ratios in partial least square regression models for assessing winter wheat biomass. Biomass could be predicted with a maximum R^2 of 0.84 and a minimum RMSE of 0.177 kg m^{-2}.

The purpose of this work was: (i) to compare the performance of multiple linear regression and partial least square models for assessing grain yield in multi-annual experiments, (ii) to investigate which traits should be included in data fusion, (iii) to investigate the advantages of data combinations, (iv) to assess improvements in yield prediction by data fusion models, and (v) to test the methods by using more and less variable cultivars.

2. Materials and methods

2.1. Field trials

The field experiments were conducted at the Dürnast research station of the Technical University of Munich in southwestern

Fig. 2. Rain-out shelter facility with spring barley crop at heading stage.

Germany (11°41′60″E, 48°23′60″N). Dürnast is characterised by a sub-oceanic climate and has mild cloudy winters and warm summers. The average yearly precipitation is 787 mm, and the average temperature is 7.8 °C (Deutscher Wetterdienst, 1971–2000). At the research station two rain-out shelter facilities (Fig. 2) are used for conducting field trials under controlled drought stress. A rain-out shelter is a moving greenhouse mounted on tracks. The crops grow under open sky, and if it rains, the shelter closes to keep the plants and the soil dry.

During four years three types of experiments were conducted in this study: In rainfed field trials, spring barley (*Hordeum vulgare* L.) was grown in a randomised block design with four replicates on calcaric Cambisol consisting of silty loam. Additional mineral fertilizer was applied to increase the level to 130 kg N ha^{-1} before sowing in all trials, based on the residual nitrogen content in the soil measured after winter. Spring barley was sown in April 10, 2010, and in March 29, 2011, in the rainfed trials, and in April 17, 2012, and in April 18, 2013, in all trials. The seeding rate in all trials was 330 seeds per m^2. The plots consisted of 14 rows spaced 15 cm apart and had a length of 10 m. In 2010, 2011 and 2012, 16 historical and modern German cultivars were grown. The use of herbicide and fungicide treatments was done in all trials when necessary. To reduce the lodging risk, especially of the historic cultivars, the growth regulator "Trinexapac" was applied at the initiation of shooting (BBCH 31) in all trials.

The mildly drought-stressed field trials were established in a randomized block design with five replicates in a rain-out shelter facility in the seasons 2012 and 2013. The soil is a calcaric Cambisol consisting of silty loam. It has a field capacity of 42%$_v$, and the permanent wilting point is at 20%$_v$. The rooting depth for fully developed spring barley was assumed to be at 1.5 m, based on previous measurements of water extraction in the subsoil (data not shown). The plant available field capacity in the rooting depth amounts to 330 mm. The rain-out shelter was kept closed during autumn and winter to keep the soil dry. The crop was spray irrigated only once every season with 20 mm for germination. There was no further irrigation used, which led to a continuous but mild drought stress. The plots consisted of eight rows spaced 15 cm apart and had a length of 1.8 m. In 2012, sixteen historical and modern German cultivars were grown. In 2013, sixteen modern international cultivars were grown.

Severely drought-stressed field trials were performed in a randomized block design with four replicates in a rain-out shelter facility in 2012 and 2013. The re-filled soil under the rain-out shelter consists of a sandy loam soil mixture (22.6% silt, 53.3% fine sand, 19.1% medium sand and 5% coarse sand) at a depth of 0–20 cm. The sandy soil which is found at a depth of 20–100 cm consists of 60% medium sand, 30.6% fine sand, 3.4% coarse sand, 4.4% silt and 1.6% clay. The soil will hereafter be termed sandy soil. The plant available field capacity in the soil profile is 90 mm. The crop was spray irrigated during vegetative growth to ensure good crop establishment on the poor soil. Irrigation was stopped before heading, and the plants were kept dry for 7 days in 2012 and 12 days in 2013. After this short but severe stress phase during heading and anthesis, the crops were re-irrigated. The total amount of water added during the whole season was 114 mm in 2012 and 132 mm in 2013. The plots consisted of eight rows spaced 15 cm apart and had a length of 1.8 m. In 2012, nine historical and modern German cultivars were grown, and in 2013, twelve modern international cultivars were grown. In 2013 a control trial under full sprinkler irrigation was carried out on sandy soil with the same twelve modern cultivars as in the severe drought stress trial without replications.

The rainfed trials were harvested with a plot combine harvester, whereas the rain-out shelter trials were harvested by hand. The total grain yield was determined and the samples were oven-dried to determine grain water content on a gravimetric basis. The yield was expressed as t ha^{-1} and was normalized to a water content of 14%$_w$.

2.2. Cultivars

German and international cultivars were selected regarding similar phenological development. An impact of variable heading and anthesis dates on sensor measurements and yield could largely be excluded. Table 2 indicates the cultivars used in the field trials.

2.3. Hyperspectral reflectance measurements

The passive bidirectional reflectance sensor system (tec5, Oberursel, Germany) used for this study contains two units of a Zeiss MMS1 silicon diode array spectrometer. One unit is linked to a diffuser detecting solar radiation as a reference signal. Simultaneously, the second unit measured the canopy reflectance with a 12° field of view. The bi-directional sensor was calibrated before measurements with a PTFE white standard. The sensor analyses the reflected radiation in 200 spectral channels. It measures reflectance and incident radiation simultaneously in the spectral detection range of 352 nm to 1000 nm with a bandwidth of 3.26 nm (Mistele and Schmidhalter, 2008; Winterhalter et al., 2013; Erdle et al., 2013).

2.4. Contour plot

A contour plot showing the coefficients of determination (R^2) of the relationships between grain yield and the narrow band normalized differential spectral indices was calculated from all possible two-band combinations in the range of 352–1000 nm (formula: (R1 − R2)/(R1 + R2)) from measurements in the mildly drought stressed field trial on 19 June 2013, 17:10 using the software R (R Foundation, Vienna, Austria) and was visualised using the lattice package for R.

2.5. Spectral indices

In this study, spectral indices were calculated from the reflectance measured by the passive hyperspectral sensor. The indices R_{780}/R_{550}, R_{780}/R_{670}, R_{760}/R_{730} (Erdle et al., 2011), R_{780}/R_{700}, R_{780}/R_{740} (Mistele and Schmidhalter, 2008, 2010), the triangular vegetation index (TVI; $0.5(120^*(R_{750}\text{-}R_{550})\text{-}200(R_{670}\text{-}R_{550}))$) (Broge and Leblanc, 2000), the modified triangular vegetation index 1 (MTVI1; $1.2\ (1.2\ (R_{800}\text{-}R_{550})\text{-}2.5\ (R_{670}\text{-}R_{550}))$) (Haboudane et al., 2004), the renormalized difference vegetation index (RDVI; $(R_{800}\text{-}R_{670})/(R_{800}+R_{670})^2$) (Haboudane et al., 2004) are mainly related

to the biomass. The red edge inflection point (REIP; ((700 + 40 (R_{670} + R_{780})/2–R_{700})/R_{740}–R_{700})) (Guyot et al., 1988), the blue green pigment index (BGI; R_{450}/R_{550}) (Zarco-Tejada et al., 2005), the modified chlorophyll absorption in reflectance index 2 (MCARI2) (as described in Haboudane et al., 2004) are related to the tissue chlorophyll content. The index R_{900}/R_{970} (Peñuelas et al., 1993) is related to the water status. The fluorescence ratio (FR2; R_{690}/R_{740}) was found to be related to chlorophyll fluorescence (Lichtenthaler, 1996).

2.6. Thermal measurements

The canopy surface temperature was determined in the field trials with two HEITRONICS KT15D infrared thermometers (Heitronics GmbH, Wiesbaden, Germany). The thermal sensors have a spectral response between 8 and 20 μm, and the temperature resolution is 0.06 °C. The diameter of the field of view is 3–10 cm at a distance of 40–100 cm. Rotating mirrors inside the case are used for referencing the case temperature and allow for highly stable environmental measurements. The infrared thermometers were mounted on the carrier vehicle PhenoTrac4, from the Chair of Plant Nutrition, Technical University of Munich. The thermometers view the canopy from two opposed oblique views at an angle of 45° from the nadir. The canopy temperature was assessed by averaging the measurements from both devices. The oblique view increases the biomass fraction and decreases the soil fraction in the field of view. Using measurements from opposite sides provides information from both sun-lit and shaded canopy.

2.7. Crop water stress index

The Crop Water Stress Index (CWSI) was calculated according to the following formula by Jackson et al. (1981):

$$CWSI = \frac{T - T_{min}}{T_{max} - T_{min}} \quad (1)$$

where, T is the actual infrared temperature measured in the canopy, T_{min} is the lowest temperature measured in the whole field trial (lower baseline) and T_{max} is the highest temperature in the whole field trial (upper baseline). Jackson et al. (1981) used the wet bulb temperature as the lower baseline and the dry bulb temperature as the upper baseline. In this experiment, canopy measurements were used. The advantage of this method is, that no additional measurements other than recording the infrared temperature are necessary. Even in drought-stressed large field trials, varying temperatures of fully transpiring and non-transpiring leaves can be found and used as baselines. CWSI makes temperature measurements interpretable as reduction in stomatal conductance and allows a comparison of drought stress levels in different field trials under different environmental conditions, including air temperature, radiation and air humidity.

2.8. Distance measurements

To measure the structure and height of the canopy, an ultrasound sensor was mounted on the carrier vehicle and pointed nadir (UM 30-14113, Sick AG, Germany). The sensor measures a maximum distance of 3.4 m and has a repetition rate of 5 measurements s^{-1}. The measured distance between the sensor and the canopy was subtracted from the mounting height of 1 m.

A laser runtime distance sensor (OWTG 4100 PE S1, Welotec GmbH, Germany) was also used. The sensor measures a maximum distance of 10 m and has a repetition rate of 5 measurements s^{-1}.

2.9. Carrier vehicle

The sensor devices were mounted 1 m above ground on the mobile multi-sensor phenotyping platform PhenoTrac4. The sensors were mounted in a nadir position to provide high-throughput measurements of all plots (Kipp et al., 2014).

The sensor data were co-registered with corresponding GPS coordinates from a Trimble RTK-GPS (real-time kinematic global positioning system) (Trimble, Sunnyvale, CA, USA).

2.10. Field measurements

The sensor measurements were taken mostly on sunny days in 2010, 2011, 2012 and 2013. Passive hyperspectral measurements are favoured at direct sunlight providing good illumination of the canopy. Weather conditions in southern Germany are variable, what explains the requirement to measure also at different phenological stages. Table 1 gives an overview of the field measurements done.

2.11. Statistical analysis

2.11.1. Multiple linear regression

We considered the following multiple linear regression model:

$$Y = X\beta + \varepsilon, \quad (2)$$

where, Y is the response variable yield and X is a n × N design matrix consisting of n observations of N covariates (predictors). The models were fitted in SPSS (IBM Corp., 2012) using the spectral indices, ultrasonic distance, laser distance and CWSI as covariates. Suitable models were chosen by using the stepwise function. Spectral indices, ultrasonic distance and laser distance variables were standardised (mean was set to zero, standard deviation was set to one). A residual analysis showed an adequate fit for all models under consideration, which indicates homoscedasticity. A 5-fold cross validation was applied using the bootstrap package in R (Efron and Tibshirani, 1993).

To test interactions between cultivars and model parameters, the cultivars were divided into three classes according to their year of release (<1950, 1950–1990, >1990). The multiple linear regression models were enlarged by these cultivar classes and their interactions with one parameter in each model. Thus, the interactions between all ten parameters and the cultivar classes were tested.

2.11.2. Partial least square regression

The underlying model for the partial least squares (PLS) method is the following:

$$\boldsymbol{X} = \boldsymbol{T}P^T + \boldsymbol{E} \quad (3)$$

$$Y = \boldsymbol{U}Q^T + \varepsilon, \quad (4)$$

where, Y and $\boldsymbol{X}$ are defined as in 2.9.1, $\boldsymbol{T}$ and U are $n \times p$ matrices of latent vectors, the $N \times p$ matrix P and the $1 \times p$ matrix Q are matrices consisting of loadings. The $n \times p$ matrix E and the n dimensional vector ε are residuals. The loadings $\boldsymbol{T}$ and $\boldsymbol{U}$ are chosen so that the sample covariance is equal to the maximal covariance between linear transformations of $\boldsymbol{X}$ and Y as described by Rosipal and Krämer (2006).

The partial least square regression models for all single measurements and data combinations were fitted and compared with multiple linear regressions. Data modelling was performed with the PLS (Mevik and Wehrens, 2007) package in R (R Core Team, 2014) software. The same predictor variables (spectral indices, ultrasonic distance, laser distance and CWSI) used in the MLR model were used for modelling yield. Spectral indices, ultrasonic distance and

Table 1
List of field measurements done at different dates and growth stages (BBCH): the drought stress treatments depicting the trial type (rainfed, mild and severe drought stress); the soil environments [sandy loam/sandy soil mixture (sandy soil) and silty loam soil]; and the cultivars being used, either historical or modern as well as their origin and their number in the respective trials, are indicated.

Date	BBCH	Treatment	Environment	Cultivars		
				No.	Origin	Type
17 June 12	59	severe stress	sandy soil	9	German	hist.[a] and modern
18 June 12	60	severe stress	sandy soil	9	German	hist. and modern
21 June 12	63	severe stress	sandy soil	9	German	hist. and modern
28 May 13	35	severe stress	sandy soil	12	International	modern
19 June 13	56	severe stress	sandy soil	12	International	modern
25 May 12	32	mild stress	silty loam	16	German	hist. and modern
18 June 13	55	mild stress	silty loam	16	International	modern
19 June 13	53	mild stress	silty loam	16	International	modern
1 July 10	71	rainfed	silty loam	16	German	hist. and modern
28 June 11	88	rainfed	silty loam	16	German	hist. and modern
12 July 12	85	rainfed	silty loam	16	German	hist. and modern
16 June 12	59	rainfed	silty loam	16	German	hist. and modern

[a] Hist. . .historical: tall cultivars, modern: short-growing cultivars with dwarfing genes.

Table 2
Cultivars used in the field trials.

Trial	Years	Cultivars
Rainfed	2010,2011,2012	Heils Franken[a], Apex, Trumpf, Isaria[a], Pflugs Intensiv[a], Ackermanns Bavaria[a], Sissy, Perun,
Mild Stress	2012	Barke, Wiebke, Djamila, Beatrix, Eunova, Victoriana, Streif, Ursa
Mild stress	2013	Sissy, Trumpf, Barke, Streif, Ursa, IPZ 24727, DH 179 (Argentine), BRS 195, Gimpel, Extract, Eunova, Djamila, Wiebke, Scarlett, Grace, Quench
Severe Stress	2012	Heils Franken[a], Apex, Ackermanns Bavaria[a], Barke, Wiebke, Djamila, Beatrix, Eunova, Ursa
Severe Stress	2013	Sissy, Trumpf, Barke, Streif, Ursa, IPZ 24727, DH 179 (Argentine), Djamila, Wiebke, Scarlett,
Irrigated Control		Grace, Quench

[a] Historical cultivars.

laser distance variables were standardised (mean was set to zero, standard deviation was set to one). The number of components was chosen according to the minimum root mean square error of validation. To assess the importance of the predictor variables their loadings on the first and second component were calculated. A 5-fold cross validation was applied within the data sets.

We also evaluated whether the inclusion of non-spectral information leads to improved yield modelling. The partial least square regressions between hyperspectral data (range 352 nm–1000 nm), spectral indices (as enumerated in Section 2.3) and CWSI were modelled for single measurements and data combinations. The calibration and validation was applied as described in the above section.

2.11.3. Data combinations

The data obtained from high-throughput measurements were combined for yield modelling in three different ways. First, the yield was assessed based on repeated measurements in one field trial at different phenological stages (rainfed trial in 2012: BBCH 59 + 85, severely drought stressed trial in the rain-out shelter (ROS) with sandy soil; BBCH 59 + 63, mildly drought stressed trial in the rain-out shelter with silty loam; BBCH 52 + 53). The yield under severe drought stress (ROS sandy soil) and under mild drought stress (ROS silty loam) was assessed by combining measurements of both trials on the same day (19 June 2013, BBCH 53). The yields obtained in the rainfed trials in the years 2010, 2011 and 2012 were assessed by combining measurements taken at similar phenological stages in each year (2010: BBCH 71, 2011: BBCH 88, 2012: BBCH 85).

2.11.4. Coefficients of determination, root mean square error, relative root mean square error and normalized root mean square error

The performance of the models based on either multiple linear regression or partial least square regression was evaluated by comparing the prediction abilities for grain yield using the following coefficient of determination:

$$R^2 = \frac{\left(\sum_{i=1}^{n}(y_i - \bar{y})\left(\hat{y}_i - \bar{\hat{y}}\right)\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^{2'}} \quad (5)$$

where, y is the harvested yield and $\hat{y}$ is the fit from the model. The coefficient of determination shows the portion of yield variance explained by the model based on the total yield variance. The root mean square error (RMSE) of cross validation/prediction was calculated using the following equation:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (6)$$

which is the standard deviation of the differences between harvested and predicted yield. A high R^2 value and a low *RMSE* mark a good yield prediction. *RMSE* does not allow comparing the predictive ability within trials with different grain yield variability. Thus, the relative root mean square error (*rRMSE*) was calculated by dividing the *RMSE* by the standard deviation of yield (as previously described in Li et al. (2008). The normalised root mean square error (*nRMSE*) was calculated by dividing the *RMSE* by the mean yield. All parameters were calculated for the calibration and validation of both the MLR and PLSR models.

3. Results

3.1. Yield and drought stress in field trials

The mean yield of the field trials ranged from 2.9 to 6.63 Mg ha^{-1}. The standard deviation of the yield within the field trials ranged from 0.74 to 1.66 Mg ha^{-1}. The rainfed field trials, on average, yielded 5.50 Mg ha^{-1}. The field trials on silty loam soil had an aver-

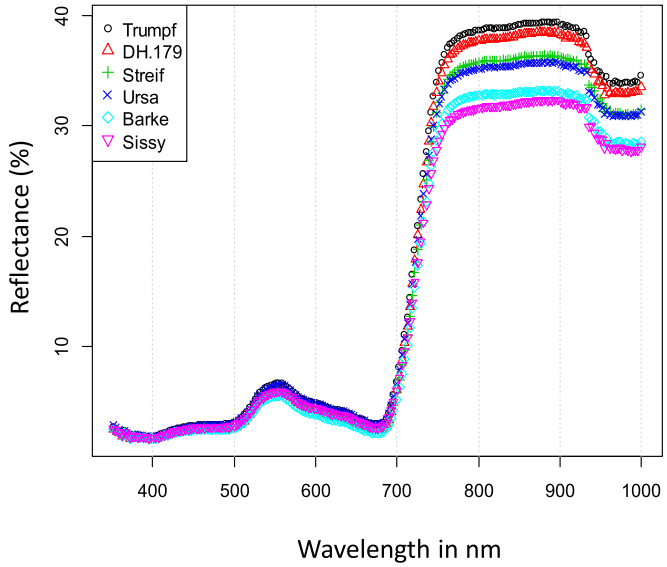

Fig. 3. Hyperspectral reflectance in the range from 352 nm to 1000 nm of six different spring barley cultivars grown in a severe drought stress trial measured on 19 June 2013 at 17:10 (BBCH 56).

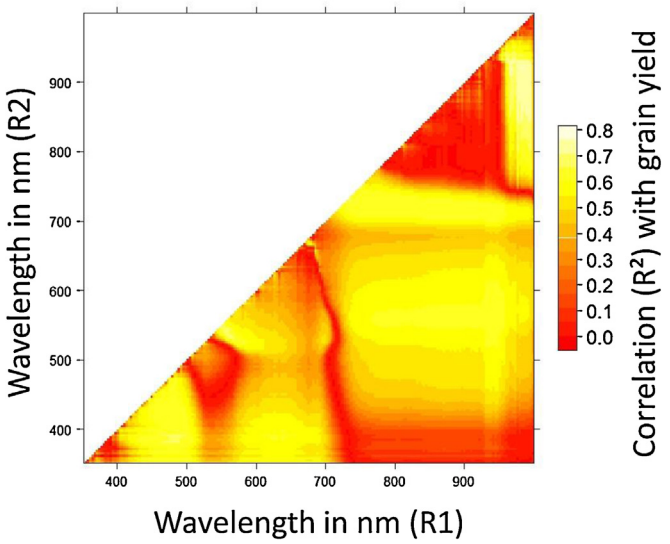

Fig. 4. Coefficients of determination (R^2) of normalised, two-band spectral indices measured in a mildly drought-stressed rain-out shelter trial, a severely drought stressed rain-out shelter trial and an irrigated control trial on 19 June 2013 at 17:10 (BBCH 56) with spring barley grain yield.

age yield of 6.42 Mg ha^{-1}, while the field trials on the sandy soil, on average, yielded 4.22 Mg ha^{-1}. The yield ranking of the cultivars within the trials was not constant (data not shown). However, more recent cultivars, such as Djamila (registration year: 2003), had significantly higher yields than historical cultivars, such as Ackermanns Bavaria (registration year: 1903), under both stressed and well-watered conditions.

The crop water stress index at the time of measurement ranged from 0.26 (mild drought stress) to 0.83 (severe drought stress). The rainfed field trials had an average crop water stress index (CWSI) of 0.35. The field trials on the silty loam soil had an average CWSI of 0.4, while the field trials on sandy soil had an average CWSI of 0.55.

3.2. Hyperspectral reflectance

Reflectance in the range 352–1000 nm (average of 4 different plots) of six different cultivars grown in a severe drought stress trial measured on 19 June 2013 at 17:10 (BBCH 56) is shown in Fig. 3. Comparably low reflectance is evident in the visible range (400–700 nm), increasing reflectance is shown in the red edge range (700–760 nm), while high reflectance is shown in the near infrared range (>760 nm). Reflectance is variable among cultivars in the visible range from 530 to 690 nm, and especially in the near infrared range. The data quality of the sensor is generally high, only at increased wavelengths >926 nm some noise occurs due to a lower sensitivity of the sensor.

3.3. Contour plots

The contour plot (Fig. 4) gives an overview on wavelength combinations of binary, normalized spectral indices showing high correlations with yield. Especially VIS/VIS combinations around R1: 486 nm and R2: 391 nm, VIS/VIS combinations around R1: 630 nm and R2: 391 nm, Red Edge/VIS or NIR/VIS combinations around R1: 732–996 nm and R2: 562 nm, Red Edge/Red Edge or NIR/Red Edge combinations around R1: 719–996 nm and R2: 715 nm and NIR/NIR combinations around R1: 983 nm and R2: 771–958 nm show high correlations with yield, with a maximum R^2 of 0.76 at R1: 983 nm and R2: 919 nm.

Table 3
Ranking of coefficients of determination of linear models between spectral indices, height and thermal parameters measured on 19 June 2013; 17:10 (BBCH 56) in a mildly drought-stressed rain-out shelter trial, a severely drought stressed rain-out shelter trial and an irrigated control trial with grain yield.

parameter	R^2	parameter	R^2	parameter	R^2
R900/R970	0.69	R780/R670	0.48	BGI	0.31
CWSI	0.64	REIP	0.44	MTVI1	0.29
R760/R730	0.62	FR2	0.42	RDVI	0.19
R780/R740	0.61	Ultrasonic	0.41	MCARI2	0.16
R780/R550	0.58	Laser	0.36		
R780/R700	0.56	TVI	0.32		

3.4. Selection of indices

All spectral indices, as well as the measurements of the CWSI and the distance parameters calculated for 19 June 2013, 17:10 (BBCH 56) in a mildly drought-stressed rain-out shelter trial, a severely drought stressed rain-out shelter trial and an irrigated control trial were ranked according to their correlations with grain yield (Table 3). Seven indices showing the highest correlations (R_{900}/R_{970}, R_{760}/R_{730}, R_{780}/R_{740}, R_{780}/R_{550}, R_{780}/R_{700}, R_{780}/R_{670} and REIP), the CWSI and the distance parameters were chosen for further statistical analysis.

3.5. Assessment and selection of predictors in multiple linear regression models

The predictors of grain yield were assessed based on the number and quality of their significance in the multiple linear regression (MLR) models (Tables S1 and Table 4).

All variables except REIP were significant in at least one of the MLR models. Based on the data significance, the water index R_{900}/R_{970} was important for yield modelling within multiple linear regressions. The other simple ratio indices and the height parameters showed intermediate importance. The CWSI had low importance, and the REIP index was not included in the models.

The presence of high absolute amounts (not regarding the algebraic sign) of β-coefficients in Table S1 is mostly in agreement with the significance of the variables. The highest average absolute amounts are indicated by the variables R_{900}/R_{970} (0.56), R_{780}/R_{550}

Table 4
Number of significances of the predictors in 12 multiple linear yield regression models of grain yield based on plant height, spectral indices and crop water stress index (CWSI) of barley cultivars grown rainfed or in rain-out shelters with different soil substrates (for details see Table S1). The counts are calculated as the sums of the number of significances multiplied by: 1($p < 0.1$), 2($p < 0.05$) or 3($p < 0.01$).

p	Ultra-sonic	Laser	R_{780}/R_{550}	R_{780}/R_{670}	R_{780}/R_{700}	R_{760}/R_{730}	R_{780}/R_{740}	R_{900}/R_{970}	REIP	CWSI
<0.1	1						1			
<0.05	2		1			1		1		2
<0.01	1	3	2	3	1	3	3	6		
counts	8	9	8	9	3	11	10	20	0	4

(0.44), R_{760}/R_{730} (0.39), R_{780}/R_{700} (0.39). The non-spectral parameters showed comparably small average absolute values for the laser distance (0.22), ultrasonic distance (0.18) and the CWSI (0.17).

The variance inflation factors of the significant variables ranged from one to eight and showed a limited collinearity between predictors. The collinearity can be less than a VIF of ten only if the number of predictors in the models is limited to four or less.

The stepwise multiple linear regression was applied in the combinations of data from two different phenological phases in the rainfed trial 2012, the trial on sandy soil in 2012, and the trial on the silty loam soil in 2013. In the rainfed trial the predictors from both phenological phases were selected in the MLR models (Table S2). The model results changed compared to the models from single dates (Table S1). In the other two trials, the parameters from only one phenological phase were selected by MLR and the models did not differ from models based on single measurements.

3.6. Calibration performance and validation of multiple linear regression models

The MLR models of grain yield based on spectral indices, plant height and CWSI reached an R^2 for calibration of 0.15–0.83 (Table 5). The *RMSE* of the calibration ranged from 0.32 to 1.13 Mg ha^{-1}. The *rRMSE* of the calibration ranged from 40 to 91.9%.

The R^2 of the validation for the MLR models of grain yield ranged from 0.1 to 0.79 (Table 5). The *RMSE* of the validation ranged from 0.36 to 1.15 Mg ha^{-1}. The *rRMSE* of the validation ranged from 45 to 96.1%.

The quality of yield prediction tended to increase across the phenological stages. A regression line of R^2 calibration plotted against BBCH showed a positive slope and a significant correlation of $R^2 = 0.23$. A regression line of *RMSEC* plotted against BBCH showed a negative slope and a significant correlation of $R^2 = 0.20$. In the yield models enlarged by the cultivar classes, no significant interactions of the cultivar classes with one of the ten parameters were found.

3.7. Assessment of predictors in partial least square regression models

All of the parameters that were included in the partial least square (PLSR) models including plant height, spectral indices and CWSI had loadings over 0.1 for most of the first and second components (except CWSI). The loadings on the first component are predominantly positive, while loadings on the second component were predominantly negative. The loadings of the spectral indices showed different patterns than loadings of the distance parameters. On the first component spectral indices had absolute loadings (averaged over all 12 single measurements) close to 0.35, while the distance parameters showed only absolute average loadings close to 0.26. On the second component spectral indices had absolute average loadings close to 0.25, while the distance parameters had absolute average loadings close to 0.58.

Fig. 5 shows loadings on the first and second component for a yield model based on hyperspectral data. The first component shows values around −0.07 for the visible range until 600 nm and

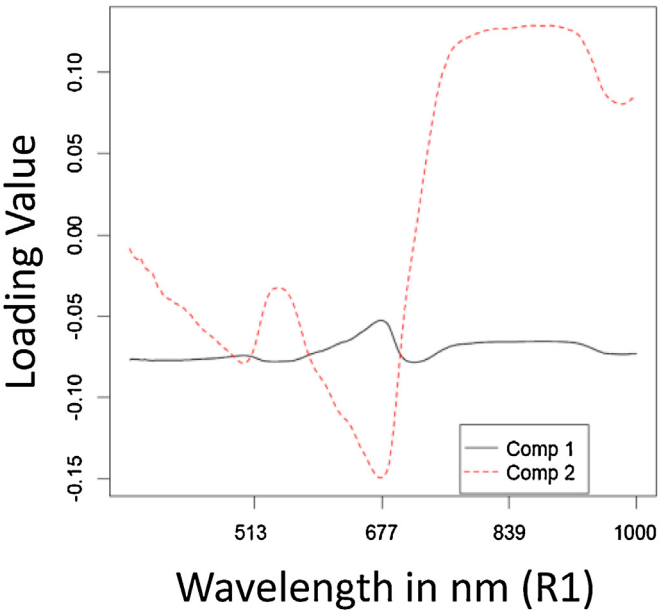

Fig. 5. Loadings on the first and second component of a partial least square regression yield model based on hyperspectral data from a severely drought-stressed rain-out shelter trial (sandy soil) measured on 18 June 2012 at 14:23.

for the near infrared range starting from 780 nm. Loadings on the second component show a peak of −0.15 at 677 nm (red) and values around 0.13 from 780 until 910 nm.

In the partial least square regression models of data combinations all of the parameters that were included in the PLSR models, including CWSI, had loadings over 0.1 for most of the first and second components. The loadings of the spectral indices showed different patterns than loadings of the distance parameters and also differed from loadings of the CWSI. The first component spectral indices had absolute average loadings (averaged over all lines) of approximately 0.27. The distance parameters had absolute average loadings of 0.19, and the CWSI had average loadings of 0.16. The second component spectral indices had absolute average loadings close to 0.14. The distance parameters had absolute average loadings of 0.44. The CWSI had average loadings of 0.24.

3.8. Calibration performance and validation of partial least square regression models

The PLSR models of grain yield based on spectral indices, plant height and CWSI reached an R^2 for calibration of 0.28–0.84 (Table 6). The *RMSE* of the calibration ranged from 0.31 to 0.82 Mg ha^{-1}. The *rRMSE* of the calibration ranged from 38.8% to 83.7%.

The R^2 for the validation of PLSR models of grain yield ranged from 0 (failure of validation) to 0.80. The *RMSE* of the validation ranged from 0.36 to 1.06 Mg ha^{-1}. The *rRMSE* of the validation ranged from 45 to 103.9%.

The quality of yield prediction tended to increase throughout the phenological stages. A regression line of R^2 calibration plotted

Table 5
Assessment of multiple linear regression models of grain yield based on plant height, spectral indices and thermal crop water stress index (CWSI) of spring barley cultivars grown rainfed or in rain-out shelters (ROS) with different soil substrates.

Field trial	Date	BBCH	CWSI	$\overline{X}$Yield	σ Yield	n	R^2 calibration	RMSEC Mg/ha	rRMSEC (%)	nRMSEC (%)	R^2 validation	RMSEV t/ha	rRMSEV (%)	nRMSEV (%)
rainfed	1 July 10, 13:34	71	0.26	6.00	0.82	73	0.32	0.67	81.7	11.2	0.22	0.73	89.0	12.1
rainfed	28 June 11, 13:36	88	0.3	4.69	0.88	48	0.56	0.58	65.9	12.3	0.48	0.63	71.6	13.4
ROS Loam	25 May 12, 15:02	32	0.33	6.63	0.98	84	0.22	0.86	87.8	12.9	0.21	0.86	87.8	13.0
ROS Loam	19 June 13, 17:10	53	0.35	6.54	1.66	84	0.57	1.08	65.0	16.5	0.54	1.13	68.1	17.2
rainfed	12 July 12, 16:13	85	0.4	5.66	1.08	64	0.71	0.58	53.7	10.2	0.63	0.66	61.1	11.6
ROS Sand	17 June 12, 15:44	59	0.4	4.78	0.74	27	0.43	0.55	74.3	11.5	0.38	0.57	77.0	12.0
rainfed	16 June 12, 11:48	59	0.45	5.66	1.08	64	0.30	0.90	83.3	15.8	0.30	0.90	83.3	16.0
ROS Sand	21 June 12, 10:09	63	0.48	4.94	0.80	48	0.54	0.53	66.3	10.8	0.49	0.56	70.0	11.4
ROS Sand	28 May 13, 16:45	30	0.48	3.56	1.02	17	0.32	0.81	79.4	22.8	0.10	0.98	96.1	27.5
ROS Loam	18 June 13, 14:27	52	0.51	6.10	1.28	64	0.58	0.83	64.8	13.6	0.54	0.86	67.19	14.1
ROS Sand	18 June 12, 14:23	59	0.56	4.94	0.80	48	0.83	0.32	40.0	6.5	0.79	0.36	45.0	7.3
ROS Sand	19 June 13, 17:10	53	0.83	2.90	1.23	48	0.15	1.13	91.9	38.9	0.12	1.15	93.5	39.6
Rainfed	12	59 + 85		5.66	1.08	64	0.79	0.50	45.81	8.77	0.74	0.55	50.61	9.68
ROS Sand	12	59 + 63		4.94	0.80	48	0.83	0.32	40.0	6.5	0.79	0.36	45.0	7.3
ROS Loam	13	52 + 53		6.1	1.28	64	0.58	0.83	64.68	13.55	0.54	0.86	67.19	14.11
ROS Sand + Loam	13	53		5.2	2.3	132	0.74	1.13	50.44	21.4	0.73	1.17	51.84	22.02
Rainfed	10 + 11 + 12	71 + 88 + 85		5.55	1.07	185	0.45	0.79	74.28	14.32	0.42	0.81	75.78	14.61

Table 6
Assessment of partial least square regression models of grain yield, including plant height, spectral indices and thermal crop water stress index (CWSI) of spring barley cultivars grown rainfed or in rain-out shelters (ROS) with different soil substrates.

Field trial	Date	BBCH	CWSI	No of Comps	$\bar{X}$Yield Mg/ha	σ Yield Mg/ha	n	R^2 calibration	RMSEC Mg/ha	rRMSEC (%)	nRMSEC (%)	R^2 validation	RMSEV Mg/ha	rRMSEV (%)	nRMSEV (%)
rainfed	1 July 10, 13:34	71	0.26	4	6.00	0.82	73	0.36	0.65	79.3	10.8	0.29	0.69	84.1	11.5
rainfed	28 June 11, 13:36	88	0.3	5	4.69	0.88	48	0.72	0.47	53.4	10.0	0.56	0.58	65.9	12.4
ROS Loam	25 May 12, 15:02	32	0.33	5	6.63	0.98	84	0.28	0.82	83.7	12.4	0.14	0.90	91.8	13.6
ROS Loam	19 June 13, 17:10	53	0.35	5	6.54	1.66	84	0.58	0.82	49.4	12.5	0.50	0.88	53.0	13.5
rainfed	12 July 12, 16:13	85	0.4	4	5.66	1.08	64	0.71	0.57	52.8	10.1	0.66	0.62	57.4	11.0
ROS Sand	17 June 12, 15:44	59	0.4	5	4.78	0.74	27	0.75	0.37	50.0	7.7	0.47	0.53	71.6	11.1
rainfed	16 June 12, 11:48	59	0.45	6	5.66	1.08	64	0.49	0.77	71.3	13.6	0.33	0.88	81.5	15.5
ROS Sand	21 June 12, 10:09	63	0.48	4	4.94	0.80	48	0.58	0.51	63.8	10.3	0.43	0.59	73.8	11.9
ROS Sand	28 May 13, 16:45	30	0.48	2	3.56	1.02	17	0.37	0.78	76.5	21.9	0.00	1.06	103.9	29.8
ROS Loam	18 June 13, 14:27	52	0.51	5	6.10	1.28	64	0.66	0.74	57.8	12.1	0.57	0.84	65.6	13.8
ROS Sand	18 June 12, 14:23	59	0.56	4	4.94	0.80	48	0.84	0.31	38.8	6.3	0.80	0.36	45.0	7.3
ROS Sand	19 June 13, 17:10	53	0.83	5	2.90	1.23	48	0.49	0.77	62.6	26.6	0.22	0.95	77.2	32.8
rainfed	12	59 + 85		6	5.66	1.08	64	0.81	0.47	43.5	8.3	0.72	0.57	52.8	10.1
ROS Sand	12	59 + 63		7	4.94	0.80	48	0.88	0.25	31.3	5.1	0.46	0.53	66.3	10.7
ROS Loam	13	52 + 53		5	6.1	1.28	64	0.70	0.70	54.7	11.5	0.54	0.86	67.2	14.1
ROS Sand + Loam	13	53		3	5.2	2.3	132	0.72	1.02	44.3	19.6	0.67	1.11	48.3	21.3
rainfed	10 + 11 + 12	71 + 88 + 85		5	5.55	1.07	185	0.47	0.77	72.0	13.9	0.44	0.80	74.8	14.4

against BBCH had a positive slope and a significant relationship of $R^2 = 0.32$ (data not shown). A regression line of *RMSEC* plotted against BBCH had a negative slope and a significant correlation of $R^2 = 0.29$. A regression line of *rRMSEC* plotted against BBCH had a negative slope and a significant correlation of $R^2 = 0.20$. A regression line of *nRMSEC* plotted against BBCH had a negative slope and a significant correlation of $R^2 = 0.21$.

Similar to the calibration, the regression line of R^2 validation plotted against BBCH had a positive slope and a significant correlation of $R^2 = 0.41$. A regression line of *RMSEV* plotted against BBCH had a negative slope and a significant correlation of $R^2 = 0.40$.

The quality of yield prediction tended to increase with increasing drought stress. A regression line of R^2 calibration plotted against CWSI had a positive slope and a significant correlation of $R^2 = 0.16$ (measurement under severe drought stress on 19 June 2013; 17:10 not included, plot not shown). A regression line of *rRMSEC* plotted against CWSI had a negative slope and a significant correlation of $R^2 = 0.12$ (measurement under severe drought stress on 19 June 13; 17:10 not included).

Soil texture also influenced the quality of yield prediction. The R^2 of grain yield averaged over all measurements on sandy soil (0.61) was higher than the average over all measurements on silty loam soil (0.50). On average, the *RMSEC* was smaller on sandy soil (0.55 Mg ha^{-1}) than on silty loam soil (0.79 Mg ha^{-1}). The highest R^2 (0.66) of grain yield on silty loam soil was reached at the beginning of heading (BBCH 52). On sandy soil, the highest R^2 (0.84) was reached at the end of heading (BBCH 59). After validation the R^2 at the beginning of heading was reduced to 0.57 on silty loam soil, while it was reduced to 0.80 at the end of heading on sandy soil.

Repeated measurements in the same field trial in different phenological stages were combined to test possible improvements in the grain yield models. The data combination in the rainfed field trial in 2012 led to an improved R^2 of 0.81 compared with an R^2 of 0.49 at the end of heading (BBCH 59) and 0.71 in the drought treatment at BBCH 85. The *RMSEC* decreased to 0.47 for the combination compared to 0.77 at BBCH 59 and 0.57 at BBCH 85 (Table 7).

The data combination in the field trial on sandy soil in 2012 led to a slightly improved R^2 of 0.88 compared with the single measurements (R^2 of 0.84 at the end of heading – BBCH 59, and 0.58 at anthesis – BBCH 63). The *RMSEC* decreased to 0.25 for the combination compared to 0.31 at BBCH 59 and 0.51 at BBCH 63.

The data combination in the field trial on silty loam soil in 2013 led to a slightly improved R^2 of 0.70 compared to two measurements at the beginning of heading with an R^2 of 0.66 (BBCH 52) and 0.58 (BBCH 53). The *RMSEC* slightly decreased to 0.70 for the combination compared to 0.74 at BBCH 52 and 0.82 at BBCH 53.

To test whether the PLSR models of grain yield are applicable under large environmental variability the data from two field trials on silty loam and sandy soil measured at the same time (19 June 2013; 17:10) were combined. The combination led to an improved R^2 of 0.72 compared with an R^2 of 0.49 on sandy soil and 0.58 on silty loam soil.

To develop more broadly adapted models of grain yield, the data from the rainfed field trials in the years 2010, 2011 and 2012 were combined. The combination led to a R^2 of 0.47 compared to an R^2 of 0.36 in 2010, 0.72 in 2011 and 0.71 in 2012. The root mean square error of calibration increased from 0.65 Mg ha^{-1} in 2010, 0.47 Mg ha^{-1} in 2011 and 0.57 Mg ha^{-1} in 2012 to 0.77 Mg ha^{-1} for the combination.

3.9. Comparison of methods to model grain yield

The PLSR models of grain yield based on spectral indices, plant height and CWSI showed a better performance in the calibration than multiple linear regression models based on the same data (Tables 4 and 7). The R^2 of the calibration averaged over all single measurements increased from 0.46 in MLR models to 0.57 in PLSR models. The average *RMSEC* of all single measurements decreased from 0.74 in MLR models to 0.63 in PLSR models.

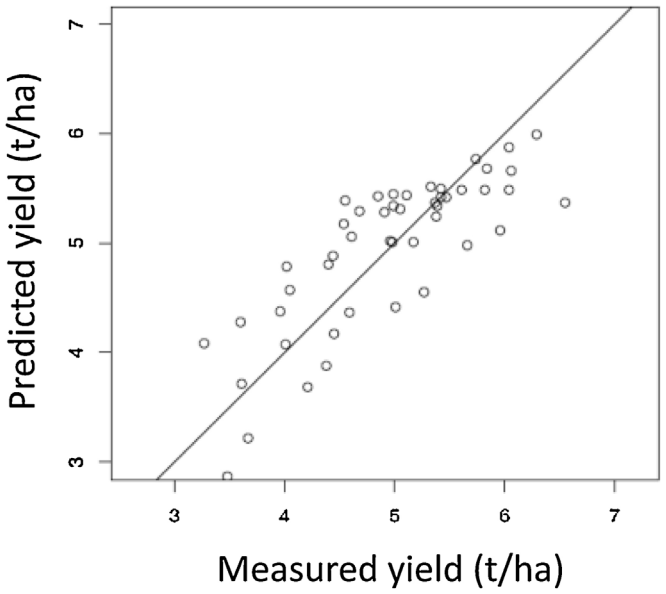

Fig. 6. Scatter plot and least square regression line between harvested yield and yield modelled by partial least square regression based on hyperspectral data measured in a rain-out shelter with sandy soil on 18 June 2012 at 14:23.

The better performance of the PLSR models was not confirmed by the results of the cross validation. The average R^2 of the validation of all single measurements was 0.40 in the MLR models and 0.41 in the PLSR models. The average *RMSEV* of all single measurements decreased slightly from 0.78 in the MLR models to 0.74 in the PLSR models.

The data combinations in PLSR partly showed advantages in comparison with MLR. The R^2 averaged over the three models combining different phenological phases increased from 0.73 for the MLR models to 0.80 for the PLSR models. The *RMSEC* averaged over the three models combining different phenological phases decreased from 0.55 for the multiple linear regression models to 0.47 for the PLSR models. The R^2 of the field trial combination remained constant (MLR: 0.74; PLSR: 0.72). However, the *RMSEC* decreased from 1.13 in the MLR model to 1.02 for the PLSR model. The R^2 of the combination of years remained constant (MLR: 0.45; PLSR: 0.47). Furthermore, the *RMSEC* also remained constant (MLR: 0.79; PLSR: 0.77).

The PLSR models of grain yield based on hyperspectral data reached an R^2 in the calibration of 0.18–0.76 (calibration data not shown). The yield modelling based on data from 28 May 2013, was not successful. The *RMSE* of the calibration ranged from 0.45 to 0.98 Mg ha^{-1}. The *rRMSE* of the calibration varied between 48.2% and 90.7%.

The R^2 of the validation for PLSR models of grain yield ranged between 0 (failure of validation) to 0.65 (Table 7, Fig. 6). The *RMSE* of the validation ranged from 0.47 to 1.06 Mg ha^{-1}. The *rRMSE* of the validation varied between 53 and 97.3%.

The PLSR models of grain yield based on data fusion of spectral indices, plant height, CWSI and hyperspectral data reached an R^2 in the calibration of 0.41–0.86 (calibration data not shown). The *RMSE* of the calibration ranged from 0.29 to 0.77 Mg ha^{-1}. The *rRMSE* of the calibration varied between 36.3 and 76.8%.

The R^2 of the validation of the PLSR models of grain yield based on data fusion ranged between 0 (failure of validation) to 0.76 (Table 7, Fig. 7). The *RMSE* of the validation ranged from 0.39 to

Table 7
Validation of partial least square regression models of grain yield, including hyperspectral data only resp. plant height, spectral indices, crop water stress index (CWSI) and hyperspectral data of spring barley cultivars grown rainfed or in rain-out shelters (ROS) with different soil substrates.

							hyperspectral data only				Data fusion			
Field trial	Date	BBCH	No of Comps	$\bar{X}$Yield Mg/ha	σ Yield Mg/ha	n	R^2 validation	RMSEV Mg/ha	rRMSEV (%)	nRMSEV (%)	R^2 validation	RMSEV Mg/ha	rRMSEV (%)	nRMSEV (%)
rainfed	1 July 10, 13:34	71	6	6.00	0.82	73	0.29	0.69	84.1	11.5	0.27	0.70	85.4	11.7
rainfed	28 June 11, 13:36	88	6	4.69	0.88	48	0.45	0.65	73.9	13.9	0.53	0.60	68.2	12.8
ROS Loam	25 May 12, 15:02	32	6	6.63	0.98	84	0.03	0.95	96.9	14.3	0.09	0.92	93.9	13.9
ROS Loam	19 June 13, 17:10	53	5	6.54	1.66	84	0.52	0.88	53.0	13.5	0.56	0.84	50.6	12.8
rainfed	12 July 12, 16:13	85	7	5.66	1.08	64	0.57	0.70	64.8	12.4	0.66	0.62	57.4	11.0
ROS Sand	17 June 12, 15:44	59	4	4.78	0.74	27	0	0.72	97.3	15.1	0.51	0.50	67.6	10.5
rainfed	16 June 12, 11:48	59	9	5.66	1.08	64	0	1.05	97.2	18.6	0.36	0.85	73.1	14.0
ROS Sand	21 June 12, 10:09	63	4	4.94	0.80	48	0.44	0.59	73.8	11.9	0.38	0.62	77.5	12.6
ROS Sand	28 May 13, 16:45	30	3	3.56	1.02	17	–	–	–	–	0.00	1.16	113.7	32.6
ROS Loam	18 June 13, 14:27	52	5	6.10	1.28	64	0.52	0.88	68.8	14.4	0.59	0.80	62.5	13.1
ROS Sand	18 June 12, 14:23	59	5	4.94	0.80	48	0.65	0.47	58.8	9.5	0.76	0.39	48.8	7.9
ROS Sand	19 June 13, 17:10	53	5	2.90	1.23	48	0.01	1.06	86.2	36.6	0.00	1.07	87.0	36.9
rainfed	12	59 + 85	6	5.66	1.08	64	0.49	0.77	71.3	13.6	0.66	0.62	57.4	11.0
ROS Sand	12	59 + 63	5	4.94	0.80	48	0.20	0.65	81.3	13.2	0.46	0.54	67.5	10.9
ROS Loam	13	52 + 53	4	6.1	1.28	64	0.52	0.88	68.8	14.4	0.52	0.88	68.8	14.4
ROS Sand + Loam	13	53	5	5.2	2.3	132	0.70	1.06	46.1	20.4	0.69	1.08	47.0	20.8
rainfed	10 + 11 + 12	71 + 88 + 85	6	5.55	1.07	185	0.47	0.78	72.9	14.1	0.57	0.70	65.4	12.6

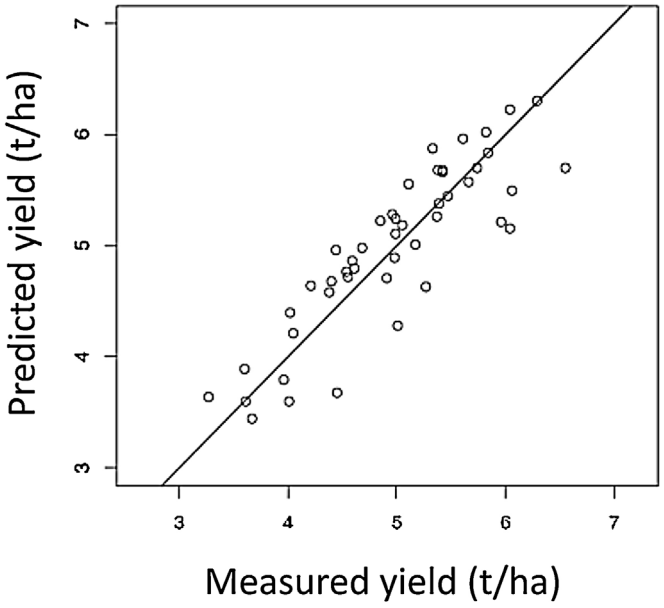

Fig. 7. Scatter plot and least square regression line between harvested yield and yield modelled by partial least square regression based on plant height, spectral indices, CWSI and hyperspectral data measured in a rain-out shelter with sandy soil on 18 June 2012 at 14:23.

1.16 Mg ha^{-1}. The *rRMSE* of the validation varied between 48.4 and 113.7%.

The overall performance of the PLSR models based on hyperspectral data could be improved by including spectral data, plant height parameters and the crop water stress index. The R^2 of the calibration averaged over all single measurements increased from 0.48 in the models based on hyperspectral data to 0.62 in the models based on data fusion. The average *RMSEC* of all single measurements decreased from 0.69 in the models based on hyperspectral data to 0.59 in the models based on data fusion (calibration data not shown).

The better performance of the models based on data fusion was confirmed by the results of the cross validation (Table 7). The R^2 of the validation averaged over all single measurements from the models based on hyperspectral data was 0.32 compared to 0.39 in the models based on data fusion. The average *RMSEV* of all single measurements in the models based on hyperspectral data was 0.79 compared to 0.76 in the models based on data fusion.

The inclusion of non-spectral data into the PLSR models led to improvements in the combined models. The R^2 of the calibration averaged over the models of repeated measurements at different phenological phases was 0.56 for grain yield models based on hyperspectral data and 0.72 for the PLSR models based on data fusion. The R^2 of the calibration of the field trial combination remained constant (PLSR, only hyperspectral: 0.76; PLSR, data fusion: 0.76). The R^2 of the calibration for the combination of years slightly improved (PLSR, only hyperspectral: 0.51; PLSR, data fusion: 0.60).

4. Discussion

Different soil and irrigation conditions led to varying degrees of drought stress, therefore different conditions were obtained in the field trials and were reflected in an extended range of the CWSI (0.26–0.83, Table S1). The quality of yield prediction tended to improve with increasing mild drought stress as indicated by the positive linear regressions obtained between the CWSI in the range between 0.26 to 0.56 and the R^2 and the negative linear regressions between the CWSI and *rRMSEC* (calculations based on figures in Table S1, regressions not shown). However, the yield prediction was not successful under severe drought stress (R^2 = 0.15 at CWSI 0.83). Weber et al. (2012) related leaf respectively canopy reflectance (350–2500 nm) to yield in individual PLSR models. They also reported an increased R^2 and a reduced *RMSE* for the models under drought stress compared to irrigated conditions. A higher variability in the water status under mild stress compared to severe stress or fully irrigated conditions within field trials facilitated the spectral assessment of yield.

In this study, the quality of yield prediction improved during crop development. This result was previously demonstrated by other authors (Marti et al., 2007; Gutierrez et al., 2010; Lobos et al., 2014). Aparicio et al. (2000) reported increasing coefficients of determination during the vegetation period of a normalised difference vegetation index ($R_{900} - R_{680}$)/($R_{900} + R_{680}$) and a simple ratio index (R_{900}/R_{680}) with durum wheat yield in an irrigated field trial. In a rainfed field trial the relationships were closest at anthesis and became slightly less accurate at maturity. Decreasing the time between yield prediction and the actual harvest lowers the risk of variable weather conditions or of other factors influencing yield.

4.1. Selection of indices

The contour plot (Fig. 4) shows four groups of indices containing information on yield, as influenced by distinct physiological traits; VIS/VIS indices indicate tissue chlorophyll content (Zarco-Tejada et al., 2001). Chlorophyll contains nitrogen and is related to nitrogen uptake and the nutritional status of the crop. NIR/Red Edge indices are also related to tissue pigment concentration and the crop nutrition status (Clevers et al., 2002). NIR/VIS indices mainly indicate soil cover and biomass (Carlson and Ripley, 1997), two factors related to spatial crop productivity. NIR/NIR indices around the water absorption bands at 970 nm are indicators of the crop water status (Peñuelas et al., 1993). Comparing the contour plot (Fig. 4) with simple ratio indices in Table 3 shows, that not only single distinct wavelength combinations, but an increased number of indices with similar wavelengths contains information relevant for grain yield. Simple ratio and normalised indices showed almost identical correlation with yield (data not shown). However, the indices chosen for further statistical analysis represent relevant NIR/NIR (R_{900}/R_{970}), NIR/VIS (R_{780}/R_{550}, R_{780}/R_{670}) and NIR/Red Edge (R_{780}/R_{740}, R_{780}/R_{700}, R_{760}/R_{730}, REIP) information.

4.2. Comparison of multiple linear regression and partial least square regression models

All multiple linear regression models based on spectral indices, distance parameters and the CWSI on single days predicted the yield. However, the quality of the yield prediction was varying, with R^2-values ranging from 0.15 to 0.83 (Table 5). These results are similar to those of Royo et al. (2003), who included the following measurements: single wavebands R_{550}, R_{680} and the indices water index (R_{970}/R_{900}); normalized differential vegetation index ($R_{900} - R_{680}$)/($R_{900} + R_{680}$); simple ratio index (R_{900}/R_{680}); photochemical reflectance index ($R_{531} - R_{570}$)/($R_{531} + R_{570}$); structural independent pigment index (R_{445}-R_{800})/(R_{680}-R_{800}); normalized total pigment to chlorophyll index (R_{680}-R_{430})/($R_{430} + R_{680}$); and normalized phaeophytinization index (R_{415}-R_{435})/($R_{415} + R_{435}$). These parameters were used in a stepwise multiple linear regression to predict the durum wheat yield under Mediterranean climate conditions. The coefficients of determination of the models for single days varied from 0.17 to 0.65. This variability in the model quality limits the feasibility of both PLSR and MLR models based on spectral information and height and temperature data for yield prediction and phenotyping. The rRMSE of 40% to 91.1% showed

that the measurement error represented a large proportion of the yield variability.

In contrast, partial least square models showed an enhanced performance compared to multiple linear regression models when considering the results from the calibration step. All of the PLSR models showed a slightly higher coefficient of determination (Table 5). All of the partial least square models also showed slightly lower root mean square errors. Other authors also compared the model performance. Hansen and Schjoerring (2003) found that the root mean square error of the partial least square models of hyperspectral data (438–883 nm) with green biomass was reduced by 22% when compared with the exponential regressions of normalised difference spectral indices with biomass.

The enhanced performance of the partial least square models compared to multiple linear regression models in the calibration could not be confirmed by improvements in the validation step (Tables 2 and 4). The PLSR model calibrations showed some tendency to overfitting, which could partly be avoided by careful selection of the number of components and had to be confirmed by validation.

Sharabian et al. (2014) compared the performance of different statistical approaches for modelling winter wheat yield based on hyperspectral data (350 nm to 2500 nm). A partial least square model was calibrated on a dataset of 40 samples obtained from two years and was validated on a set of 40 samples from a different year. The samples were taken in an incremental nitrogen fertilisation trial receiving 0–90 kg N ha^{-1}. The validation showed R^2-values of 0.87, an *RMSE* of 0.30 Mg ha^{-1} and an *nRMSE* of 4.53%. A stepwise multiple linear regression was performed based on ten wavelengths that were selected according to their correlation with yield. The validation showed improved results, indicating an R^2 of 0.89, an *RMSE* of 0.29 Mg ha^{-1} and an *nRMSE* of 4.28%. Sharabian et al. (2014) preferred multiple linear regressions because of the lower number of wavelengths and the easier interpretation of the results. The repeated and independent spectral yield assessments in breeding nurseries rely on the results of model validation. Thus, there is no consistent recommendation for one of the applied modelling methods. However, the partial least square model has other advantages. In stepwise multiple linear regression analysis, the inclusion of predictors is limited by the collinearity to four predictor variables varying for different models (Table S1, VIF-factors). Thus, there is no concise selection of predictors. In contrast, the partial least square analysis is not affected by collinearity. All of the available predictors can be included, and the only requirements are simple steps of data transformation (z-transformation).

4.3. Which traits to select?

In analysing the B-coefficients in partial least square analysis, Sharabian et al. (2014) found the wavelengths ranges from 405 to 480 nm, 520–750 nm and 1010–1350 nm were of major importance for yield prediction. Hansen and Schjoerring (2003) analysed the loadings of hyperspectral wavelengths in the range 438–883 nm in partial least square models of the green biomass. They found that NIR wavelengths had high positive loadings on the first component, while the wavelength 550 nm and the range 700–740 nm showed peaks of negative loadings on the second and third component. The PLSR models of yield indicated that the wavelengths 550 nm and 700–740 nm were particularly important.

The distance parameters and the spectral and thermal indices of the twelve multiple linear regression models are compared in Table 4 by evaluating their numbers of occurrence and significance levels in the models. The spectral indices R_{900}/R_{970}, R_{760}/R_{730}, R_{780}/R_{740}, R_{780}/R_{670} and R_{780}/R_{550} were found to be important for yield prediction. In the PLSR models, these spectral indices and the index R_{780}/R_{700}, which had no importance in MLR, dominated the first component. The indices R_{760}/R_{730}, R_{780}/R_{740} and R_{780}/R_{550} contained wavelengths which were also adopted by Sharabian et al. (2014) and Hansen and Schjoerring (2003). The indices R_{900}/R_{970} and R_{780}/R_{670} were found to be important in this study, but were not included in these wavelength ranges.

The water index R_{900}/R_{970} (Peñuelas et al., 1993) reacts sensitively to changes in the water status. Our results highlight its importance in predicting yield under drought stress and are in agreement with other authors. Lobos et al. (2014) found close relationships with coefficients of determination with $R^2 = 0.66$ under mild water stress for linear models between a normalised water index $(R_{970}\text{-}R_{920})/(R_{970} + R_{920})$ and wheat grain yield. Under severe water stress and full irrigation, the models had lower coefficients of determination, with $R^2 = 0.58$. Serrano et al. (2011) found coefficients of determination with $R = 0.61$ ($R^2 = 0.37$) in linear models of the water index with grape yield.

The CWSI reacts to the cooling effect of canopy transpiration. In several other studies it was found to have close relationships with yield (Pinter et al., 1983; Olufayo et al., 1996). The physiological background for both the WI and the CWSI is their connection to stomatal conductance, which determines both transpiration and CO_2 uptake. In stepwise MLR models due to the collinearity between the WI and CWSI, the CWSI was mostly excluded from the models. In PLSR models describing single day measurements, the CWSI showed no loadings > |0.1| on the first and second component. Thus, the inclusion of thermal based data and data fusion with spectral information is not recommended.

A similar effect could be observed for the REIP, and it was excluded by the other five spectral indices. Both indices and the REIP are related to traits such as biomass and chlorophyll concentration and seem to contain redundant information. The indices react to absorption in the visible range, while the REIP is affected by a broadening of the major chlorophyll absorption feature centred at 680 nm, which causes a shift in the red edge slope and wavelengths of maximum slope (Cho and Skidmore, 2006). However, taken alone Mistele et al. (2003) found that the REIP has a close relationship to maize yield ($R^2 = 0.80$) in a field trial receiving incremental nitrogen fertilisation from 25 to 220 kg N ha^{-1}.

Independent information on yield was added by the ultrasonic (8 counts) and the laser (9 counts) (Table 4) distance parameters in the MLR models. In the PLSR models, the distance parameters were mainly found on the second component. There were functional differences offered by the distance sensors. The laser distance ranging delivers a clear cut vertical profile of the canopy. Ultrasound reacts to the density of the touched material, it penetrates deeper into the canopy and its reflections are smoothed. The distance sensors offer a high repetition rate of five measurements per second and provides information on the canopy height and the structure of the measured plots. Despite their technically different principles, no differences were found in the loadings of the PLSR models. Thus, it is not necessary to include both devices in the data fusion models.

The set of cultivars grown in the year 2012 represents the German breeding history since 1895. The yields were improved by including dwarf genes and by altering the harvest index towards a higher proportion of grain in the biomass and by increasing biomass growth (White and Wilson, 2006). Breeding progress is associated with shorter canopies and higher yield. The single linear regressions between distance parameters and yield showed negative slopes for the field trials grown in 2012. These trials included both the tall historical cultivars and shorter, younger cultivars. Tall canopies are related to low yields. Conversely, the modern cultivars grown in 2013 have positive slopes, and the taller cultivars were associated with higher yields. This result is also expected when only one cultivar is grown and high stands indicate better soil conditions and are related to higher yields.

4.4. Instability of water status measurements

Spectral water status (WI) and temperature measurements are dependent on mild drought stress due to favourable soil and weather conditions. Rischbeck et al. (2014) assessed the relative leaf water content with a normalised index $(R_{730}\text{-}R_{457})/(R_{730}+R_{457})$ measured at daytime ($R^2 = 0.47$) and by dividing the index by the same index measured at night ($R^2 = 0.54$). They obtained significant relationships under warm or hot and sunny weather conditions, but not on overcast and cool days. The spectral and distance parameters can also be measured on cloudy or overcast days. The fusion of short term WI and medium term drought stress traits (SI, distance parameters) enhances the model flexibility under different weather and drought stress conditions.

4.5. How broad should the models be?

The combination of data across time, years and trials creates more broadly adapted yield models. Weber et al. (2012) recommended combining data from different years to gain broader calibration. However, only data from the same water regime (e.g., mild drought stress) should be combined. Under different water regimes, genotype and environment interactions occur that can decrease the robustness of grain yield models. In this study, a gain in the broadness of the model resulted in a reduction of the prediction quality.

The combination of data from two phenological phases in multiple linear regression analyses can be beneficial for yield prediction. We showed this result by combining data from BBCH 59 and BBCH 85 in the rainfed field trial in 2012. This approach includes more information from the life of the plant. Hansen et al. (2002) collected spectral measurements from spring barley with high yield variability due to the use of two plant densities (150 and 450 plants m^{-2}) and eleven strategies of nitrogen application (0–150 kg total N ha^{-1}). They obtained a coefficient of determination with $R^2 = 0.93$ by combining eight wavebands at 560 nm, 650 nm, 690 nm, 740 nm, 760 nm, 810 nm, 900 nm and 970 nm and 10 derived spectral indices from measurements in the phenological stages BBCH 28, 30, 33 and 40 in PLSR models with grain yield. In the stepwise multiple linear regressions the algorithm preferably selects data that have closer relationships with yield. When combining data from distant phenological phases these data are usually derived from the later measurements. Another shortcoming of MLR models is their failure to address the collinearity that occurs in time series measurements (Nguyen and Lee, 2006). The partial least square model has the advantage of a better inclusion of data from earlier phenological phases independent of collinearity.

In this study, none of the parameters used for yield modelling showed interactions with the cultivar classes. The heterogeneity of cultivars did not substantially affect the quality of yield assessments and the parameters could be used without restrictions for yield modelling. There was no separate calibration of models required for the very heterogeneous classes of cultivars that represent the German breeding history from 1895 until 2007. This is an advantage for possible applications in remote sensing when various cultivars are grown in neighbouring fields and separate calibrations are not feasible.

4.6. Advantages of data fusion

Mouazen et al. (2014) reported an application of data fusion in agronomy. They used multiple linear regressions and artificial neural networks to model the soil water holding capacity (WHC) based on the following parameters: organic carbon, clay content, bulk density, plasticity index and electrical conductivity. The parameters were measured by spectral and mechanical devices attached to tillage instruments carried by a tractor. Using a carrier vehicle, allows different devices to be used simultaneously and data can be allocated to the same positions in the field trials by the use of a GPS system. To the best of our knowledge, data fusion from different sensor sources has not yet been adopted for yield modelling and prediction in field trials. As shown in this study (Table 7), data fusion can improve the quality and robustness of yield predictions compared to models that are only based on spectral data.

The fusion of data from different spectral ranges or different physical sources can be done by multiple linear regressions, partial least square and other methods, such as support vector machines (Montes et al., 2011). Additional improvements can be made when traits physiologically related in different ways to yield are combined.

The calibration of PLSR models is improved when comparing coefficients of determination and root mean square errors by including non-spectral data in the models. However, the improvements could not fully be confirmed by validation. In this study, the combination of spectral data related to the water status, morphology and the nutritional status with plant height information proved to be successful.

Hyperspectral measurements are important for identifying wavelengths relevant for yield. However, as shown by Sharabian et al. (2014) and Hansen and Schjoerring (2003) and this study, there is substantial redundancy in hyperspectral data. Much of the spectral information can be discarded without loss of quality in yield models. Fusing data from different sources can be used for combining relevant spectral bands and non-spectral data based on empirical evaluations and identification of the physiological importance for yield of the selected traits.

In this study, in 2012 data fusion techniques were applied in trials with larger variability in grain yield due to the use of a heterogeneous set of cultivars. In 2013, in the field trials a lower variability of grain yield was observed and more uniform cultivars were grown. Although the quality of yield prediction decreased with lower variability, the data fusion still showed advantages in the quality of prediction.

5. Conclusions and outlook

It can be concluded that partial least square regression models of spring barley grain yield based on spectral indices, distance parameters and the CWSI led to slightly better results than multiple linear regression in yield model calibration but not in model validation. Distance sensors reflect variability in growth height of cultivars and environmental growth conditions, which adds substantial information to yield models. Thermal information as influenced by crop transpiration adds information to a lesser extent. Some redundancy exists with spectral information from bands being sensitive to the water status of the crop (e.g. 970 nm). Regarding the environmental conditions yield modelling of drought-stressed crops is improved under conditions of mild stress compared to severe or non-stress conditions. Under mild stress a differentiation in water status occurs, that leads to variable impacts on growth and crop yield. Assessments at later phenological stages led to improved yield predictions, since less time and fewer weather events occurred between the assessment and the determination of yield.

Data combinations from different phenological phases enhance information from the crops life cycle, and improve yield models. Including data from crops grown on different soils or in different years enables broadly adapted yield models.

It can be stated, that data fusion of hyperspectral, thermal and distance information shows advantages for yield prediction compared to models based on hyperspectral data only. The quality of yield models can be improved regarding the explained variance of

yield (R^2) as well as the root mean square error of prediction. The quality of yield prediction decreased with lower variability of cultivars, but data fusion still showed advantages in the quality of the prediction.

Recent technological developments like reduced weight of hyperspectral and thermal sensors and cameras, improvements in stereoscopic distance measurement and radiometric calibration (Honkavaara et al., 2013) enable to gather spectral, thermal and distance information from unmanned aerial vehicles. Airborne campaigns and satellites used for environmental monitoring are regularly equipped with sensors operating in the visible, near and thermal infrared range (Calderón et al., 2013; Wright et al., 2013; Lucas et al., 2008). Data fusion using multivariate statistical approaches can improve yield modelling in remote sensing and can be applied in precision farming as well as in phenotyping.

Acknowledgements

Special thanks to the team at the experimental station in Dürnast for conducting the field trials and for many technical suggestions for improving the measurements. The authors would like to thank Stephan Haug from the Zentrum Mathematik, Technical University of Munich for his valuable comments regarding statistics. The authors thankfully acknowledge the financial support by the German Federal Ministry of Education and Research (BMBF) for the project 'CROP.SENSe.net' under support code 0315530C.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eja.2016.04.013.

References

Agüera Vega, F., Carvajal Ramírez, F., Pérez Saiz, M., Orgaz Rosúa, F., 2015. Multi-temporal imaging using an unmanned aerial vehicle for monitoring a sunflower crop. Biosyst. Eng. 132, 19–27.

Aparicio, N., Villegas, D., Casadesus, J., Araus, J.L., Royo, C., 2000. Spectral vegetation indices as nondestructive tools for determining durum wheat yield. Agron. J. 92, 83–91.

Bradford, K.J., Hsiao, T.C., 1982. Physiological responses to moderate water stress, in Encyclopedia of Plant Physiology. In: Lange, O.L., et al. (Eds.), Physiological Plant Ecology II, Water Relations and Carbon Assimilation, vol. 12B. Springer-Verlag, New York, Heidelberg, Berlin, pp. 263–324.

Broge, N.H., Leblanc, E., 2000. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote Sens. Environ. 76 (2), 156–172.

Calderón, R., Navas-Cortés, C., Lucena, C., Zarco-Tejada, P.J., 2013. High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of olive using fluorescence, temperature and narrow-band spectral indices. Remote Sens. Environ. 139, 231–245.

Carlson, T.N., Ripley, D.A., 1997. On the relation between NDVI, fractional vegetation cover and leaf area index. Remote Sens. Environ. 62 (3), 241–252.

Cho, M.A., Skidmore, A.K., 2006. A new technique for extracting the red edge position from hyperspectral data: the linear extrapolation method. Remote Sens. Environ. 101, 181–193.

Clevers, J.G., De Jong, S.M., Epema, G.F., Van der Meer, F.D., Bakker, W.H., Skidmore, A.K., Scholte, K.H., 2002. Derivation of the red edge index using the MERIS standard band setting. Int. J. Remote Sens. 23 (16), 3169–3184.

Comar, A., Burger, P., de Solan, B., Baret, B., Daumard, F., Hanocq, J.F., 2012. A semi-automatic system for high throughput phenotyping wheat cultivars in-field conditions: description and first results. Funct. Plant Biol. 39, 914–924.

Deutscher Wetterdienst, 2015. Deutscher Klimaatlas.

Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap. Chapman and Hall, New York, London.

Erdle, K., Mistele, B., Schmidhalter, U., 2011. Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. Field Crops Res. 124, 74–84.

Erdle, K., Mistele, B., Schmidhalter, U., 2013. Spectral high-throughput assessments of phenotypic differences in biomass and nitrogen partitioning during grain filling of wheat under high yielding Western European conditions. Field Crops Res. 141, 16–26.

Fu, Y., Yang, G., Wang, J., Song, X., Feng, H., 2013. Winter wheat biomass estimation based on spectral indices, band depth analysis and partial least squares regression using hyperspectral measurements. Comput. Electron. Agric. 100, 51–59.

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.

Gutierrez-Rodriguez, M., Reynolds, M.P., Escalante-Estrada, J.A., Rodriguez-Gonzalez, M.T., 2004. Association between canopy reflectance indices and yield and physiological traits in bread wheat under drought and well-irrigated conditions. Aust. J. Agric. Res. 55, 1139–1147.

Guyot, G., Baret, F., Major, D.J., 1988. High spectral resolution: determination of spectral shifts between the red and the near infrared. Int. Arch. Photogramm. Remote Sens. 11, 750–760.

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., Strachan, I.B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. Remote Sens. Environ. 90 (3), 337–352.

Hansen, P.M., Schjoerring, J.K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sens. Environ. 86 (4), 542–553.

Hansen, P.M., Jorgensen, J.R., Thomsen, A., 2002. Predicting grain yield and protein content in winter wheat and spring barley using repeated canopy reflectance measurements and partial least squares regression. J. Agric. Sci. 139, 307–318.

Heil, K., Schmidhalter, U., 2012. Characterisation of soil texture variability using the apparent soil electrical conductivity at a highly variable site. Comput. Geosci. 39, 98–110.

Herrmann, I., Pimstein, A., Karnieli, A., Cohen, Y., Alchanatis, V., Bonfil, D.J., 2010. Assessment of leaf area index by the red edge inflection point derived from VENμS bands. In: Proceedings of the 'Hyperspectral 2010 Workshop, Frascati, Italy, 17–19 March 2010 (ESA SP-683, May 2010).

Honkavaara, E., Saari, H., Kaivosoja, J., Pölönen, I., Hakala, T., Litkey, P., Mäkynen, J., Pesonen, L., 2013. Processing and assessment of spectrometric, stereoscopic Imagery collected using a lightweight UAV spectral camera for Precision Agriculture. Remote Sens. 5, 5006–5039, http://dx.doi.org/10.3390/rs5105006.

IBM Corp 2012 IBM SPSS Statistics for Windows Version 21 0. Armonk, NY : IBM Corp.

Idso, S.B., Jackson, R.D., Pinter, P.J., Reginato, R.J., Hatfield, J.L., 1981. Normalizing the stress degree-day parameter for environmental variability. Agric. Meteorol. 24, 45–55.

Jackson, R.D., Idso, S.B., Reginato, R.J., Pinter Jr., P.J., 1981. Canopy temperature as a crop water stress indicator. Water Resour. Res. 17, 1133–1138.

Jones, H.G., 2004. Application of thermal imaging and infrared sensing in plant physiology and ecophysiology. Adv. Bot. Res. 41, 107–163.

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

Laurent, A., Loyce, C., Makowski, D., Pelzer, E., 2015. Using site-specific data to estimate energy crop yield. Environ. Model. Softw. 74, 104–113.

Li, F., Martin, L., Jia, L., Miao, Y., Yu, Z., Kopp, W., Bareth, G., Chen, X., Zhang, F., 2008. Estimating N status of winter wheat using a handheld spectrometer in the North China plain. Field Crops Res. 106 (1), 77–85.

Li, Z., Jin, X., Zhao, C., Wang, J., Xu, X., Yang, G., Li, C., Shen, J., 2015. Estimating wheat yield and quality by coupling the DSSAT-CERES model and proximal remote sensing. Eur. J. Agron. 71, 53–62.

Lichtenthaler, H.K., 1996. Vegetation stress: an Introduction to the stress concept in plants. J. Plant Physiol. 148 (1–2), 4–14.

Lobos, G.A., Matus, I., Rodriguez, A., Romero-Bravo, S., Araus, J.L., del Pozo, A., 2014. Wheat genotypic variability in grain yield and carbon isotope discrimination under mediterranean conditions assessed by spectral reflectance. J. Integr. Plant Biol. 56, 470–479.

Lucas, R., Bunting, P., Paterson, M., Chisholm, L., 2008. Classification of Australian forest communities using aerial photography, CASI and HyMap data. Remote Sens. Environ. 112 (5), 2088–2103.

Marti, J., Bort, J., Slafer, G.A., Araus, J.L., 2007. Can wheat yield be assessed by early measurements of normalized difference vegetation index? Ann. Appl. Biol. 150, 253–257.

Mevik, B.H., Wehrens, R., 2007. The pls package: principal component and partial least squares regression in R. J. Stat. Softw. 18 (2), 1–24.

Mistele, B., Schmidhalter, U., 2008. Spectral measurements of the total aerial N and biomass dry weight in maize using a quadrilateral-view optic. Field Crops Res. 106, 94–103.

Mistele, B., Schmidhalter, U., 2010. Tractor-based quadrilateral spectral reflectance measurements to detect biomass and total aerial nitrogen in winter wheat. Agron. J. 102 (2), 499–506.

Mistele, B., Bredemeier, C., Gutser, R., Schmidhalter, U., 2003. Spectral detection of nitrogen status, biomass and yield of field-grown maize plants. In: Werner, A., Jarfe, A. (Eds.), Programme Book of the Joint Conference of EPCA-EPCLF. Wageningen Academic Publishers, EPCA Berlin, pp. 497–498, 15–19.6.03.

Mistele, B., Gutser, R., Schmidhalter, U., 2004. Validation of Field-Scaled Spectral Measurements of the Nitrogen Status in Winter Wheat. In: Programm book of the joint conference ICPA, Minneapolis, pp. 1187–1195.

Mistele, B., Elsayed, S., Schmidhalter, U., 2012. Assessing water status in wheat under field conditions using laser induced chlorophyll fluorescence and hyperspectral measurements. In: 11th International Conference on Precision Agriculture, Indianapolis, Indiana USA.

Montes, J.M., Technow, F., Dhillon, B.S., Mauch, F., Melchinger, A.E., 2011. High-throughput non-destructive biomass determination during early plant development in maize under field conditions. Field Crops Res. 121, 268–273.
Mouazen, A.M., Alhwaimel, S.A., Kuang, B., Wayne, T., 2014. Multiple online soil sensors and data fusion approach for delineation of water holding capacity zones for site-specific irrigation. Soil Tillage Res. 143, 95–105.
Nguyen, T.A., Lee, B.W., 2006. Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression. Eur. J. Agron. 24, 349–356.
Olufayo, A., Baldy, C., Ruelle, P., 1996. Sorghum yield, water use and canopy temperatures under different levels of irrigation. Agric. Water Manage. 30, 77–90.
Peñuelas, J., Filella, I., Serrano, L., 1993. The reflectance at the 950–970 nm region as an indicator of plant water status. Int. J. Remote Sens. 14, 1887–1905.
Pinter, P.J., Fry, K.E., Guinn, G., Mauney, J.R., 1983. Infrared thermometry: a remote sensing technique for predicting yield in water-stressed cotton. Agric. Water Manage. 6, 385–395.
R Core Team, 2014 R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
Rebetzke, G.J., Chenu, K., Biddulph, B., Moeller, C., Deery, D.M., Rattey, A.R., Bennett, D., Barrett-Lennard, E.G., Mayer, J.E., 2013. A multisite managed environment facility for targeted trait and germplasm phenotyping. Funct. Plant Biol. 40 (1), 1–13, http://dx.doi.org/10.1071/FP12180.
Rischbeck, P., Baresel, P., Elsayed, S., Mistele, B., Schmidhalter, U., 2014. Development of a diurnal dehydration index for spring barley phenotyping. Funct. Plant Biol. 41 (12), 1249–1260, http://dx.doi.org/10.1071/FP14069, FP14069 (published online).
Overview and Recent Advances in Partial Least Squares. In Saunders, C., Grobelnik, M., Gunn, S., Shawe-Taylor, J., 2006. editors. 'Subspace, Latent Structure and Feature Selection Techniques', Lecture Notes in Computer Science, 34–51.
Rostamza, M., Chaichi, M.R., Jahansouz, M.R., Alimadadi, A., 2011. Forage quality, water use and nitrogen utilization efficiencies of pearl millet (*Pennisetum americanum* L.) grown under different soil moisture and nitrogen levels. Agric. Water Manage. 98, 1607–1614.
Royo, C., Aparicio, N., Villegas, D., Casadesus, J., Monneveux, P., Araus, J.L., 2003. Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean conditions. Int. J. Remote Sens. 24, 4403–4419.
Schmidhalter, U., Evéquoz, M., Camp, K.H., Studer, C., 1998. Sequence of drought response of maize seedlings in drying soil. Physiol. Plant. 104, 159–168.
Schmidhalter, U., Glas, J., Heigl, R., Manhart, R., Wiesent, S., Gutser, R., Neudecker, E., 2001. Application and testing of a crop scanning instrument – field experiments with reduced crop width, tall maize plants and monitoring of cereal yield. In: 3rd. European Conference for Precision Agriculture, Montpellier, pp. 953–958.
Serrano, L., Gonzalez-Flor, C., Gorchs, G., 2011. Assessment of grape yield and composition using the reflectance based water index in mediterranean rainfed vineyards. Remote Sens. Environ. 118, 249–258.
Sharabian, V.R., Noguchi, N., Ishi, K., 2014. Significant wavelengths for prediction of winter wheat growth status and grain yield using multivariate analysis engineering in agriculture. Environ. Food 7, 14–21.
Siddique, M.R., Hamid, A., Islam, M.S., 2000. Drought stress effects on water relations of wheat. Bot. Bull. Acad. Sin. 41, 35–39.
Wall, G.W., Garcia, R.L., Wechsung, F., Kimball, B.A., 2011. Elevated atmospheric CO_2 and drought effects on leaf gas exchange properties of barley agriculture. Ecosyst. Environ. 144, 390–404.
Weber, V.S., Araus, J.L., Cairns, J.E., Sanchez, C., Melchinger, A.E., Orsini, E., 2012. Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes. Field Crops Res. 128, 82–90.
White, E.M., Wilson, F.E.A., 2006. Responses of grain yield, biomass and harvest index and their rates of genetic progress to nitrogen availability in ten winter wheat varieties. Irish J. Agric. Food Res. 45, 85–101.
White, J.W., Andrade-Sanchez, P., Gore, M.A., Bronson, K.F., Coffelt, T.A., Conley, M.M., Feldmann, K.A., French, A.N., Heun, J.T., Hunsaker, D.J., Jenks, M.A., Kimball, B.A., Roth, R.L., Strand, R.J., Thorp, K.R., Wall, G.W., Wang, G., 2012. Field-based phenomics for plant genetics research. Field Crops Res. 133, 101–112.
Winterhalter, L., Mistele, B., Jampatong, S., Schmidhalter, U., 2011. High throughput phenotyping of canopy water mass and canopy temperature in well-watered and drought stressed tropical maize hybrids in the vegetative stage. Eur. J. Agron. 35, 22–32.
Winterhalter, L., Mistele, B., Schmidhalter, U., 2013. Evaluation of active and passive sensor systems in the field to phenotype maize hybrids with high-throughput. Field Crops Res. 154, 236–245.
Wright, R., Lucey, P., Crites, S., Horton, K., Wood, M., Garbeil, H., 2013. BBM/EM design of the thermal hyperspectral imager: an instrument for remote sensing of earth's surface, atmosphere and ocean, from a microsatellite platform.
Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H., Sampson, P.H., 2001. Scaling-up and model inversion methods with narrow-band optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. IEEE Trans. Geosci. Remote Sens. 39, 1491–1507.
Zarco-Tejada, P.J., Berjón, A., López-Lozano, R., Miller, J.R., Martín, P., Cachorro, V., González, M.R., de Frutos, A., 2005. Assessing vineyard condition with hyperspectral indices: leaf and canopy reflectance simulation in a row structured discontinuous canopy. Remote Sens. Environ. 99, 271–287.
Zhao, T., Komatsuzaki, M., Okamoto, H., Sakai, K., 2010. Cover crop nutrient and biomass assessment system using portable hyperspectral camera and laser distance sensor. Eng. Agric. Environ. Food 3, 105–112.
Zhao, Y., Chen, X.P., Cui, Z.L., Lobell, D.B., 2015. Using satellite remote sensing to understand maize yield gaps in the North China Plain. Field Crops Res. 138, 31–42.