Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes

V.S. Weber\textsuperscript{a}, J.L. Araus\textsuperscript{b,c}, J.E. Cairns\textsuperscript{c}, C. Sanchez\textsuperscript{c}, A.E. Melching\textsuperscript{a}, E. Orsini\textsuperscript{a,∗}

\textsuperscript{a} Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Fruwirthstr. 21, D-70593 Stuttgart, Germany
\textsuperscript{b} Unitat de Fisiologia Vegetal, Faculat de Biologia, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain
\textsuperscript{c} Global Maize Program, International Maize and Wheat Improvement Center (CIMMYT), Mexico D.F., Mexico

\section*{A R T I C L E   I N F O}

\textbf{Article history:}  
Received 16 August 2011  
Received in revised form 5 December 2011  
Accepted 22 December 2011

\textbf{Keywords:}  
Maize  
Leaf and canopy spectral reflectance  
PLSR regression  
Grain yield  
Anthesis and milk-grain stage

\section*{A B S T R A C T}

The ability to accurately estimate grain yield using spectral reflectance measurements prior harvest could be used to reduce phenotyping time and costs. In this study, grain yield of 300 maize testcrosses grown under different water and temperature regimes in the dry season 2010 was predicted using spectral reflectance (495–1853 nm) of both leaves and canopy measured between tassel emergence until milk-grain stage. Partial least square regression (PLSR) was used for data analysis. Coefficients of determination ($R^2$) between predicted and actual grain yield were highest for measurements conducted at anthesis and milk-grain stage, explaining at maximum 23% and 40% of the genotypic variation in grain yield after validation, respectively. PLSR models explained a higher proportion of the genetic variation in grain yield under drought stress compared to well-watered conditions. The association between predicted and actual grain yield was stronger in spectral reflectance measurements taken at the leaf level compared to canopy level. By combining the most predictive PLSR models across trials, at maximum of 40% of the variation in grain yield could be explained in each trial with a relative efficiency of selection of 0.88 and 0.68 using leaf and canopy reflectance, respectively. The most relevant wavelengths for predicting grain yield were associated with photosynthetic capacity (495–680 nm), red inflection point (680–780 nm) and plant water status (900 nm, 970, 1450 nm, 1150–1260 nm, and 1520–1540 nm). Additional wavelengths based on leaf (800 nm, 1000, and 1260–1830 nm) and canopy (988–999 nm and 1430–1640 nm) reflectance of unknown physiological relevance were also identified for prediction of grain yield. Caution must be exercised before integrating our spectral reflectance approach into a breeding program because this is a pilot study based on a single location and season.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Drought stress is a major constraint of raised maize (\textit{Zea mays} L.) yields (\textit{FAO, 2010}). Projections of decreasing precipitation and increasing evaporative demand by 2050 will further exacerbate yield losses due to drought in many raised maize areas (\textit{IPCC, 2007}). To offset predicted yield losses, the development of improved breeding pipelines is a priority. The ability to accurately predict grain yield under drought stress could be used to select superior lines at anthesis for crossing, thereby increasing selection gain while reducing phenotyping costs within the early stages of a breeding program. Reduced anthesis-silking interval (ASI) combined with selection for grain yield has been resulted in gains of up to 144 kg ha\textsuperscript{-1} yr\textsuperscript{-1} under drought stress (\textit{Bolaños and Edmeades, 1993; Edmeades et al., 1999}). Genotypes with a short ASI have a higher efficiency of biomass partitioning to ear and tassels at flowering than those with a long ASI (\textit{Edmeades et al., 1999}). However, repeated selection for reduced ASI within elite germplasm has reduced genotypic variation for this trait (\textit{Monneveux et al., 2008}). Thus, the development of novel phenotyping tools for the drought screening are now required. Besides ASI, traits related to high photosynthetic capacity (i.e., chlorophyll content) and plant water content (i.e., stomatal conductance) were reported to contribute to higher grain yields under drought stress (\textit{Araus et al., 2002}). Several devices have been developed to indirectly measure those traits, such as SPAD-meter for chlorophyll content or leaf...
Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Trial codes</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E-DS</td>
<td>L-DS</td>
<td>E-WW</td>
<td>L-WW</td>
<td>L-WW+H</td>
</tr>
<tr>
<td>AD (days after sowing)</td>
<td>81 ± 2.6</td>
<td>84 ± 2.8</td>
<td>79 ± 2.9</td>
<td>80 ± 2.6</td>
<td>78 ± 3.0</td>
</tr>
<tr>
<td>SD (days after sowing)</td>
<td>83 ± 2.5</td>
<td>85 ± 2.5</td>
<td>82 ± 2.5</td>
<td>83 ± 2.6</td>
<td>80 ± 3.0</td>
</tr>
<tr>
<td>Total precipitation (mm)</td>
<td>124.2</td>
<td>124.2</td>
<td>124.2</td>
<td>124.2</td>
<td>123.5</td>
</tr>
<tr>
<td>Total irrigation (mm ha⁻¹)</td>
<td>521.9</td>
<td>489.3</td>
<td>848.1</td>
<td>935.1</td>
<td>841.3</td>
</tr>
<tr>
<td>Minimum temperature (°C)</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Maximum temperature (°C)</td>
<td>37 (8)</td>
<td>37 (8)</td>
<td>37 (8)</td>
<td>37 (8)</td>
<td>40 (52)</td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td>56.5 ± 11.5</td>
<td>56.4 ± 11.5</td>
<td>56.5 ± 11.5</td>
<td>56.5 ± 11.5</td>
<td>47.6 ± 11.5</td>
</tr>
</tbody>
</table>

* E and L refer to early and late maturity group, respectively; DS, WW and WW + H indicate the drought, well-watered and well-watered plus heat condition, respectively.

* Average ± standard deviation.

* Maximum temperature and number of days (in parenthesis) with an air temperature higher than 36 °C, from anthesis to harvest.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A promising technique, applicable to a large number of plants, is the measurement of leaf and canopy spectral reflectance which can give information on traits related to grain yield such as: photosynthetic capacity, aboveground biomass and plant water content. Chlorophyll content (Sims and Gamon, 2002), aboveground biomass (Liu et al., 2010) and plant water status (Peñuelas et al., 1993; Linke et al., 2008) have been estimated using spectral indices, developed on the basis of ratios or differences in the reflectance at given wavelengths. To date, spectral indices for grain yield have only been used in wheat, with spectral reflectance more closely associated with grain yield under well-watered conditions compared to drought stress (Aparicio et al., 2000; Gutierrez-Rodriguez et al., 2004; Gutiérrez et al., 2010). In bread and durum wheat grown across a wide range of water regimes and years, more than 90% of grain yield variation could be explained using canopy reflectance measured at 550 nm or spectral indices of wavelengths from 850 to 970 nm (Royo et al., 2003; Gutiérrez et al., 2010). In maize grown under well-watered conditions, stepwise regressions combining six different wavelengths were able to explain over 95% of the variation in grain yield (Osborne et al., 2002). However, the wavelengths used to estimate grain yield varied with growth stage and the environmental conditions. Further, regression models are prone to over-fitting as the number of wavelengths included is increased (Thenkabail et al., 2000). Partial least square regression (PLSR) decomposes the variability of the spectrum matrix into a small number of factors to describe the response variable (Naes et al., 2004) and may provide a more suitable alternative to stepwise regression for the identification of key wavelengths within the entire spectrum. In durum wheat, prediction of grain yield was generally stronger and more robust based on PLSR models than with previously assayed spectral indices (Ferrio et al., 2005). At present, PLSR models for grain yield under different water regimes have only been developed in durum wheat for single environments (Ferrio et al., 2004, 2005). To be implemented into a breeding program, however, PLSR models have to predict grain yield with high heritability independently of the environments wherein they are applied. To the best of our knowledge, such information is entirely lacking in maize.

The goal of this study was to evaluate the ability of PLSR models to predict grain yield of maize under drought, heat and well-watered conditions. The specific objectives were to compare the potential of spectral reflectance at both the leaf and canopy level to predict grain yield, and identify the most relevant traits and the related wavelengths for predicting grain yield across different water conditions. Spectral indices and regions previously shown to be correlated with the water status, photosynthetic capacity and biomass of several plant species, were used in this study. Briefly, regions included were (i) from 400 to 700 nm (photosynthetic active radiation i.e., PAR, Cer and Bornman, 1990; Aparicio et al., 2000), (ii) from 680 to 780 nm (the red inflection point i.e., REIP, Filella and Peñuelas, 1994), (iii) from 750 to 800 nm (brown pigments, Peñuelas and Filella, 1998), (iv) at 900 and 970 nm (water index, Peñuelas et al., 1993, 1997), (v) from 1150 to 1260 nm and from 1520 to 1540 nm (water content, Sims and Gamon, 2003) and (vi) at 1450 nm (water absorbance, Sims and Gamon, 2003).

2. Materials and methods

2.1. Experimental setup

A total of 300 single cross maize hybrids were used in this study, comprised of 292 lines, chosen to represent the genetic diversity within tropical and subtropical maize germplasm, crossed to the common tester CML-312-SR and eight hybrids of the same set as checks. Hybrids were separated into two maturity groups, early (E, n = 150) and late (L, n = 150) based on information derived from previous experiments. A total of five trials were sown comprising of three different water regimes; well-watered (E-WW and L-WW), drought stress imposed at anthesis (E-DS and L-DS) and heat stress (L-WW+H). Trials were planted at the CIMMYT experimental station at Tlaltizapán, Morelos, Mexico, during the dry season of 2010. Well-watered and drought trials were sown in November 2009 and the well-watered trial with heat stress was sown in February 2010 (the early maturity group was not included in this treatment). Drought stress was imposed at flowering by stopping irrigation two weeks prior to anthesis until 30 days after anthesis (Bänziger et al., 2000). Weather data were recorded at an adjacent weather station and are presented in Table 1. Trials were laid out as alpha lattice designs with two replications in 2 row 5 m long plots, with a plot size of 7.5 m² and a plant density of 66,000 plants ha⁻¹. Appropriate fertilization was provided according to common agronomical practices. Weeds and diseases were controlled, when necessary. Plots were harvested manually, ear yield was determined in the whole plot at 10% moisture level and a shelling percentage of 80% was assumed (Betrán et al., 2003).

2.2. Reflectance measurements

Leaf and canopy reflectance were measured using a spectroradiometer (HR-1024, Spectra Vista Corporation, NY, USA). The spectroradiometer measures radiance in 1024 continuous bands with a sampling interval of 3.5 nm from 350 to 1000 nm and of 16 nm from 1000 to 2500 nm, thereby covering the ultra-violet, visible, and near infra-red region of the spectra. The reflectance...
spectrum was calculated in real time as the ratio between the reflected and the incident spectra, where the incident spectrum was periodically obtained from the light reflected by a white reference panel.

Leaf reflectance was measured in 100 randomly selected genotypes within the early and late maturity groups using the reference probe including a light source (Spectra Vista Corporation, NY, USA). Six plants were randomly selected per plot and tagged. Measurements were taken in the centre of the ear leaf blade in each plant. The same leaves were used for each measurement. Measurements were taken at anthesis and silking in the early maturity group and four times from anthesis until the beginning of grain filling (milk-grain stage) in the late maturity group. In the late maturity group, measurements were also taken at tassel emergence.

Canopy reflectance for all \( n = 300 \) genotypes was measured per plot with the spectroradiometer fitted with a 25° field-of-view optic placed about 50 cm above the top of the canopy at a 50° angle in both field replications. Three measurements were taken during tassel emergence until dough stage. Measurements were performed in the 4 h interval around noon on cloudless days.

2.3. Statistical analysis

The relationship between leaf and canopy reflectance spectra and grain yield were determined using partial least square regression (PLSR). The spectra were scaled and the first derivative of the mean spectra per plot was calculated for individual time points. The first derivative was calculated as the difference between the reflectance at one wavelength and the reflectance at the next wavelength. Each second wavelength was excluded because wavelengths close to each other are known to be highly correlated (Naes et al., 2004). In total 326 spectral bands from 494.6 to 1852.9 nm were included in the models; spectra outside these limits were not included as they were highly variable and less sensitive. The number of components used in PLSR was chosen according to the minimum root mean square error of prediction (RMSE) calculated by leave one out validation as described by Naes et al. (2004). For comparing the coefficient of determination \( R^2 \) for grain yield of PLSR models for individual trials, the relative RMSE \((rRMSE)\) was calculated as the ratio between RMSE and the standard error of grain yield, as previously described (Ferrio et al., 2005).

For estimating the variance components of actual and predicted grain yield using canopy and leaf reflectance, an alpha lattice analysis of variance was employed using a mixed-model approach. The effective error mean square of actual grain yield was calculated according to Cochran and Cox (1992).

Heritability \( (h^2) \) of actual and predicted grain yield was calculated as the ratio between the genotypic and the phenotypic variance (Hallauer et al., 2010). For predicted grain yield using the combined model, \( h^2 \) was calculated for each trial and then averaged. The relative efficiency of selection \( (RE) \) of grain yield predicted with leaf or canopy spectral reflectance was estimated according to Falconer and Mackay (1996) assuming that the selection intensity is the same for predicted and actual grain yield:

\[ RE = \frac{r_g h_p}{h_h} = \frac{r_g}{h_h} \]

where \( r_g \) and \( r_p \) are the genotypic and phenotypic correlation, respectively, between the predicted and actual grain yield, \( h_p \) is the square root of the \( h^2 \) of the predicted grain yield and \( h_h \) is the square root of the \( h^2 \) of the actual grain yield. The \( r_g \) was calculated as the ratio between \( r_p \) and the product between \( h_p \) and \( h_h \) (Cooper et al., 1996).

For the grain yield predicted with the combined models, the \( r_g \) and \( RE \) were calculated for each trial and are given as an average.

To determine, whether grain yield prediction is more robust with increasing number of trials and genotypes, data collected under several water regimes were combined into one model. PLSR models with a high \( R^2 \) for grain yield, low RMSE, and \( h^2 \) above 0.44, were selected to build a combined model used for grain yield prediction across field trials. Loading values of each wavelength of the first principle component of the combined model were used to identify the most relevant spectral wavelengths for prediction of grain yield and its related traits assessed with spectral reflectance measurements. To estimate, whether a model restricted to relevant spectral wavelengths is able to predict grain yield reliably, an arbitrary threshold loading value below \(-0.08\) and above \(0.08\) was applied to select and combine wavelengths into one PLSR model.

All statistical analyses were conducted in R (version 2.9.0, R Development Core Team, 2009). PLSR analyses was performed using the pls package (Mevik and Wehrens, 2007) and the variance components of actual and predicted grain yield were estimated using the ASREML package (Butler et al., 2007).

3. Results

3.1. Genotypic performance under well-watered and drought stress conditions

Grain yield was significantly reduced under drought relative to the well-watered conditions \((P < 0.001)\). Under drought stress mean grain yield was reduced, by almost 50% relative to the well-watered control, to around 4 t ha\(^{-1}\) (Table 2). Mean grain yield under heat stress was 6.9 t ha\(^{-1}\) and was significantly less than grain yield under optimal temperature, well-watered conditions \((P < 0.001)\). There was no significant difference in grain yield between the early and the late maturity group under both well-watered and drought stress conditions. Differences in phenology between the two maturity groups under both well-watered and drought stress conditions were small. Genotypic variance \((\sigma^2_g)\) for grain yield was significant in all trials and it was lower under drought stress compared to well-watered conditions and higher in the late compared to the early maturity group. The highest effective error mean square \((\sigma^2_\varepsilon)\) was observed in the WW + H trial. Heritability of grain yield ranged from 0.80 to 0.89 under well-watered conditions, and it was below 0.77 under drought stress. The genotype-by-treatment of grain yield was significant \((P < 0.001, \text{data not shown})\).

3.2. Spectral reflectance of leaves and canopies

Leaf and canopy reflectance were highest in the range between 750 and 1300 nm, with mean leaf reflectance approximately 20% higher than canopy reflectance within this region (Fig. 3). Differences outside this range were marginal. Leaf reflectance was higher under drought stress than under well-watered conditions. Canopy reflectance was highest in the L-WW trial compared to the L-DS, E-WW and E-DS trials. The highest leaf and canopy reflectance in the range of 750–1300 nm was observed within the L WW + H trial.

Both leaf and canopy reflectance was low within the photosynthetic active radiation (PAR) (from 400 to 700 nm). Within this range no differences were observed between trials for canopy reflectance while at the leaf level, reflectance in the L-DS trial was approximately 10 percentage points higher relative to other trials. The increase in canopy and leaf reflectance from red to near-infrared (REIP) was higher under well-watered conditions and shifted towards the near-infrared region.

Both leaf and canopy reflectance decreased after 1332 nm, reaching a minimum at 1450 nm. In this range differences between the early and the late maturity group were marginal for all trials. Within the region 1450–1853 nm no differences were observed.
in canopy reflectance between trials. Leaf reflectance within this range was approximately 5 percentage points higher in the L-DS trial compared to all other trials.

3.3. Calibration performance and validation of models

PLSR models based on leaf reflectance explained between 14.2% and 68.5% of the variability in grain yield (Table 3). The $R^2$ of grain yield was higher under drought stress than under well-watered conditions. The highest $R^2$ (60.8%) of grain yield under well-watered conditions was observed with measurements at anthesis while under drought stress, the highest $R^2$ (68.5%) was observed with measurements taken at the milk-grain stage. After validation the $R^2$ was reduced to between 22.8% and 39.8%, respectively. The RMSE of calibration ranged from 0.47 to 1.38 t ha$^{-1}$, whereas $r$RMSE varied between 6.72 and 13.44. Furthermore, RMSE and rRMSE were larger under well-watered conditions compared to drought stress conditions. For leaf reflectance, $h^2$ and RE of predicted grain yield ranged from 0.10 to 0.77 and 0.43 to 1.11, respectively.

PLSR models based on canopy reflectance explained between 1.8% and 43.5% of the variability in grain yield. The $R^2$ of grain yield under drought stress was again higher than under well-watered conditions. Under well-watered conditions the highest $R^2$ (33.8%) for grain yield was obtained with measurements taken at anthesis while under drought stress the highest $R^2$ (43.5%) was associated with measurements taken at the milk-grain stage. After validation the $R^2$ was reduced to between 7.9 and 38.1%, respectively. RMSE of calibration ranged from 0.62 to 1.51 t ha$^{-1}$, whereas RMSE varied between 9.68 and 15.12. RMSE and rRMSE were larger under well-watered conditions compared to drought stress conditions. Heritability and RE of predicted grain yield ranged from 0.11 to 0.69 and 0.15 to 0.88, respectively.

PLSR models with the greatest ability to predict grain yield were selected based on the highest $R^2$, $h^2$, and low RMSE. A total of 3 and 4 models (highlighted in bold, Table 3) were selected to build a combined model for leaf and canopy reflectance measurements, respectively, collected over different water regimes. The combined PLSR model accounted for 82.0% and 79.2% of the variation in grain yield at the leaf and canopy level, respectively. When the combined model was used to predict grain yield within individual trial, RE for PLSR models equalled 0.88 and 0.68 for leaf and canopy reflectance, respectively. A combined model was also calculated for all measurements conducted at anthesis. The $R^2$ of calibration of this model was 71.4% based on leaf and 68.2% based on canopy reflectance.

For all PLSR models based on leaf or canopy reflectance, $R^2$ decreased after cross validation while RMSE and rRMSE increased. In particular, the reduction in $R^2$ after cross validation was larger for individual models compared to the combined model while the increase of RMSE and rRMSE was comparable among all models.

The correlation coefficient between actual and predicted grain yield using combined models was 0.91 and 0.89 for leaf and canopy measurements, respectively (Fig. 2). Within individual trials predicted grain yield was positively correlated with actual grain yield and explained between 30 and 40% of grain yield variation based on leaf reflectance and between 19 and 34% based on canopy reflectance. Thus, both individual and combined PLSR models explained less than 40% of grain yield variation in individual trials.

3.4. Main spectral wavelengths related to grain yield

Selected wavelengths (i.e. above and below the loading value of 0.08 and –0.08, Fig. 3) of the combined model coincided with those selected in each single reflectance (results not shown). To evaluate the ability of selected peaks to predict grain yield, the combined PLSR model was recalculated using only selected wavelengths. While the number of wavelengths included in the PLSR model was one third less than the model using all wavelengths, the correlation coefficient between predicted and actual grain yield decreased only slightly (leaf reflectance, $r = 0.89$; canopy reflectance, $r = 0.83$) (Fig. 2). All correlations between predicted and actual grain yield within individual trials were significant.

4. Discussion

4.1. Leaf and canopy reflectance under different water regimes

The mean spectra for both leaf and canopy reflectance (from 495 to 1853 nm) are similar to those previously reported for tree and grass species (Carter, 1991), and bread and durum wheat under both drought stress and well-watered conditions (Linke et al., 2008). In this study the greatest differences between trials for both the leaf and canopy reflectance were observed within 750–1300 nm. Within this range, a low reflectance between 750 and 800 nm indicates a lower brown pigment content and thus less leaf senescence (Peñuelas and Filella, 1998). Further, a low reflectance in the range of the water index has been reported to indicate a better plant water status (Peñuelas et al., 1993, 1997). Moreover, a low canopy reflectance between 1150 and 1260 nm has been reported to be associated with reduced stomatal conductance and high relative water content in a wide range of species (Sims and Gamon, 2003). In the current study, the leaf reflectance in the ranges of brown pigments, the water index and the reflectance between 1150 and 1260 nm were higher under drought stress compared to well-watered conditions. This suggests an association between high leaf reflectance and limited leaf water content, and agrees with the results of Linke et al. (2008) that leaf reflectance of bread and durum wheat within the range of 350–2500 nm was higher under drought stress conditions compared to well-watered conditions. Thus, the reflectance measurements in the range between 750 and 1300 nm can be used to detect differences in leaf/plant water status of genotypes under well-watered and drought stress conditions.

Within the spectral range between 750 and 1300 nm reflectance was found to be higher under drought stress conditions compared to well-watered conditions at the leaf level only. This may be related to the lower water content of leaf tissues under drought stress (Linke et al., 2008). At the canopy level, reflectance was higher in the L-WW trial relative to E-WW, L-DS and E-DS trials. The different results of both the comparison between canopy and
leaf reflectance and between maturity groups under well-watered conditions may, in part, be a result of environmental noise when taking the measurements above the canopy including differences in light intensity and leaf orientation. Under drought stress, the leaf blade rolls to reduce water loss, resulting in a larger fraction of the sunlight absorbed by the soil and surrounding environment. In contrast, when measurements are taken at the leaf level, the sensor is in direct contact with the leaf surface thereby reducing environmental noise. Previously, Hütte (1988) proposed a method to correct spectral reflectance errors introduced by the background soil, however, to date, this methodology has only been applied in correcting normal difference vegetation index (NDVI) (Royo et al., 2003).

Under heat stress (WW + H) the leaf and canopy reflectance from 750 to 1300 nm were higher than reflectance in the other trials. Higher leaf reflectance under heat stress has previously been associated with the deposition of epicuticular waxes on the leaf surface, a strategy to limit water loss from plant cells while avoiding excess radiation (Febro et al., 1998; Shepherd and Griffiths, 2006).

4.2. Maize grain yield prediction based on spectral reflectance models

For individual PLSR models, the highest $R^2$ for grain yield for leaf and canopy reflectance were both reached at the milk-grain stage. Similarly, the strongest association between spectral indices measured at the canopy level and grain yield was previously identified at the heading and milk-grain stage in bread and durum wheat (Aparicio et al., 2000; Royo et al., 2003; Babar et al., 2006a,b). In our study, $R^2$ for grain yield for leaf measurements taken at anthesis was higher than that for canopy measurements taken at the milk-grain stage. However, $R^2$ for grain yield decreased to less than 40% after validation for leaf and canopy reflectance with $RE < 1$. Thus, the leaf reflectance measurements could be used as a rough screening tool to identify superior genotypes already at anthesis. However, careful consideration must be taken into account as more than 60% of grain yield variation could not be explained with individual spectral measurements.

<table>
<thead>
<tr>
<th>Model$^a$</th>
<th>Growth stage</th>
<th>$m^b$</th>
<th>$N^c$</th>
<th>Calibration$^d$</th>
<th>$h^d$</th>
<th>$RE^d$</th>
<th>$R^2$ (%)</th>
<th>Validation$^e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-DS</td>
<td>Anthesis</td>
<td>1</td>
<td>98</td>
<td>49.3 0.63</td>
<td>10.48</td>
<td>0.44</td>
<td>0.92</td>
<td>22.5 1.12</td>
</tr>
<tr>
<td>E-DS</td>
<td>Silking</td>
<td>2</td>
<td>95</td>
<td>61.2 0.55</td>
<td>9.12</td>
<td>0.28</td>
<td>1.04</td>
<td>20.3 1.14</td>
</tr>
<tr>
<td>L-DS</td>
<td>Anthesis</td>
<td>1</td>
<td>100</td>
<td>48.1 0.61</td>
<td>8.72</td>
<td>0.44</td>
<td>0.92</td>
<td>6.5 0.82</td>
</tr>
<tr>
<td>L-DS</td>
<td>Silking</td>
<td>2</td>
<td>100</td>
<td>35.8 0.68</td>
<td>7.68</td>
<td>0.10</td>
<td>0.80</td>
<td>14.7 0.78</td>
</tr>
<tr>
<td>L-DS</td>
<td>Silking</td>
<td>3</td>
<td>100</td>
<td>40.2 0.66</td>
<td>9.36</td>
<td>0.24</td>
<td>0.84</td>
<td>14.4 0.78</td>
</tr>
<tr>
<td>L-DS</td>
<td>Silking</td>
<td>4</td>
<td>100</td>
<td>68.5 0.47</td>
<td>6.72</td>
<td>0.51</td>
<td>1.11</td>
<td>39.8 0.64</td>
</tr>
<tr>
<td>E-WW</td>
<td>Anthesis</td>
<td>1</td>
<td>100</td>
<td>54.0 0.67</td>
<td>7.52</td>
<td>0.49</td>
<td>0.90</td>
<td>11.9 1.34</td>
</tr>
<tr>
<td>E-WW</td>
<td>Silking</td>
<td>2</td>
<td>74</td>
<td>39.5 0.73</td>
<td>8.16</td>
<td>0.65</td>
<td>0.78</td>
<td>12.1 1.57</td>
</tr>
<tr>
<td>L-DS</td>
<td>Anthesis</td>
<td>1</td>
<td>100</td>
<td>52.3 1.00</td>
<td>10.00</td>
<td>0.77</td>
<td>0.81</td>
<td>13.4 1.34</td>
</tr>
<tr>
<td>L-DS</td>
<td>Silking</td>
<td>2</td>
<td>100</td>
<td>24.4 1.26</td>
<td>12.66</td>
<td>0.57</td>
<td>0.55</td>
<td>7.5 1.38</td>
</tr>
<tr>
<td>L-DS</td>
<td>Milk-grain stage</td>
<td>3</td>
<td>100</td>
<td>41.9 1.10</td>
<td>11.04</td>
<td>0.48</td>
<td>0.73</td>
<td>5.2 1.40</td>
</tr>
<tr>
<td>L-DS</td>
<td>Milk-grain stage</td>
<td>4</td>
<td>100</td>
<td>14.2 1.34</td>
<td>13.44</td>
<td>0.25</td>
<td>0.43</td>
<td>0.7 1.44</td>
</tr>
<tr>
<td>L-DS</td>
<td>Tassel emergence</td>
<td>1</td>
<td>100</td>
<td>25.6 1.38</td>
<td>10.64</td>
<td>0.58</td>
<td>0.64</td>
<td>12.3 1.50</td>
</tr>
<tr>
<td>L-DS + H</td>
<td>Tassel emergence</td>
<td>2</td>
<td>100</td>
<td>28.8 1.35</td>
<td>10.40</td>
<td>0.56</td>
<td>0.68</td>
<td>14.5 1.46</td>
</tr>
<tr>
<td>L-DS + H</td>
<td>Anthesis</td>
<td>3</td>
<td>100</td>
<td>60.8 1.00</td>
<td>7.68</td>
<td>0.56</td>
<td>0.98</td>
<td>22.8 1.40</td>
</tr>
<tr>
<td>L-DS + H</td>
<td>Silking</td>
<td>4</td>
<td>100</td>
<td>45.1 1.19</td>
<td>9.12</td>
<td>0.33</td>
<td>0.84</td>
<td>23.0 1.40</td>
</tr>
<tr>
<td>Combined model</td>
<td>Anthesis to milk-grain stage</td>
<td>498</td>
<td>15</td>
<td>82.0 1.01</td>
<td>7.76</td>
<td>0.41(0.11)$^f$</td>
<td>0.88(0.22)$^f$</td>
<td>70.7 1.28</td>
</tr>
<tr>
<td>Combined model</td>
<td>Anthesis</td>
<td>498</td>
<td>13</td>
<td>71.4 1.27</td>
<td>9.76</td>
<td>0.26(0.19)$^f$</td>
<td>0.80(0.06)$^f$</td>
<td>52.0 1.64</td>
</tr>
</tbody>
</table>

---

$^a$ F and L refer to early and late maturity group, respectively; DS, WW and WW + H indicate the drought, well-watered and well-watered plus heat condition, respectively.

$^b$ m: measurement in time; N: number of samples.

$^c$ RMSE: root mean square error, and $RMSE$: ratio between RMSE and the standard error of grain yield.

$^d$ $h^2$: heritability of predicted grain yield; RE: relative efficiency of selection based on predicted grain yield.

$^e$ Average and standard deviation of $h^2$ and $RE$ of predicted grain yield based on the combined model.
For breeding programs, the ability to predict lines with superior grain yield at anthesis would allow the number of crosses to be reduced. Canopy-based measurements will encompass crop performance at the organizational level, providing a more integrative measure of plant performance (Araus et al., 2008). In general, the relationship between predicted and actual grain yield was higher in spectral reflectance measured at the leaf level rather than the canopy level. The weaker association between canopy reflectance and grain yield may be related to environmental background noise while measuring the canopy reflectance as discussed above, and the fact that the canopy reflectance is saturated for values of leaf area index higher than 3 (Royo et al., 2003). For individual PLSR models of leaf and canopy, we observed a higher $R^2$ and $h^2$, together with a reduced rRMSE and RMSE under drought stress conditions compared to well-watered conditions. A similar trend was previously reported for canopy reflectance measured in irrigated and rainfed trials in durum wheat (Ferro et al., 2005). Together, these results suggest that reflectance measurements are of greater practical value under drought stress conditions than under well-watered conditions.

The PLSR model combined across trials and water regimes showed stronger relationships with grain yield compared to the individual PLSR models and to any previously reported spectral indices developed in wheat (Aparicio et al., 2000; Royo et al., 2003; Gutiérrez et al., 2010). The decrease of $R^2$ after validation was smaller for the combined model compared to any other PLSR models we developed indicating a higher robustness of grain yield prediction with increasing number of trials and genotypes. However, when the combined models were used to predict grain yield in individual trials, the $R^2$ between predicted and actual grain yield was below 40% and thus comparable to the individual models. Therefore, direct selection for grain yield was more efficient ($R^2 < 1$) than indirect selection for predicted grain yield based on individual and combined models. Consequently, selection for predicted grain yield would only be suitable for pre-screening, while final yield evaluation would still be necessary.

The $R^2$ of validation of the combined model was overestimated relative to that between predicted and actual grain yield in individual trials. This was probably caused by the significant genotype-by-treatment interaction of grain yield and its related traits measured with spectral reflectance, implying genotypic rank changes between water regimes.

In particular, the reflectance between 750 and 1300 nm indicated higher plant water content under well-watered than under drought stress conditions (Table A1). Thus, predicting grain yield using the combined model is not recommended, because the genotype-by-treatment interaction across water regimes is not considered and the physiological mechanisms contributing to yield performance are assumed to be the same. Consequently, rather than combining reflectance measurements across different water regimes, a model should be developed for each water regime separately. Prediction of grain yield within water regimes may be limited because of random changes in the relative performance of genotypes from environment to environment and year to year. This form of random genotype-by-environment variation within water regimes can only be reduced by testing the genotypes across several environments. Consequently, for each water regime, reflectance measurements in several trials, locations and years should be combined, if the genotype-by-environment interaction is not significant, to increase the robustness of grain yield prediction. Whether these models are predictive enough for pre-harvest genotype selection has to be investigated in further research because the current study is a pilot study based on a limited amount of data collected in one location during one season. To be applied in a breeding program the models should be able to explain more than 50% of grain yield variation. If the prediction of grain yield proves to be valuable at anthesis and milk-grain stage, this would give further evidences continuing the selection of genotypes under managed drought conditions, wherein drought stress is imposed at this stage.

The number of PLSR factors of the combined model was higher than those of individual PLSR models but similar to that reported for prediction of maize biomass during early plant development (Montes et al., 2011). A high number of PLSR factors may result in over-fitting of models, however, we chose the number according to the first local minimum of RMSE. A reduction of PLSR factors led to increased RMSE and decreased $R^2$ for grain yield (data not shown), suggesting the optimal number for model calibration was 15 for leaf reflectance and 14 for canopy reflectance.

By using a threshold of loading values above 0.08 and below –0.08 the number of wavelengths of the combined PLSR model was reduced to one third. However, correlations between predicted and actual grain yield were only slightly lower than those found using all wavelengths within the combined model (Fig. 2), suggesting selected peaks accounted for a larger proportion of grain yield variability. A number of the selected peaks have previously been reported in the literature related to photosynthetic capacity of leaves or the photosynthetic size of the canopy in the red ranges, red inflection point (REIP), and plant water status (Table A1). Thus, for leaf and canopy reflectance, selected wavelengths were located
in the same spectral regions as the photosynthetic active radiation (PAR, from 495 to 700 nm, Cen and Bornman, 1990) and the REIP from 680 to 780 nm, Filella and Peñuelas, 1994) (Fig. 2). Additional peaks selected for grain yield prediction based on leaf reflectance were related to (i) brown pigment content (from 750 to 800 nm, Peñuelas and Filella, 1998), (ii) water index (at 900 and 970 nm, Peñuelas et al., 1993, 1997), (iii) water content (from 1150 to 1260 nm, and from 1520 to 1540 nm, Sims and Gamon, 2003) and, (iv) water absorbance (at 1450 nm, Sims and Gamon, 2003). Consequently, genotypes with a low reflectance from 400 to 780 nm and from 1150 to 1540 nm have probably a high photosynthetic activity and high leaf/plant water content (Table A.1), respectively, and thus a higher grain yield than those with high reflection in the spectra mentioned (Figs. 1 and 3). To date, these peaks have been measured individually, through the formulation of different spectral indices, to assess drought response in wheat (Royo et al., 2003; Babar et al., 2006a; Gutiérrez et al., 2010). Alternatively, genotypes whose spectral reflectance pattern indicate a low photosynthetic activity and leaf/plant water content relative to a high yielding control could be discarded considering the leaf reflectance spectra at anthesis and the canopy reflectance spectra at anthesis and/or milk-grain stage. Usually, the assessment of yield related traits such as photosynthetic activity and leaf/plant water content is very laborious. Reflectance measurements in the range between 750 and 1300 nm might be of use to measure those traits and incorporate them as novel selection criteria for drought tolerant genotypes Figs. 2 and 3.

In this study, combining wavelengths across a wide spectral rather than the use of individual, pre-selected wavelengths, improved the ability to predict grain yield using spectral reflectance. In addition we used additional peaks from another four spectral regions (leaf reflectance, from 800 to 1000 nm and from 1260 to 1830 nm; canopy reflectance, from 988 to 999 nm and from 1430 to 1640 nm) which have not previously been reported and (to date) are of unknown physiological function.

4.3. Implementation of spectral reflectance in the maize breeding program

According to Falconer and Mackay (1996), the selection gain based on an indirect trait (predicted grain yield) is expected to
be higher than selection based on the direct trait (actual grain yield) if the \( R_g \) between the two traits is very high and the \( h^2 \) is higher for the indirect trait than for the direct trait. In practice, this combination is rarely obtained (Monneveux et al., 2008). In this study the RE based on predicted grain yield was also lower than that based on actual grain yield (Table 3). Nevertheless, if selection intensity of predicted grain yield could be increased relative to actual grain yield, the use of spectral reflectance to predict grain yield prior to harvesting still has potential as a high-throughput phenotyping tool. Higher selection intensity could be obtained by increasing number of genotypes and/or decreasing plot size. For example, by decreasing the plot size from two-row to one-row plots, two times more genotypes (\( n = 300 \) in one trial) could be evaluated while maintaining two replications, or 30% more genotypes (\( n = 200 \) in one trial) could be evaluated using three replications. Moreover, since RE is a function of \( h^2 \) and \( R_g \) between predicted and actual grain yield, it can be maximized by increasing both terms. An increase in \( R_g \) could be obtained by enlarging the number of data points used for calibration, thereby reducing the distance between data points and improving the fit of the regression lines. However, this would increase the work needed and undermine the potential as a high-throughput phenotyping tool. To increase \( h^2 \), improving the precision of measurements is essential. For leaf and canopy reflectance, higher \( h^2 \) could be achieved by increasing the number of replications. Additionally, \( h^2 \) of predicted grain yield based on canopy reflectance could also be increased by correcting measurements for environmental noise (e.g. soil background). Furthermore, the RE for predicted grain yield with spectral reflectance could also be increased when selecting female and male parents for crossing prior to anthesis. Using this method, undesirable male and female genotypes could be discarded thereby increasing RE two-fold compared to selection based on actual grain yield after anthesis (Falconer and Mackay, 1996).

Finally, practical considerations make indirect selection preferable to direct selection of grain yield because using an indirect selection technique prior to anthesis would allow the early identification of superior lines for crossing, thereby speeding up the selection process and allowing reducing costs associated with growing plants till maturity. Considering that measuring reflectance at the leaf level requires approximately three times the length of time relative to measuring reflectance at the canopy level (i.e. 4.5 h for 600 plants compared to 1.5 h for 300 plots), the latter could be used for screening plots in multiple environments in later generations (F\(_5\)–F\(_{10}\)) within a pedigree breeding scheme, while the former could be used for the selection of individual plants with early generation testing (e.g. F\(_2\)–F\(_3\)). However, as our study was based on one location and season only, the results have to be confirmed across several trials, locations and years.

5. Conclusions

By integrating all available spectral information into one PLSR model, more robust assessments of grain yield can be achieved compared to the use of individual, pre-determined spectral indices. The results of this pilot study suggest anthesis and milk-grain stage as the most appropriate stages for the measurement of leaf and canopy reflectance, respectively. PLSR models for grain yield prediction based on leaf and canopy reflectance could explain at maximum 40% of grain yield variation after cross validation. Consequently, selection based on predicted grain yield would only be suitable for pre-screening, while final yield evaluation would still be necessary. In maize breeding, canopy reflectance could be used as a rough screening tool to predict grain yield within early generations, whereas leaf reflectance could be applied during later breeding stages. RE of predicted grain yield could be further increased by increasing \( h^2 \), through increasing the number of data points used for calibration and the selection intensity of spectral measurements. The results obtained in our pilot study are based on one location and season and should therefore be corroborated in future research across several trials, locations and years in order to assess the robustness and power of spectral reflectance measurements in predicting grain yield.

Acknowledgments

This study was supported in part by the BMZ 08.78603-001.00 research project and the Regional Fund for Agriculture Technology (FONTAGRO). We thank J.M. Montes for his advices on the experimental setup, H.F. Utz for his advices for the statistical analysis and G. Romano, W. Spreer and J. Müller for their technical advice on spectroradiometry. We are grateful to the physiology group at Tlatizapán experimental station for their helpful contribution in field experiments. Two anonymous referees are also thanked for very helpful comments and suggestions.

Appendix A.

See Table A.1.
References


