



Insights on multi-spectral vegetation indices derived from UAV-based high-throughput phenotyping for indirect selection in tropical wheat breeding

Caique Machado e Silva · Henrique Caletti Mezzomo · João Paulo Oliveira Ribeiro · Victor Silva Signorini · Gabriel Wolter Lima · Eduardo Filipe Torres Vieira · Marcelo Fagundes Portes · Gota Morota · Lucas de Paula Corredo · Maicon Nardino

Received: 26 June 2023 / Accepted: 21 January 2024
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract High-throughput phenotyping (HTP) approaches are potentially useful for designing indirect selection strategies, which consists in the selection of a primary target trait X based on secondary trait Y. Usually, the secondary trait Y is correlated to the target trait X and is easier, faster and cheaper to measure. In this context, the use of secondary traits (such as vegetation indices derived from HTP platforms) could potentially lead to a fast, non-invasive, accurate and efficient selection of superior genotypes. Considering the lack of information in the literature regarding high-throughput phenotyping approaches

in tropical wheat breeding, this study aimed to (i) determine the best stages to carry out image acquisition for applying multi-spectral vegetation indices for genotype evaluation and selection; (ii) evaluate the heritability and accuracy of multi-spectral vegetation indices; (iii) understand the relationships between vegetation indices and target agronomic traits; and (iv) evaluate the efficiency of indirect selection via UAV-based high-throughput phenotyping. A diversity panel of 49 tropical wheat cultivars was evaluated during the 2022 winter season. Weekly flight campaigns were performed to further build the dataset with multi-spectral vegetation indices, which were then analyzed together with four target agronomic traits. Statistical analysis based on Mixed Effect Model was performed to estimate genetic parameters and predict genetic values, which were subjected to correlation analysis. Additionally, factor analysis was applied, and the factorial scores were used in an indirect selection strategy (indirect via HTP). This strategy was compared to three alternative strategies: direct via grain yield, direct via days to heading, and the multi-trait genotype-ideotype distance index. The results indicate that vegetation indices are suitable for indirect selection strategies and highly efficient for the indirect selection of grain yield and cycle. These findings will help in the decision making regarding the adoption of remote or proximal sensing-based approaches in Brazilian public wheat breeding programs.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10681-024-03299-1>.

C. M. e Silva · J. P. O. Ribeiro · V. S. Signorini · G. W. Lima · E. F. T. Vieira · L. de Paula Corredo · M. Nardino (✉)
Department of Agronomy, Federal University of Viçosa, Viçosa, MG, Brazil
e-mail: nardino@ufv.br

H. C. Mezzomo
GDM Seeds, Lucas do Rio Verde, MT, Brazil

M. F. Portes
Department of Agricultural Engineering, Federal University of Viçosa, Viçosa, MG, Brazil

G. Morota
Department of Animal and Poultry Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Keywords Phenomics · Mixed-model · Multivariate-analyses · *Triticum aestivum* L.

Introduction

Advances in the development of molecular markers and genotype sequencing technologies in the last decades have made high-throughput genotyping possible at relatively affordable costs (You et al. 2018). Nevertheless, plant breeding programs still depend on the acquisition of high-quality phenotypic information to better understand the contribution of genetic variation to phenotypic variation. In this context, efforts in the phenomics field have been focused on developing phenotyping platforms to allow accurate evaluation of genotypes for further integration of phenotypic information with available genomic data (Mir et al. 2019).

Traditional phenotyping methods are commonly invasive, laborious, time-consuming, and inefficient, limiting genetic gains. Alternatively, phenotyping techniques based on remote or proximal sensing allow fast and easy acquisition of high-resolution data, as well as efficient and accurate evaluation and selection of genotypes. Additionally, advances in the development of unmanned aerial vehicle (UAV) platforms, combined with improvements in computational capabilities for processing UAV-derived image data, have enabled the evaluation of several important traits in plant breeding programs, including plant height, biomass, leaf area index, tolerance to biotic and abiotic stress, nutritional and vegetation indices, for example (Volpato et al. 2021; Feng et al. 2021). Vegetation indices are by themselves of particular interest because they allow for indirect selection.

In the indirect selection method, a target trait *X* is indirectly selected based on a secondary trait *Y*, which has genetic correlation with the target trait *X* and can be measured via HTP (e.g., vegetation indices) (Morota et al. 2022). Consequently, selection on trait *Y* can cause corresponding changes on the target trait *X*. In some cases, the indirect selection is a suitable strategy to rapidly achieve desired genetic gains in the breeding program (Moreira et al. 2019). This is true, especially when measuring the target primary trait is costly, laborious and time-consuming.

Several vegetation indices are suitable for being used as secondary traits in indirect selection strategies. The normalized difference vegetation index (NDVI) (Rouse

et al. 1974), one of the most commonly used vegetation indices in agriculture, is based on the differences in reflectance between the near-infrared (NIR) and red wavelengths, and can be used to identify vegetation and make inferences about vegetative status (Khan et al. 2018). NDVI is often associated with higher biomass accumulation and growth rate when measured in the vegetative stage, and with longer duration of the grain filling phase and lower leaf senescence during grain filling (Babar et al. 2006). Despite its widespread use, NDVI is generally sensitive to the irradiation reflected by the soil in areas with low canopy cover (Prudnikova et al. 2019). Therefore, the soil-adjusted vegetation index (SAVI) (Huete et al. 1988) has been proposed to overcome this problem.

The green normalized difference vegetation index (GNDVI), which is based on the differences between the NIR and green spectral regions (Gitelson et al. 1997), is also widely used. GNDVI, the normalized difference red edge index (NDRE) and the simplified chlorophyll canopy content index (SCCCI) are sensitive to chlorophyll concentration and highly correlated with leaf nitrogen content. For this reason, they have been used to monitor nitrogen leaf status (Li et al. 2020; Barzin et al. 2022).

In Brazil, some reports highlight the suitability of UAV-based image analysis for the evaluation and selection of genotypes in soybean (Santana et al. 2021; Casagrande et al. 2022; Santana et al. 2022a), corn (Santana et al. 2022b), elephant grass (Ferreira et al. 2022) and eucalyptus (Borges et al. 2022). However, studies on high-throughput phenotyping approaches are needed for the evaluation of tropical wheat trials and the design of selection strategies. Thus, the present study aimed to (i) determine the best stages to carry out image acquisition for applying multi-spectral vegetation indices; (ii) evaluate the heritability and accuracy of multi-spectral vegetation indices; (iii) understand the relationships between vegetation indices and target agronomic traits; and (iv) evaluate the efficiency of indirect selection using UAV-based high-throughput phenotyping.

Material and methods

Diversity panel and field trial

Figure 1 summarizes the UAV-based high-throughput phenotyping process, from the field trial

implementation to the analysis of agronomic data and vegetation indices.

A diversity panel of 49 tropical wheat genotypes released between 1800 and 2021 (Supplementary Table S1) was evaluated in the 2022 winter crop season at the Universidade Federal de Viçosa (20° 45' 14" S; 42° 52' 55" W; 648 m altitude), Viçosa, Minas Gerais State, Brazil. The climate of the region is monsoon-influenced humid subtropical climate with wet winters and hot summers, average annual precipitation between 1300 and 1600 mm, and average annual temperature of 21 °C (Alvares et al. 2013). The genotypes were evaluated in a complete 7×7 lattice design with two replications. The plots were two lines with two meters, spaced at 0.20 m, with a population density of 350 seeds m⁻². The agronomic practices were performed according to the technical recommendations for wheat cultivation in the Brazilian Central-South region.

Flight campaign, image acquisition and processing

Eight ground point controls were distributed in the experimental area for further geometric correction of the orthomosaic map. Their coordinates were registered using a global navigation satellite system receiver with real-time kinematic correction (Trimble T10 GNSS system, Trimble Inc. Sunnyvale, CA, EUA).

The flights were performed using a DJI Matrice 100 platform (DJI Innovations, Shenzhen, China) equipped with a MicaSense RedEdge MX multi-spectral camera (MicaSense, Seattle, WA, EUA). The camera has five individual sensors that simultaneously capture five spectral bands (one band per sensor): 475 nm ± 20 nm (Blue), 560 nm ± 20 nm (Green), 668 nm ± 10 nm (Red), 840 nm ± 40 nm (NIR) and 717 nm ± 10 nm (RedEdge).

Precision Flight for DJI Drones (PrecisionHawk, Raleigh, NC, USA) was used for flight planning. During the planning, an 80% overlap between frontal and side photos was defined, as well as a flight height of 30 m starting from the take-off point. The flights were performed once a week, starting from the tillering stage until harvesting. All the flights were performed under stable light conditions between 11:00 and 13:00. The radiometric calibration of all the bands was performed using images of a reference panel provided by the manufacturer, which were taken before

each flight, and the appropriate correction factor was extracted from these images with the reference calibration panel. The orthomosaic and georeferencing of the images were then performed using Agisoft Metashape (Agisoft, 2019) software system.

Agronomic traits and vegetation indices

Days to heading (DH) was evaluated when at least 50% of the plants in the plots presented exposed ears. Five representative plants from each plot were collected to evaluate plant height (PH, cm) using a graduate ruler. One hundred grains from each plot were weighed to assess their mass (HGM, g). After physiological maturity, the plots were manually harvested to determine grain yield (GY, kg ha⁻¹). Humidity was adjusted to 13% in all plots.

After obtaining and processing the multi-spectral images, the wavelengths available were used to calculate six vegetation indices (Table 1) in QGIS software, version 3.2. These indices were calculated in seven different phenological stages, during the development of the crop (Table 2).

Genetic and statistical analyses

Mixed model

The phenotypic values were subjected to mixed effect model analysis (Patterson and Thompson 1971; Henderson 1975) to estimate the variance components using restricted maximum likelihood (REML) and to predict genotypic values using best linear unbiased prediction (BLUP). Mixed effect model analysis was performed in R software, version 4.2.2 (R Core Team 2022), using the R package metan (Olivoto and Lúcio 2020). The following model was considered:

$$y = Xr + Zg + Wb + e$$

where **y** is the vector of phenotypic data; **r** is the vector of fixed effects of replications added to the overall mean; **g** is the vector of random effects of genotypes, $\mathbf{g} \sim N(0, \sigma_g^2)$; **b** is the vector of random effects of blocks nested within replication, $\mathbf{b} \sim N(0, \sigma_b^2)$; **e** is the vector of random effects of errors, $\mathbf{e} \sim N(0, \sigma_e^2)$; **X**, **Z** and **W** are the incidence matrices of these effects.

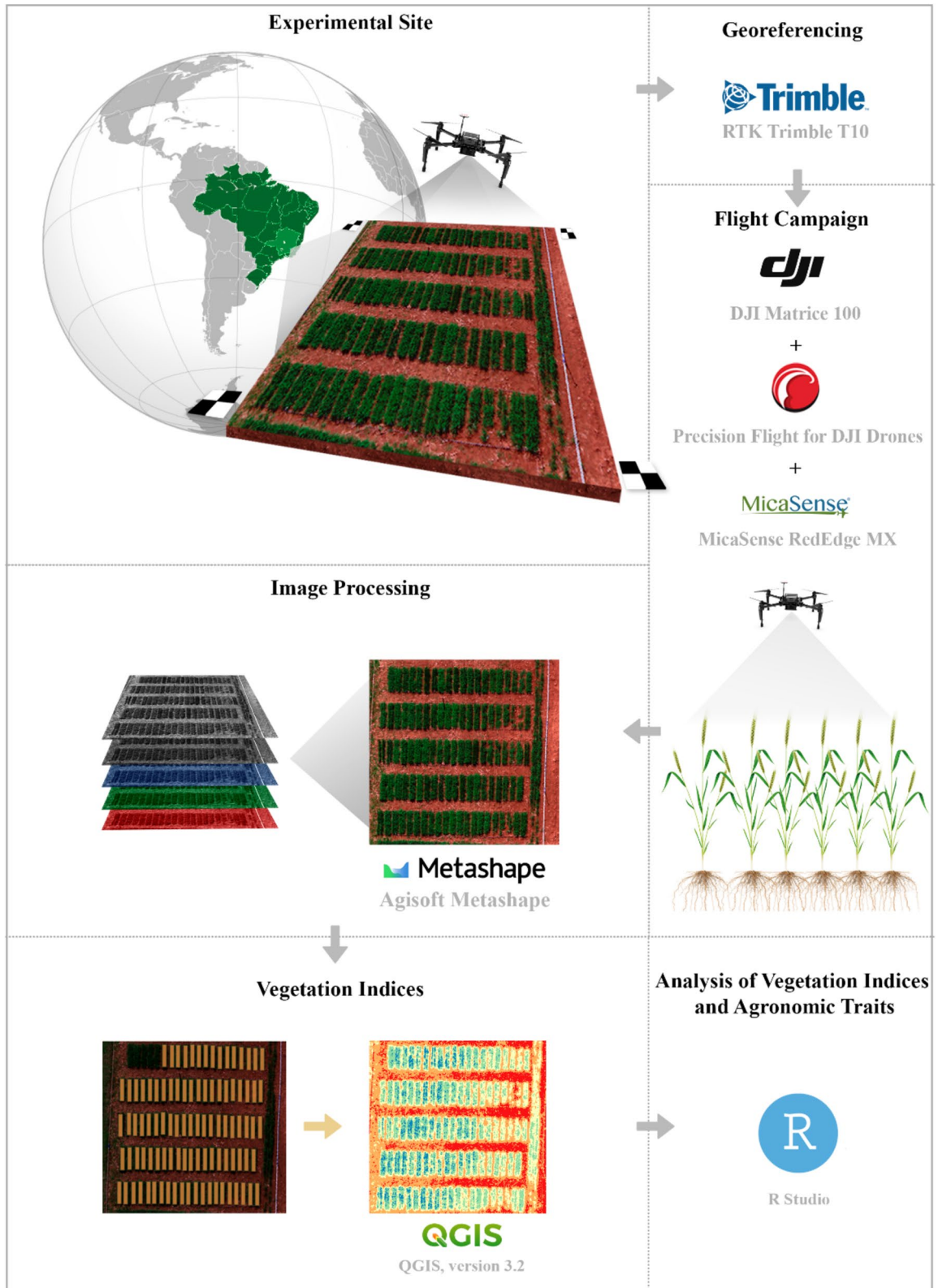


Fig. 1 UAV-based high-throughput phenotyping workflow of a tropical wheat diversity panel

The significance of the genotype effect was tested using the likelihood ratio test (LRT) (Wilks 1938; Resende and Alves 2020) as follows:

$$LRT = -2(\text{Log}L_F - \text{Log}L_R)$$

where $\text{Log}L_F$ is the logarithm of the restricted likelihood function of the full model, and $\text{Log}L_R$ is the logarithm of the restricted likelihood function of the reduced model. The significance of the genotype effect was tested by the chi-square test at the threshold of 5% and 1%.

The variance components were used to estimate the broad-sense heritability as follows:

$$h^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_p^2}$$

where h^2 is the broad-sense heritability; $\hat{\sigma}_g^2$ is the genotypic variance and $\hat{\sigma}_p^2$ is the phenotypic variance.

The selection accuracy was obtained as follows:

$$\hat{r}_{gg} = \sqrt{1 - \frac{PEV}{\hat{\sigma}_g^2}}$$

where \hat{r}_{gg} is the selection accuracy; PEV is the prediction error variance; and $\hat{\sigma}_g^2$ is the genotypic variance.

Genetic correlation analysis

The genotypic values were used to estimate the genetic correlation coefficients among the traits. Genetic correlation analysis was performed using basic functions of R (R Core Team 2022), and plots were generated using the ggcorrplot package. Correlation coefficients were estimated as follows:

$$\hat{r}_{g_{xy}} = \frac{Cov_g(x, y)}{\sqrt{\hat{\sigma}_g^2(x) \times \hat{\sigma}_g^2(y)}}$$

where $\hat{r}_{g_{xy}}$ is the genetic correlation coefficient between two phenotypes; Cov_g is the genotypic covariance; $\hat{\sigma}_g^2$ is the genotypic variance; and x and y are two phenotypes.

Factor analysis

Factor analysis was performed to reduce the dimensionality of the data and better understand the relationship among the traits from a multivariate perspective. Factor analysis was performed in R software, version 4.2.2 (R Core Team 2022), using the R psych package (Revelle 2020). The following model was considered (Momen et al. 2021; Yu et al. 2019):

$$Y = \Lambda F + U$$

where Y is the $t \times n$ matrix of phenotypic observations; Λ is the $t \times q$ matrix of factor loadings; F is the $q \times n$ matrix of factor scores; and U is the $t \times n$ vector of unique effects that is not explained by underlying q common factors. The variance–covariance matrix of Y is:

$$\Sigma = \Lambda \Phi \Lambda' + \Psi$$

where Σ is the $t \times t$ variance–covariance matrix of phenotypes; Φ , is the variance of factor scores; and Ψ is a $t \times t$ diagonal matrix of unique variance. The elements of Λ , Φ , and Ψ are parameters of the model to be estimated from the data. It was assumed that $\Phi = I$ yielding factors each with unit variance (Anderson 2003; Jöreskog 1967). Parameters Λ and Ψ were estimated by maximizing the log-likelihood of $\mathcal{L}(\Lambda, \Psi|Y)$ along with a varimax rotation (Kaiser 1958).

Selection strategies

Four different selection strategies were designed: direct via grain yield, direct via days to heading, indirect via HTP (using the factor scores obtained in the factor analysis), and the multi-trait genotype-ideotype distance index (MGIDI) (Olivoto and Nardino 2021). For selecting the most suitable genotypes, the 20% with best performance were retained in all strategies.

The response to direct selection was obtained as follows:

$$SG_j = DS_j \times h_j^2$$

where SG_j is the selection gain in trait j ; DS_j is the selection differential in trait j , calculated as the difference between the mean of the selected genotypes (20% best performing genotypes) and the original

Table 1 Vegetation indices calculated using the multi-spectral images

Vegetation Indices	Equation	References
NDVI	$\frac{Nir-Red}{Nir+Red}$	Rouse et al. (1974)
SAVI	$(1 + 0.5) \frac{Nir-Red}{Nir+Red+0.5}$	Huete et al. (1988)
GNDVI	$\frac{Nir-Green}{Nir+Green}$	Gitelson et al. (1997)
NDRE	$\frac{Nir-RedEdge}{Nir+RedEdge}$	Gitelson et al. (1996)
SCCCI	$\frac{NDRE}{NDVI}$	Raper and Varco (2014)
EVI	$\frac{Nir-Red}{(Nir+6\times Red-7.5\times Green)+1}$	Justice et al. (1998)

NDVI, normalized difference vegetation index; SAVI, soil adjusted vegetation index; GNDVI, green normalized difference vegetation index; NDRE, normalized difference red edge index; SCCCI, simplified chlorophyll canopy content index; MSAVI, modified soil adjusted vegetation index; EVI, enhanced vegetation index

mean of the diversity panel; and h_j^2 is the broad-sense heritability of trait j.

The correlated response to indirect selection was obtained as follows:

$$SG_{j(i)} = DS_{j(i)} \times h_j^2$$

where $SG_{j(i)}$ is the selection gain in trait j when selection was carried out based on trait i; $DS_{j(i)}$ is the selection differential in trait j when selection was carried out based on trait i; and h_j^2 is the broad-sense heritability of trait j.

The efficiency of the indirect selection was assessed as follows:

$$ESI = \frac{SG_{j(i)}}{SG_j}$$

MGIDI requires the planning of an ideotype for the analyzed traits. In this work, positive desired gains were assumed for GY and HGM, and negative desired gains were assumed for PH and DH. MGIDI was performed as follows (Olivoto and Nardino 2021):

$$MGIDI_i = [\sum_{j=1}^f (\gamma_{ij} - \gamma_j)^2]^{0.5}$$

where $MGIDI_i$ is the multi-trait genotype-ideotype distance index for the i -th genotype; γ_{ij} is the score of the i -th genotype in the j -th factor ($i = 1, 2, \dots, g$; $j = 1, 2, \dots, f$); g and f are the number of genotypes and factors, respectively; and γ_j is the j th ideotype score.

Results

Descriptive analysis of vegetation indices

A descriptive analysis of tendency, dispersion, symmetry and outliers for the phenotypic values of the vegetation indices is presented in Fig. 2. The median phenotypic values widely varied across the different phenological stages for most vegetation indices throughout the crop development, except for NDVI and SAVI. Overall, the greatest phenotypic variability was observed in the maturation stage, suggesting the occurrence of significant differences in vegetative status among the genotypes at this stage. The boxplots also indicate a symmetric distribution of vegetation indices values in some phenological stages. This trend suggests the existence of suitable genetic variability between the genotypes in particular stages that could be explored in further selection strategies.

Mixed model

The results of the LRT for the random effect of genotype, considering all the analyzed traits, are given in Table 3. The random effect of genotype was significant at the threshold of 1 or 5% for all the agronomic traits and vegetation indices, considering all the phenological stages, except for GNDVI, NDRE, NDVI and SAVI at tillering; and NDVI and SAVI at stem elongation. The highest LRT values were observed for the vegetation indices on maturation.

In Fig. 3 it is presented the relative contribution of the genotypic variance to the total phenotypic variance observed for the random effect genotype, considering the agronomic traits and the vegetation indices in the different phenological stages. As expected, considering the LRT test, the genotypic variance was proportionally greater than the residual variance for the agronomic traits. The genotypic variance was relatively small for most vegetation indices at tillering and stem elongation stages. The highest contribution of genotypic variance to the total phenotypic variance was observed at the early flowering stage.

Heritability and accuracy

The estimates of heritability and accuracy values for the vegetation indices from tillering to maturation are shown in Fig. 4. The heritability estimates ranged

Table 2 Codes used for the vegetation indices corresponding to different phenological stages in a diversity panel of tropical wheat genotypes

Vegetation indices	Phenological stage	Code	Phase according to Zadoks et al. (1974)
EVI	Tillering	EVI.TIL	20–29
GNDVI		GNDVI.TIL	
NDRE		NDRE.TIL	
NDVI		NDVI.TIL	
SAVI		SAVI.TIL	
SCCCI	Stem Elongation	SCCCI.TIL	30–39
EVI		EVI.SE	
GNDVI		GNDVI.SE	
NDRE		NDRE.SE	
NDVI		NDVI.SE	
SAVI	SAVI.SE	40–49	
SCCCI	SCCCI.SE		
EVI	EVI.BO		
GNDVI	GNDVI.BO		
NDRE	NDRE.BO		
NDVI	NDVI.BO	50–59	
SAVI	SAVI.BO		
SCCCI	SCCCI.BO		
EVI	EVI.HE		
GNDVI	GNDVI.HE		
NDRE	NDRE.HE	60–69	
NDVI	NDVI.HE		
SAVI	SAVI.HE		
SCCCI	SCCCI.HE		
EVI	EVI.FLO		
GNDVI	GNDVI.FLO	70–89	
NDRE	NDRE.FLO		
NDVI	NDVI.FLO		
SAVI	SAVI.FLO		
SCCCI	SCCCI.FLO		
EVI	Grain Filling	EVI.GF	70–89
GNDVI		GNDVI.GF	
NDRE		NDRE.GF	
NDVI		NDVI.GF	
SAVI		SAVI.GF	
SCCCI	SCCCI.GF	90–99	
EVI	EVI.MA		
GNDVI	GNDVI.MA		
NDRE	NDRE.MA		
NDVI	NDVI.MA		
SAVI	SAVI.MA	SCCCI.MA	
SCCCI	SCCCI.MA		

EVI, enhanced vegetation index; GNDVI, green normalized difference vegetation index; MSAVI, modified soil adjusted vegetation index; NDRE, normalized difference red edge index; NDVI, normalized difference vegetation index; SAVI, soil adjusted vegetation index; SCCCI, simplified chlorophyll canopy content index; TIL, tillering; SE, stem elongation; BO, booting; HE, heading; FLO, flowering; GF, grain filling; MA, maturation

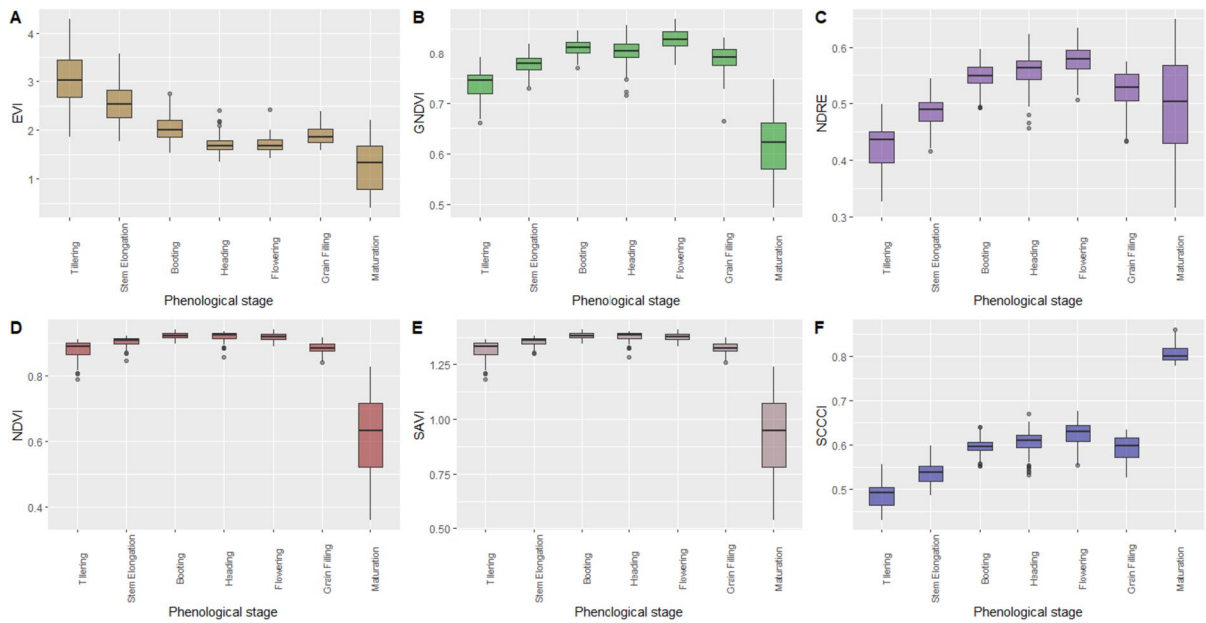


Fig. 2 Boxplot of the phenotypic values of six vegetation indices on different phenological stages in tropical wheat diversity panel. EVI, enhanced vegetation index; GNDVI, green normalized difference vegetation index; MSAVI, modified soil

adjusted vegetation index; NDRE, normalized difference red edge index; NDVI, normalized difference vegetation index; SAVI, soil adjusted vegetation index; SCCC, simplified chlorophyll canopy content index

from 0.01 for NDVI and SAVI at tillering to 0.92 for EVI at booting. The highest mean for heritability estimate was observed for the indices in the maturation stage (0.87). The accuracy values ranged from 0.10 for NDVI and SAVI at tillering to 0.96 for EVI at booting. The highest mean values for accuracy were also observed for the vegetation indices at maturation stage (0.93). EVI and SCCC presented the highest and most consistent heritability and accuracy values across the different phenological stages.

Genetic correlation

In Fig. 5, it is presented the genetic correlations among the agronomic traits and the vegetation indices on different phenological stages. DH showed a high negative correlation with GY (-0.76) and a low negative correlation with HGM (-0.29). PH showed no significant correlation with any of the agronomic traits. Most of the vegetation indices showed a high positive correlation among themselves from tillering to maturation, except for the correlations with SCCC in maturation, where the association trends changed. The associations among

the agronomic traits and most of the vegetation indices were not significant from tillering to heading, however, they started to be significant from flowering and reached the highest correlation values in maturation. From flowering to maturation, GY and HGM showed a moderate to high negative correlation with the vegetation indices, while DH showed a moderate to high positive correlation with the vegetation indices.

Factor analysis

Four factors were obtained in the factor analysis (Table 4), which explained 80% of the total variance observed. In the first factor, GY and DH were grouped with NDVI and SAVI at heading; GNDVI, NDVI, and SAVI at flowering; and most vegetation indices at grain filling and maturation. In the second factor, PH was grouped with most vegetation indices at booting and heading; and with some vegetation indices at flowering. In the third and fourth factors, the vegetation indices at tillering and stem elongation were grouped together.

Table 3 Likelihood ratio test for four agronomic traits and six vegetation indices on different phenological stages in a tropical wheat diversity panel

Trait	χ^2	<i>p</i> value
DH	31.45	<i>p</i> < 0.01
GY	25.79	<i>p</i> < 0.01
HGM	37.63	<i>p</i> < 0.01
PH	14.95	<i>p</i> < 0.01
EVI.TIL	21.39	<i>p</i> < 0.01
GNDVI.TIL	3.36	<i>p</i> < 0.05
NDRE.TIL	2.38	<i>p</i> < 0.05
NDVI.TIL	0.00	<i>p</i> < 0.05
SAVI.TIL	0.00	<i>p</i> < 0.05
SCCCI.TIL	15.65	<i>p</i> < 0.01
EVI.SE	43.86	<i>p</i> < 0.01
GNDVI.SE	11.13	<i>p</i> < 0.01
NDRE.SE	14.84	<i>p</i> < 0.01
NDVI.SE	2.02	<i>p</i> < 0.05
SAVI.SE	1.94	<i>p</i> < 0.05
SCCCI.SE	29.99	<i>p</i> < 0.01
EVI.BO	60.76	<i>p</i> < 0.01
GNDVI.BO	24.27	<i>p</i> < 0.01
NDRE.BO	38.41	<i>p</i> < 0.01
NDVI.BO	7.39	<i>p</i> < 0.01
SAVI.BO	7.39	<i>p</i> < 0.01
SCCCI.BO	12.07	<i>p</i> < 0.01
EVI.HE	39.94	<i>p</i> < 0.01
GNDVI.HE	10.92	<i>p</i> < 0.01
NDRE.HE	16.17	<i>p</i> < 0.01
NDVI.HE	22.39	<i>p</i> < 0.01
SAVI.HE	26.80	<i>p</i> < 0.01
SCCCI.HE	10.44	<i>p</i> < 0.01
EVI.FLO	13.87	<i>p</i> < 0.01
GNDVI.FLO	15.68	<i>p</i> < 0.01
NDRE.FLO	12.05	<i>p</i> < 0.01
NDVI.FLO	10.78	<i>p</i> < 0.01
SAVI.FLO	10.77	<i>p</i> < 0.01
SCCCI.FLO	16.34	<i>p</i> < 0.01
EVI.GF	26.73	<i>p</i> < 0.01
GNDVI.GF	21.76	<i>p</i> < 0.01
NDRE.GF	14.12	<i>p</i> < 0.01
NDVI.GF	10.23	<i>p</i> < 0.01
SAVI.GF	10.23	<i>p</i> < 0.01
SCCCI.GF	17.18	<i>p</i> < 0.01
EVI.MA	46.40	<i>p</i> < 0.01
GNDVI.MA	46.84	<i>p</i> < 0.01
NDRE.MA	48.08	<i>p</i> < 0.01
NDVI.MA	48.37	<i>p</i> < 0.01

Table 3 (continued)

Trait	χ^2	<i>p</i> value
SAVI.MA	48.37	<i>p</i> < 0.01
SCCCI.MA	25.87	<i>p</i> < 0.01

Bold values indicate significance for the random effect of genotype at 5 or 1% error probability by the chi-square test: $\chi^2_{5\%} = 3.84$; $\chi^2_{1\%} = 6.63$ GY, grain yield; HGM, hundred grain mass; PH, plant height; EVI, enhanced vegetation index; GNDVI, green normalized difference vegetation index; MSAVI, modified soil adjusted vegetation index; NDRE, normalized difference red edge index; NDVI, normalized difference vegetation index; SAVI, soil adjusted vegetation index; SCCCI, simplified chlorophyll canopy content index; TIL, tillering; SE, stem elongation; BO, booting; HE, heading; FLO, flowering; GF, grain filling; MA, maturation

Selection gains

In Table 5, it is presented the predicted genetic gains for the four agronomic traits using the different selection strategies: direct via GY, direct via DH, indirect via HTP, and MGIDI. The direct selection via GY provided the highest percentage genetic gain for GY (32.68%) and a desired gain for DH (6.69%). The selection gain for PH and HGM by the indirect selection via GY was a low undesired gain in the case of PH, and a low gain in the case of HGM. The selection gain achieved for the direct selection via DH, considering GY and DH, was 25.33% and 10.27%, respectively. The selection gain achieved for this strategy was 1.14% for PH and 2.49% for HGM.

The use of HTP resulted in balanced genetic gains for all the agronomic traits. For GY, the selection gain observed was 27.63%, while for DH, the selection gain was 8.78%. The selection gain obtained with this strategy was a low desired gain for PH (1.77%) and a low gain for HGM (3.52%). The efficiency of indirect selection via HTP for GY and DH was 0.85. Similar to HTP, MGIDI also resulted in balanced desired gains for the agronomic traits: the genetic gains for GY, DH, PH and HGM were 26.67, 8.45, 3.58 and 1.04%, respectively.

Discussion

The significance of genotype as random effect in the LRT test for the agronomic traits and most of the vegetation indices during wheat development

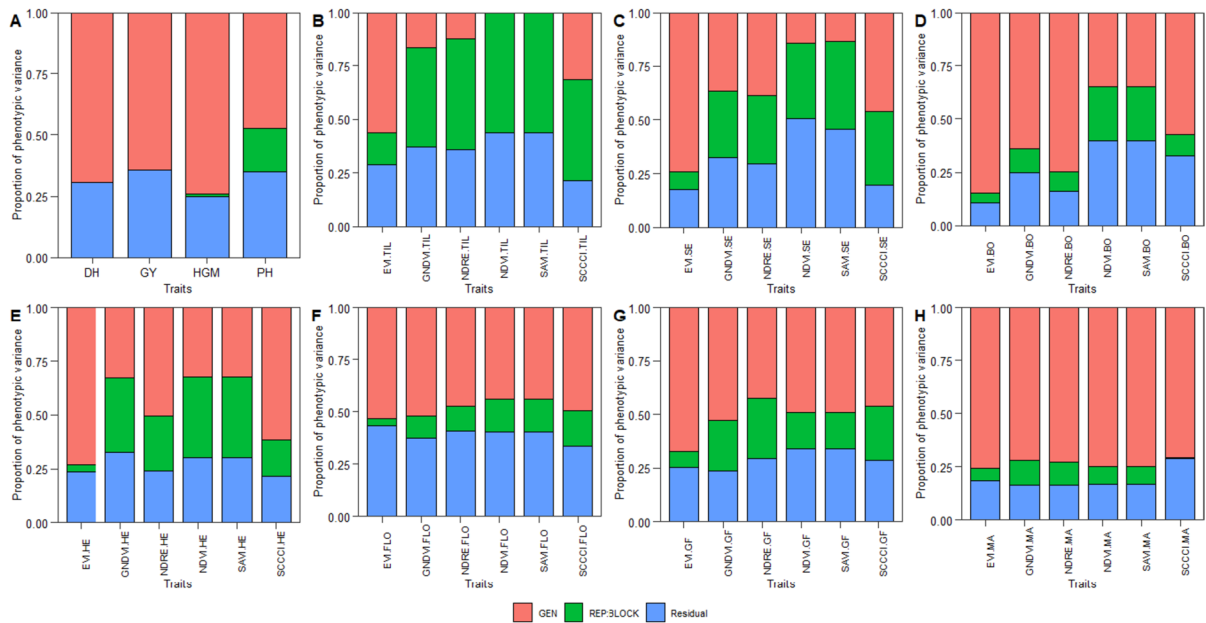


Fig. 3 Contribution of the genotypic, replication/block and residual variances to the total phenotypic variance for the random effect of genotype for four agronomic traits (**A**) and six vegetation indices applied in different phenological stages (from **B** to **H**) in a tropical wheat diversity panel. DH, days to heading; GY, grain yield; HGM, hundred grain mass; PH, plant height; EVI, enhanced vegetation index; GNDVI, green

normalized difference vegetation index; MSAVI, modified soil adjusted vegetation index; NDRE, normalized difference red edge index; NDVI, normalized difference vegetation index; SAVI, soil adjusted vegetation index; SCCCI, simplified chlorophyll canopy content index; TIL, tillering; SE, stem elongation; BO, booting; HE, heading; FLO, flowering; GF, grain filling; MA, maturation

indicates the presence of a wide genetic variability, which can be described not only directly measuring crop traits but also indirectly via vegetation indices, allowing the design of appropriate high-throughput selection strategies. The elevated chi-square test statistics observed in maturation and the large contribution of the genotypic variance to the phenotypic variance regarding the genotype effect for vegetation indices from flowering to maturation suggest that these stages may be appropriate to capture the genetic variability among genotypes using the canopy spectral response.

The significance of genotype as random effect for most of the vegetation indices (especially in the maturation stage) is related to the wheat varieties included in the diversity panel used in this study. Since the diversity panel consisted of a set of genotypes released in different periods in the Brazilian wheat breeding programs (Lima et al. 2022), it has a huge genetic variability for several traits, especially for cycle and grain yield. This variability resulted also in large variation of different phenotypic traits that

could be captured by the vegetation indices, resulting in different spectral signatures according to genotype.

The heritability and accuracy estimates obtained for the vegetation indices can be classified as moderate or high (Resende and Alves 2020) for most of the phenological stages evaluated, which can be considered satisfactory. In practice, these estimates can be used to compare alternative selection methods, calculate genetic gains with selection and design experiments (Resende and Alves 2022). Some studies have been conducted to evaluate the heritability and accuracy of vegetation indices in different phenological stages to make genetic predictions. Hassan et al. (2019) reported moderate or high heritability estimates for NDVI in reproductive stages in a set of Chinese wheat germplasm evaluated under full and limited irrigation systems. Using a panel of elite hard winter wheat genotypes, Frels et al. (2018) reported low heritability estimates for most vegetation indices evaluated at the heading stage, with increasing heritability estimates as the season progressed. As a result, measurements of vegetation indices collected during

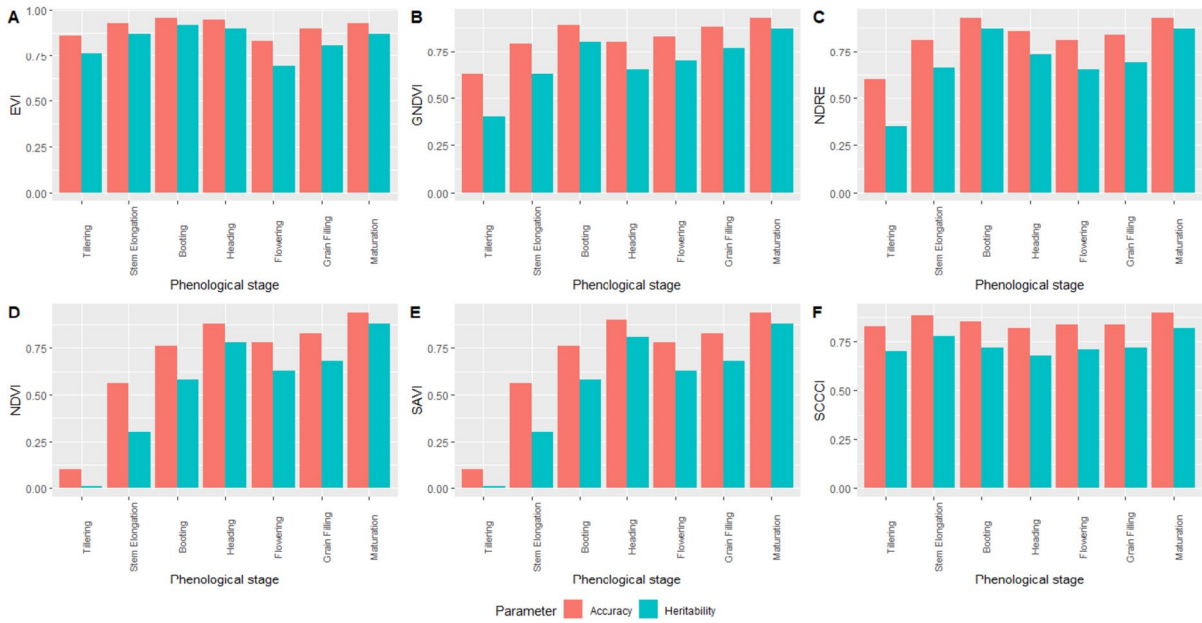


Fig. 4 Heritability and accuracy values of six vegetation indices applied in different phenological stages in a tropical wheat diversity panel. EVI, enhanced vegetation index; GNDVI, green normalized difference vegetation index; MSAVI, modi-

fied soil adjusted vegetation index; NDRE, normalized difference red edge index; NDVI, normalized difference vegetation index; SAVI, soil adjusted vegetation index; SCCCI, simplified chlorophyll canopy content index

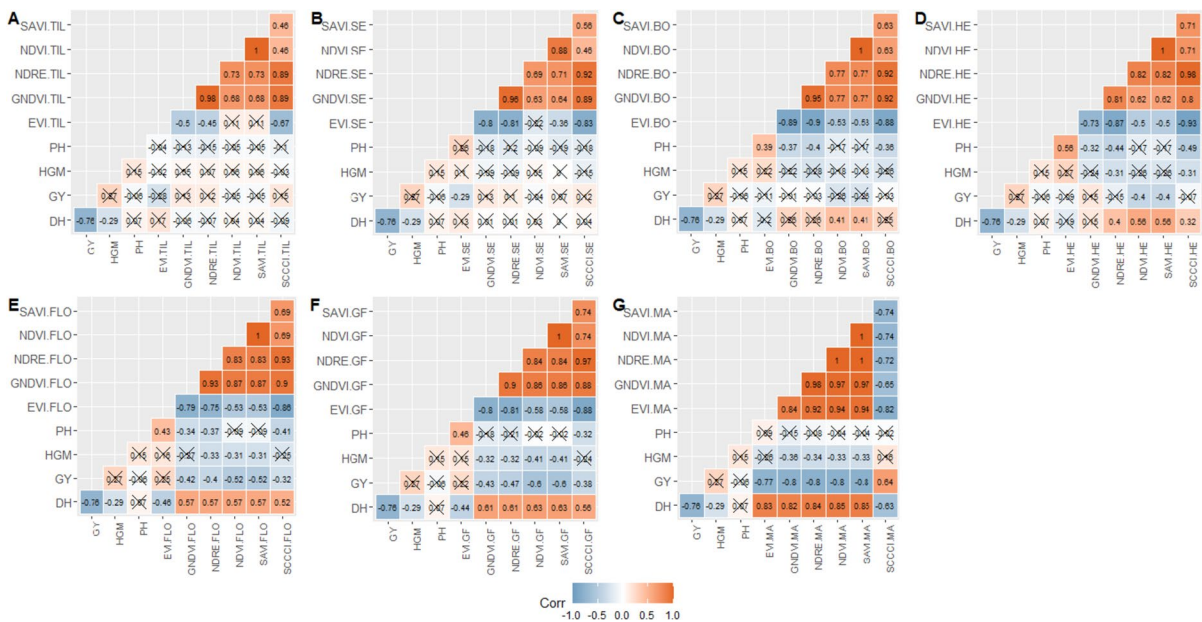


Fig. 5 Genetic correlation among four agronomic traits and six vegetation indices applied on different phenological stages of a tropical wheat diversity panel. DH, days to heading; GY, grain yield; HGM, hundred grain mass; PH, plant height; EVI, enhanced vegetation index; GNDVI, green normalized difference vegetation index; MSAVI, modified soil adjusted veg-

etation index; NDRE, normalized difference red edge index; NDVI, normalized difference vegetation index; SAVI, soil adjusted vegetation index; SCCCI, simplified chlorophyll canopy content index; (A) TIL, tillering; (B) SE, stem elongation, (C) BO, booting; (D) HE, heading; (E) FLO, flowering; (F) GF, grain filling; (G) MA, maturation

Table 4 Factor loadings of four agronomic traits and six vegetation indices applied in different phenological stages and cumulative variance in the first four factors using the factor analysis

Trait	FA1	FA2	FA3	FA4
DH	0.83	0.14	-0.04	-0.08
GY	-0.83	0.13	0.10	0.05
HGM	-0.30	-0.23	-0.06	0.12
PH	0.04	-0.53	0.00	-0.03
EVI.TIL	0.22	-0.06	-0.78	0.12
GNDVI.TIL	-0.03	0.25	0.64	0.66
NDRE.TIL	-0.03	0.24	0.60	0.70
NDVI.TIL	0.13	0.05	0.04	0.99
SAVI.TIL	0.13	0.05	0.04	0.99
SCCCI.TIL	-0.03	0.31	0.71	0.42
EVI.SE	0.19	-0.34	-0.86	0.00
GNDVI.SE	0.03	0.27	0.81	0.44
NDRE.SE	0.04	0.32	0.82	0.44
NDVI.SE	0.12	0.17	0.29	0.70
SAVI.SE	0.06	0.31	0.33	0.66
SCCCI.SE	0.06	0.43	0.83	0.24
EVI.BO	-0.07	-0.77	-0.47	-0.05
GNDVI.BO	0.23	0.77	0.50	0.13
NDRE.BO	0.21	0.76	0.48	0.21
NDVI.BO	0.43	0.55	0.26	0.21
SAVI.BO	0.43	0.55	0.26	0.21
SCCCI.BO	0.21	0.65	0.63	0.18
EVI.HE	-0.06	-0.87	-0.36	-0.06
GNDVI.HE	0.10	0.70	0.48	0.12
NDRE.HE	0.37	0.86	0.27	0.18
NDVI.HE	0.61	0.59	0.08	0.21
SAVI.HE	0.61	0.59	0.08	0.21
SCCCI.HE	0.27	0.88	0.32	0.16
EVI.FLO	-0.37	-0.72	-0.23	-0.02
GNDVI.FLO	0.60	0.60	0.13	0.17
NDRE.FLO	0.57	0.64	0.09	0.29
NDVI.FLO	0.68	0.33	0.02	0.29
SAVI.FLO	0.68	0.33	0.02	0.29
SCCCI.FLO	0.51	0.74	0.10	0.20
EVI.GF	-0.44	-0.66	-0.27	-0.18
GNDVI.GF	0.65	0.58	0.09	0.18
NDRE.GF	0.74	0.54	0.12	0.20
NDVI.GF	0.74	0.32	0.03	0.22
SAVI.GF	0.74	0.32	0.03	0.22
SCCCI.GF	0.66	0.62	0.13	0.21
EVI.MA	0.93	0.06	-0.13	-0.02
GNDVI.MA	0.95	0.19	-0.02	0.03
NDRE.MA	0.98	0.18	-0.05	-0.01

Table 4 (continued)

Trait	FA1	FA2	FA3	FA4
NDVI.MA	0.98	0.16	-0.06	-0.02
SAVI.MA	0.98	0.16	-0.06	-0.02
SCCCI.MA	-0.75	-0.01	-0.03	-0.04
Accumulated variance	0.28	0.53	0.68	0.80

Factor loadings in bold indicate traits grouped within the same factor. Acronyms: FA1, first factor; FA2, second factor; FA3, third factor; FA4, fourth factor; GY, grain yield; HGM, hundred grain mass; PH, plant height; EVI, Enhanced vegetation index; GNDVI, Green normalized difference vegetation index; MSAVI, modified soil adjusted vegetation index; NDRE, Normalized difference red edge index; NDVI, Normalized difference vegetation index; SAVI, Soil adjusted vegetation index; SCCCI, simplified chlorophyll canopy content index; TIL, tillering; SE, stem elongation; BO, booting; HE, heading; FLO, flowering; GF, grain filling; MA, maturation

the late grain filling stage were considered to be the most accurate and least affected by the environment. These reports are consistent with the results of this study.

Heritability estimates are the proportion of phenotypic variance explained by genotypic variance, while accuracy is the correlation among predicted and observed genotypic values. Although the vegetation indices are mathematical functions among wavelengths, the estimation of heritability and accuracy for these traits is justified because vegetation indices describe the spectral response of plants to incident light, which depends on the vegetative material biochemical and biophysical properties. In fact, vegetation indices vary according to multiple factors, for example: biophysical and biochemical aspects, disease incidence/resistance, chlorophyll content, hydric stress and canopy cover (Candiago et al. 2015). These traits are largely under genetic control and can potentially be improved. Since the vegetation indices showed moderate to high heritability and accurate estimates, they can be used as secondary traits in indirect selection strategies, as long as they have a high correlation with the target primary trait (Rutkowski et al. 2016).

The correlation analysis showed that the magnitude of the relationship among the agronomic traits and the vegetation indices may not be consistent across the different phenological stages. In earlier stages (from tillering to heading), the correlations were not significant. However, as the season

Table 5 Predicted genetic gains for four agronomic traits, considering four different selection strategies

Selection strategy	Response	\bar{X}_o	\bar{X}_s	DS	h^2	SG	SG%
Direct via GY	GY	5004.90	7101.97	2097.07	0.78	1635.71	32.68
	DH	61.02	56.04	4.98	0.82	4.08	6.69
	PH	108.64	109.27	0.63	0.70	0.44	0.41
	HGM	4.10	4.12	0.02	0.85	0.02	0.41
Direct via DH	GY	5004.90	6630.10	1625.20	0.78	1267.66	25.33
	DH	61.02	53.38	7.64	0.82	6.26	10.27
	PH	108.64	110.41	1.77	0.70	1.24	1.14
	HGM	4.10	4.22	0.12	0.85	0.10	2.49
Indirect via HTP	GY ^a	5004.90	6777.67	1772.77	0.78	1382.76	27.63
	DH ^b	61.02	54.49	6.53	0.82	5.35	8.78
	PH	108.64	105.89	2.75	0.70	1.93	1.77
	HGM	4.10	4.27	0.17	0.85	0.14	3.52
MGIDI	GY	5004.90	6716.27	1711.37	0.78	1334.87	26.67
	DH	61.02	54.73	6.29	0.82	5.16	8.45
	PH	108.64	103.08	5.56	0.70	3.89	3.58
	HGM	4.10	4.15	0.05	0.85	0.04	1.04

\bar{X}_o , original mean; \bar{X}_s , mean of the selected genotypes (20% best performing genotypes); DS, differential of selection; h^2 , broad sense heritability; SG, predicted selection gain. HTP, high-throughput phenotyping; MGIDI, multi trait genotype ideotype distance index; DH, days to heading; GY, grain yield; HGM, hundred grain mass; PH, plant height

^aEfficiency of indirect selection of GY via HTP=0.85; ^bEfficiency of indirect selection of DH via HTP=0.85

progressed, the correlation coefficients among the vegetation indices and the agronomic traits (especially GY and DH) increased. In early stages, genotypes seem to have similar reflectance patterns, which may lead to the absence of association among vegetation indices and agronomic traits. As the season progresses, a series of biochemical, metabolic, and physiological dynamics triggers different responses from genotypes, which are reflected in cycle, disease resistance, canopy cover, etc. The vegetation indices can capture these features, making them associated to the target agronomic traits and contributing to the significance of the correlation coefficients.

In this study, negative genetic correlations among most of the vegetation indices and GY and positive correlations among vegetation indices and DH were found from flowering to maturation. These results agree with previous studies (e.g., Kyratzis et al. 2017). Other studies, however, reported positive correlations among vegetation indices and target agronomic traits (Gizaw et al. 2016; Hassan et al. 2019). Most vegetation indices provide insight into the status of plants photosynthetic capacity. They are related to the duration of the vegetative period, grain filling or

maturation phases, as well as biomass accumulation. Genotypes presenting longer permanence in some specific development stages are generally the oldest in the panel used in this study and have higher cycle and lower grain yield genetic values (Lima et al. 2022). For this reason, genotypes with high genetic values for most vegetation indices also present the highest genetic values for DH and lower values for GY, which explains the genetic correlation results in this study.

Factor analysis is a procedure used to reduce the dimensionality of the data and obtain relevant information about the relationships between traits. This approach was successfully used by Santana et al. (2022b) for a comprehensive analysis of the relationship between vegetation indices and morphological and agronomic traits in maize. The results of the factor analysis obtained in the present study are in line with the genetic correlation analysis, i.e., most of the vegetation indices from flowering to maturation were significantly correlated with GY and DH.

An interesting feature of factor analysis is that, based on factor loadings, it allows calculating factor scores, which are new variables that can be used in selection strategies. In this study, the factor scores

obtained by the first factor, called HTP, were used to design an indirect selection strategy. In addition, the gains achieved in the target agronomic traits were compared with other selection strategies. Higher balance and increased desired genetic gains were observed for the agronomic traits when indirect selection was carried out via HTP. The genetic gains achieved with HTP for the target agronomic traits were as balanced and desirable as those achieved with MGIDI, which is an optimal selection index based on multiple traits that overcomes the fragility of classical linear indices (Olivoto and Nardino 2021). These results confirm the potential use of vegetation indices for the indirect selection of superior tropical wheat genotypes for target agronomic traits, which may deliver high-throughput, non-invasive, and less expensive selection.

The findings in this study demonstrate how the use of indirect selection strategies using vegetation indices derived from HTP can impact positively the genetic gains in a breeding program. Nevertheless, it is important to highlight that the gains with indirect selection is directly connected to the correlation between the traits. This fact has important implications and must be considered by plant breeders before adopting the strategies proposed here. Such implications are discussed below.

It is important to mention that divergences on the nature of the correlation coefficients found in different studies, as discussed previously, might be explained by the fact that genetic correlation can be estimated from the resemblance between relatives (Lynch and Walsh 1998). As demonstrated by Hill (2013) if $cov(X, Y)$ is the simple covariance between trait X on the parent and trait Y on the offspring, $2cov(X, Y)$ is an estimate of the (additive) genetic covariance (cov_A), and $\sqrt{\frac{[cov(X, Y)cov(Y, X)]}{[cov(X, X)cov(Y, Y)]}}$ is an estimate of the genetic correlation (r_A). Consequently, because of real differences among populations and species and due to sampling error, genetic correlation estimates are not all consistent (Hill 2013).

In light of this, the nature and magnitude of the genetic correlation among vegetation indices and agronomic traits may vary according to the panel and the growing and environmental conditions. Thus, it is not reasonable to assume that other panels (from other species or either wheat panels) will follow the same trend of the current panel.

Consequently, plant breeders and statisticians are encouraged to concentrate efforts on the examination of the interrelationships among the traits in their panels prior to any decision-making regarding selection.

Because the estimates of the genetic correlations between the traits across panels might present different patterns, a comprehensive analysis of the relationship between primary target and secondary traits in wheat breeding programs is essential for designing selection strategies to achieve desired genetic gains (Silva et al. 2022; Mezzomo et al. 2020). The use of approaches that allow the decomposition of the association between the traits into total, direct and indirect effects (Momen et al. 2019; Suela et al. 2023) could be potentially used in the intermediate stages of the plant breeding programs to better understand the interrelationships between the traits. This could guide the plant breeders regarding the decisions in the selection process, i.e., whether to use direct, indirect or index-based selection.

Some previous studies highlighted the suitability of UAV-based high-throughput phenotyping approaches for optimizing the selection process in wheat (Krause et al. 2020; Volpato et al. 2021). However, in tropical wheat breeding, information regarding the use of these approaches is scarce. The current study is the first attempt to provide insights for tropical wheat genotypes regarding: the optimal stages to perform image acquisition and capture the variability among genotypes, using vegetation indices; the heritability and accuracy of several vegetation indices across wheat development; the magnitude and direction/nature of the relationship among vegetation indices and target agronomic traits; and the efficiency of indirect selection to optimize evaluation and selection.

Further studies should also be conducted to evaluate the suitability of vegetation indices derived from UAV-based imagery for selection of tropical wheat segregating populations, progenies, and elite lines. Also, other target agronomic, physiological, and morphological traits should be evaluated. In the future, the available phenotypic data may also be associated with genomic information, such as molecular markers, to better understand the associations between phenotype and genotype in tropical wheat.

Conclusion

The optimal stages to carry out image acquisition using UAV-platforms for building multi-spectral vegetation indices are from flowering to maturation. Heritability and accuracy estimated for most vegetation indices across these wheat development stages were high, confirming that vegetation indices can be used in selection strategies. The most significant correlations among vegetation indices and target agronomic traits were also observed from flowering to maturation. Vegetation indices are potentially suitable to be used in indirect selection strategies, in particular for grain yield and cycle.

The findings of this study provide exciting insights on the application of vegetation indices derived from UAV platforms for the indirect selection of tropical wheat genotypes, and the methods describe here could be potentially adopted by other plant breeding programs. Nevertheless, we reiterate that prior to any decision-making regarding selection, plant breeders, along with their multidisciplinary team, should better investigate and understand the interrelationships between the traits to be improved in their panels.

Author contributions Conceptualization: Caique Machado e Silva, Maicon Nardino; Data Curation: Caique Machado e Silva, Henrique Caletti Mezzomo, João Paulo Oliveira Ribeiro, Victor Silva Signorini, Gabriel Wolter Lima, Marcelo Fagundes Portes; Formal analysis: Caique Machado e Silva, Marcelo Fagundes Portes; Investigation: Caique Machado e Silva, Maicon Nardino, Lucas de Paula Corredo, Maicon Nardino; Methodology: Caique Machado e Silva, Marcelo Fagundes Portes, Maicon Nardino; Software: Caique Machado e Silva, Eduardo Filipe Torres Vieira, Marcelo Fagundes Portes; Visualization: Caique Machado e Silva, Eduardo Filipe Torres Vieira; Funding acquisition: Maicon Nardino; Project administration: Maicon Nardino; Supervision: Gota Morota, Maicon Nardino; Writing-review and editing: Gota Morota, Lucas de Paula Corredo, Maicon Nardino; Writing-original draft: Caique Machado e Silva, Gota Morota, Maicon Nardino.

Funding This work was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – Finance Code 312791/2021-6 and by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) – Finance Code APQ 00520-21 and APQ-02668-21.

Data availability The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose. The authors declare no conflicts of interest.

References

- Agisoft (2019) Agisoft Metashape. In: Agisoft.com. <https://www.agisoft.com/>
- Alvares CA, Stape JL, Sentelhas PC et al (2013) Köppen's climate classification map for Brazil. *Meteorol Z* 22:711–728. <https://doi.org/10.1127/0941-2948/2013/0507>
- Anderson T (2003) An introduction to multivariate statistical analysis, 3rd edn. Wiley, New York
- Babar MA, Reynolds MP, van Ginkel M et al (2006) Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. *Crop Sci* 46:1046–1057. <https://doi.org/10.2135/cropsci2005.0211>
- Barzin R, Lotfi H, Varco JJ, Bora GC (2022) Machine learning in evaluating multispectral active canopy sensor for prediction of corn leaf nitrogen concentration and yield. *Remote Sens* 14:120. <https://doi.org/10.3390/rs14010120>
- Borges MVV, de Oliveira GJ, Batista TS et al (2022) High-throughput phenotyping of two plant-size traits of Eucalyptus species using neural networks. *J For Res* 33:591–599. <https://doi.org/10.1007/s11676-021-01360-6>
- Candiago S, Remondino F, De Giglio M et al (2015) Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sens* 7:4026–4047. <https://doi.org/10.3390/rs70404026>
- Casagrande CR, Sant'ana GC, Meda AR et al (2022) Association between unmanned aerial vehicle high-throughput canopy phenotyping and soybean yield. *Agron J* 114:1581–1598. <https://doi.org/10.1002/agj2.21047>
- de Resende MDV, Alves RS (2020) Linear, generalized, hierarchical, bayesian and random regression mixed models in genetics/genomics in plant breeding. *Funct Plant Breed J* 2:1–13
- de Resende MDV, Alves RS (2022) Statistical significance, selection accuracy, and experimental precision in plant breeding Statistical significance, selection accuracy, and experimental precision in plant breeding. *Crop Breed Appl Biotechnol* 22(3):e42712238. <https://doi.org/10.1590/1984-70332022v22n3a31>
- de Silva CM, Lima GW, Mezzomo HC et al (2022) Canonical correlations between high and low heritability wheat traits via mixed models. *Ciência Rural* 53:e20210798. <https://doi.org/10.1590/0103-8478cr20210798>
- Feng L, Chen S, Zhang C et al (2021) A comprehensive review on recent applications of unmanned aerial vehicle remote sensing with various sensors for high-throughput plant phenotyping. *Comput Electron Agric* 182:106033. <https://doi.org/10.1016/j.compag.2021.106033>
- Ferreira FM, Leite RV, Malikuski RG et al (2022) Bioenergy elephant grass genotype selection leveraged by spatial modeling of conventional and high-throughput

- phenotyping data. *J Clean Prod* 363:132286. <https://doi.org/10.1016/j.jclepro.2022.132286>
- Frels K, Guttieri M, Joyce B et al (2018) Evaluating canopy spectral reflectance vegetation indices to estimate nitrogen use traits in hard winter wheat. *Field Crop Res* 217:82–92. <https://doi.org/10.1016/j.fcr.2017.12.004>
- Gitelson AA, Merzlyak MN (1997) Remote estimation of chlorophyll content in higher plant leaves. *Int J Remote Sens* 18:2691–2697. <https://doi.org/10.1080/014311697217558>
- Gitelson AA, Kaufman YJ, Merzlyak MN (1996) Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens Environ* 58:289–298. [https://doi.org/10.1016/s0034-4257\(96\)00072-7](https://doi.org/10.1016/s0034-4257(96)00072-7)
- Gizaw SA, Garland-Campbell K, Carter AH (2016) Use of spectral reflectance for indirect selection of yield potential and stability in Pacific Northwest winter wheat. *Field Crop Res* 196:199–206. <https://doi.org/10.1016/j.fcr.2016.06.022>
- Hassan MA, Yang M, Rasheed A et al (2019) A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Sci* 282:95–103. <https://doi.org/10.1016/j.plantsci.2018.10.022>
- Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31:423. <https://doi.org/10.2307/2529430>
- Hill WG (2013) Genetic correlation. *Brenner's Encyclopedia of Genetics* 237–239. <https://doi.org/10.1016/b978-0-12-374984-0.00611-2>
- Huete AR (1988) A soil-adjusted vegetation index (SAVI). *Remote Sens Environ* 25:295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-x](https://doi.org/10.1016/0034-4257(88)90106-x)
- Jöreskog KG (1967) Some contributions to maximum likelihood factor analysis. *Psychometrika* 32:443–482. <https://doi.org/10.1007/bf02289658>
- Justice CO, Vermote E, Townshend JRG et al (1998) The moderate resolution imaging spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Trans Geosci Remote Sens* 36:1228–1249. <https://doi.org/10.1109/36.701075>
- Kaiser HF (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23:187–200. <https://doi.org/10.1007/bf02289233>
- Khan Z, Rahimi-Eichi V, Haefele S et al (2018) Estimation of vegetation indices for high-throughput phenotyping of wheat using aerial imaging. *Plant Methods* 14:20. <https://doi.org/10.1186/s13007-018-0287-6>
- Krause MR, Mondal S, Crossa J et al (2020) Aerial high-throughput phenotyping enabling indirect selection for grain yield at the early-generation seed-limited stages in breeding programs. *Crop Sci* 60:3096. <https://doi.org/10.1002/csc2.20259>
- Kyratzis AC, Skarlatos DP, Menexes GC et al (2017) Assessment of vegetation indices derived by UAV imagery for durum wheat phenotyping under a water limited and heat stressed mediterranean environment. *Front Plant Sci* 8:1114. <https://doi.org/10.3389/fpls.2017.01114>
- Li H, Zhang Y, Lei Y et al (2020) Evaluating different non-destructive estimation methods for winter wheat (*Triticum aestivum* L.) nitrogen status based on canopy spectrum. *Remote Sens* 12:95. <https://doi.org/10.3390/rs12010095>
- Lima GW, e Silva CM, Mezzomo HC et al (2022) Genetic diversity in tropical wheat germplasm and selection via multitrait index. *Agron J* 114:887–899. <https://doi.org/10.1002/agj2.20991>
- Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Inc. Publishers, Sunderland, p 980
- Mezzomo HC, Casagrande CR, e Silva CM et al (2020) Association between physiological and agronomic traits and selection of tropical wheat. *J Crop Sci Biotechnol* 24:167–177. <https://doi.org/10.1007/s12892-020-00069-y>
- Mir RR, Reynolds M, Pinto F et al (2019) High-throughput phenotyping for crop improvement in the genomics era. *Plant Sci* 282:60–72. <https://doi.org/10.1016/j.plantsci.2019.01.007>
- Momen M, Campbell MT, Walia H, Morota G (2019) Utilizing trait networks and structural equation models as tools to interpret multi-trait genome-wide association studies. *Plant Methods* 15:107. <https://doi.org/10.1186/s13007-019-0493-x>
- Momen M, Bhatta M, Hussain W et al (2021) Modeling multiple phenotypes in wheat using data-driven genomic exploratory factor analysis and Bayesian network learning. *Plant Direct* 5:00304. <https://doi.org/10.1002/pld3.304>
- Moreira FF, Hearst AA, Cherkauer KA, Rainey KM (2019) Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. *Plant Methods*. <https://doi.org/10.1186/s13007-019-0519-4>
- Morota G, Jarquin D, Campbell MT, Iwata H (2022) Statistical methods for the quantitative genetic analysis of high-throughput phenotyping data. *Methods Mol Biol* 2539:269–296. https://doi.org/10.1007/978-1-0716-2537-8_21
- Olivoto T, Lúcio AD (2020) metan: an R package for multi-environment trial analysis. *Methods Ecol Evol* 11:783–789. <https://doi.org/10.1111/2041-210x.13384>
- Olivoto T, Nardino M (2021) MGIDI: toward an effective multivariate selection in biological experiments. *Bioinformatics* 37:1383–1389. <https://doi.org/10.1093/bioinformatics/btaa981>
- Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika* 58:545. <https://doi.org/10.2307/2334389>
- Prudnikova E, Savin I, Vindeker G et al (2019) Influence of soil background on spectral reflectance of winter wheat crop canopy. *Remote Sens* 11:1932. <https://doi.org/10.3390/rs11161932>
- R Core Team (2022) R: the R project for statistical computing. In: R-project.org. <https://www.r-project.org/>
- Raper TB, Varco JJ (2014) Canopy-scale wavelength and vegetative index sensitivities to cotton growth parameters and nitrogen status. *Precis Agric* 16:62–76. <https://doi.org/10.1007/s11119-014-9383-4>
- Revelle W (2020) psych: procedures for psychological, psychometric, and personality research. In: R-packages. <https://cran.r-project.org/web/packages/psych/index.html>
- Rouse JW, Haas RH, Schell JA, Deering DW (1974) Monitoring vegetation systems in the Great Plains with ERTS. NASA Goddard Space Flight Center 3d ERTS-1 Symp, vol 1, Sect A

- Rutkoski J, Poland J, Mondal S et al (2016) Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *J Mater* 6:2799–2808. <https://doi.org/10.1534/g3.116.032888>
- Santana DC, da Flores MS, Cotrim MF et al (2021) Genotype \times trait biplot and canonical correlations for spectral and agronomic traits in corn. *Agron J* 113:1197–1204. <https://doi.org/10.1002/agj2.20581>
- Santana DC, de Oliveira Cunha MP, dos Santos RG et al (2022a) High-throughput phenotyping allows the selection of soybean genotypes for earliness and high grain yield. *Plant Methods*. <https://doi.org/10.1186/s13007-022-00848-4>
- Santana DC, dos Santos RG, Teodoro LPR et al (2022b) Structural equation modelling and factor analysis of the relationship between agronomic traits and vegetation indices in corn. *Euphytica*. <https://doi.org/10.1007/s10681-022-02997-y>
- QGIS Software. In: [qgis.org](https://qgis.org/en/site/). <https://qgis.org/en/site/>
- Suela MM, Azevedo CF, Nascimento ACC et al (2023) Genome-wide association study for morphological, physiological, and productive traits in *Coffea arabica* using structural equation models. *Tree Genet Genomes*. <https://doi.org/10.1007/s11295-023-01597-8>
- Volpato L, Pinto F, González-Pérez L et al (2021) High throughput field phenotyping for plant height using UAV-based RGB imagery in wheat breeding lines: feasibility and validation. *Front Plant Sci* 12:591587. <https://doi.org/10.3389/fpls.2021.591587>
- Wilks SS (1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann Math Stat* 9:60–62
- You Q, Yang X, Peng Z et al (2018) Development and applications of a high throughput genotyping tool for polyploid crops: single nucleotide polymorphism (SNP) array. *Front Plant Sci* 9:104. <https://doi.org/10.3389/fpls.2018.00104>
- Yu H, Campbell MT, Zhang Q et al (2019) Genomic Bayesian confirmatory factor analysis and Bayesian network to characterize a wide spectrum of rice phenotypes. *G3: Genes|Genomes|Genetics* g3 400154:2019. <https://doi.org/10.1534/g3.119.400154>
- Zadoks JC, Chang TT, Konzak CF (1974) A decimal code for the growth stages of cereals. *Weed Res* 14:415–421. <https://doi.org/10.1111/j.1365-3180.1974.tb01084.x>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.