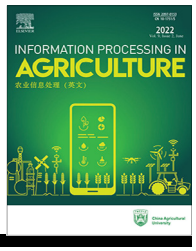
Available at www.sciencedirect.com

INFORMATION PROCESSING IN AGRICULTURE xxx (xxxx) xxx

journal homepage: www.keaipublishing.com/en/journals/information-processing-in-agriculture/



Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content

Joel Segarra [a,b], *Fatima Zahra Rezzouk* [a,b], *Nieves Aparicio* [c], *Jon González-Torralba* [d], *Iker Aranjuelo* [e], *Adrian Gracia-Romero* [a,b], *Jose Luis Araus* [a,b], *Shawn C. Kefauver* [a,b,*]

[a] *Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Barcelona 08028, Spain*
[b] *AGROTECNIO (Center for Research in Agrotechnology), Lleida 251981, Spain*
[c] *Agro-technological Institute of Castilla y León (ITACyL), Valladolid 47071, Spain*
[d] *Grupo AN, Tajonar 31192, Navarre, Spain*
[e] *Instituto de Agrobiotecnología (IdAB), CSIC- Gobierno de Navarra, Mutilva Baja 31192, Pamplona, Navarre, Spain*

ARTICLE INFO

Article history:
Received 3 December 2020
Received in revised form
5 May 2022
Accepted 17 May 2022
Available online xxxx

Keywords:
Wheat
Remote sensing
Sentinel-2
Grain nitrogen content
Phenotyping

ABSTRACT

Wheat grain quality characteristics have experienced increasing attention as a central factor affecting wheat end-use products quality and human health. Nonetheless, in the last decades a reduction in grain quality has been observed. Therefore, it is central to develop efficient quality-related phenotyping tools. In this sense, one of the most relevant wheat features related to grain quality traits is grain nitrogen content, which is directly linked to grain protein content and monitorable with remote sensing approaches. Moreover, the relation between nitrogen fertilization and grain nitrogen content (protein) plays a central role in the sustainability of agriculture. Both aiming to develop efficient phenotyping tools using remote sensing instruments and to advance towards a field-level efficient and sustainable monitoring of grain nitrogen status, this paper studies the efficacy of various sensors, multispectral and visible red–greenblue (RGB), at different scales, ground and unmanned aerial vehicle (UAV), and phenological stages (anthesis and grain filling) to estimate grain nitrogen content. Linear models were calculated using vegetation indices at each sensing level, sensor type and phenological stage. Furthermore, this study explores the up-scalability of the best performing model to satellite level Sentinel-2 equivalent data. We found that models built at the phenological stage of anthesis with UAV-level multispectral cameras using red-edge bands outperformed grain nitrogen content estimation (R^2 = 0.42, RMSE = 0.18%) in comparison with those models built with RGB imagery at ground and aerial level, as well as with those built with widely used ground-level multispectral sensors. We also demonstrated the possibility to use UAV-built multispectral linear models at the satellite scale to determine grain nitrogen content effectively (R^2 = 0.40, RMSE = 0.29%) at actual wheat fields.

* Corresponding author at: Permanent postal address at: Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Avinguda Diagonal, Barcelona, Spain.
E-mail address: sckefauver@ub.edu (S.C. Kefauver).
Peer review under responsibility of China Agricultural University.
https://doi.org/10.1016/j.inpa.2022.05.004
2214-3173 © 2022 China Agricultural University. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

© 2022 China Agricultural University. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

One of United Nations' Sustainable Development Goals is achieving improved nutrition while promoting sustainable agriculture [1]. Wheat (*Triticum*) together with maize (*Zea mays*) and rice (*Oryza sativa*) is one of the three most-cultivated cereals globally [2]. Moreover, wheat in the form of bread, as reported by Peña [3], has historically provided more nutrients to the world population than any other single food source. Nonetheless, in the last decades cereal breeder focus on improving grain yield has resulted in a genetic erosion of quality traits [4]. Meanwhile, products made with low-quality grains have been associated with nutritious and health issues for consumers [5–7]. Protein content, together with grain hardness or starch properties, among other factors, are grain characteristics that affect the end-use quality of wheat-based products [8]. Furthermore, reductions in grain protein content (GPC) have been closely linked to reductions in grain nitrogen content (GNC) after breeding programs focused on high-yielding ideotypes [9–11]. Vogel et al. [12], after analysing 12 600 genotypes at the USDA World Wheat Collection, reported a range of GPC between 7 and 22%, corresponding to 1.2–3.7% of nitrogen in grains, which accounts for an average value of around 16% of grain protein composition [14]. Specifically in Spain, where this study takes place, Sánchez-Garcia et al. [11] observed a decrease of 0.21% per year in GPC during the XX century and Acreche and Slafer [13] observed that GNC was reduced in modern varieties after breeding for increased grain yields.

GPC is determined by genetic and environmental factors (notably the availability of nitrogen fertilization and water [14–16]) and is directly linked to GNC, which can be monitored with non-destructive remote sensing approaches at ground, unmanned aerial vehicle (UAV) and satellite levels. GNC cannot be directly observed as there are resolution (grain size) and spectral sensing limitations (nitrogen content in plants is not discernible because nitrogen absorption features are obscured by liquid water in the crop canopy [17–19]). However, chlorophyll spectral features can be sensed within the visible and red-edge parts of the spectrum and chlorophyll-related vegetation indices can be used as proxies of crop nitrogen concentration [20], which in turn may be related to GNC. In this sense, a precise assessment of the spectral features could likely allow differentiating nitrogen-based proteins from other constituents at canopy level [21].

Accessible and effective ways to estimate GNC can provide optimized strategies for both phenotyping high-quality lines at experimental fields and to assist farmers with crops management at field level. High-quality bread wheat lines will experience an increasing demand due to current nutritious and health challenges [22,23]. Therefore, low-cost and accessible quality-oriented phenotyping is pivotal. Furthermore, developing an improved monitoring of nitrogen accumulation in wheat grains at field-level is central for crops management. It is a key aspect regarding both the importance of environmental factors on GNC and the aim to advance towards high-quality grains within a sustainable agriculture [24].

Various studies related with GNC monitoring and remote sensing have been published in recent years. At the ground level, different authors [25–28] showed significant positive correlations between canopy spectral indices and nitrogen accumulation in wheat grains. At the UAV level, several studies have successfully correlated vegetation indices, calculated with multispectral instruments, with nitrogen accumulated in wheat grains [29–31]. At the satellite level, until recently GPC or GNC monitoring was limited due to the spatial, spectral and temporal resolutions of the orbiting satellites [32]. Several GNC related studies, showed potentialities regarding regional/field-level GNC monitoring, for instance with the Landsat satellite [33–35] or its combination with MODIS satellite [36]. Notwithstanding, Sentinel-2's recent fully operational and improved spectral, spatial and temporal resolutions have opened up opportunities [37], such as for the case of GNC monitoring [38].

Besides focusing on single sensing levels, few studies have dealt with various platforms such as satellite/ground [39] or UAV/satellite [40] spectral data to monitor GNC. The study here presented goes a step forward and works across phenotypes with remotely sensed data on three sensing scales (ground, UAV and satellite). Moreover, it discusses the most efficient sensors, either visible red–green–blue (RGB) or multispectral, and phenological stages for GNC monitoring. Vegetation indices sensitive to chlorophyll content, greenness and biomass were calculated at each level, and linear models were developed to find the most optimal GNC estimators. This research paper aims to develop accessible and reliable GNC empirical models with remote sensing data that can support cereal phenotyping and crop management. In order to achieve this, in experimental plots growing various wheat lines, RGB and multispectral sensors were deployed at both ground and UAV levels. Vegetation indices calculated from the sensed spectral reflectance data were correlated with GNC using linear models. Regarding the satellite level, the most suitable GNC estimation model was applied to actual farmer fields with equivalent Sentinel-2 data over two years. For each farmer's field, carbon isotopic composition ($\delta^{13}C$) was obtained as an indicator for photosynthetic performance, in this study we aim to use this indicator as a proxy to characterize water conditions across agroclimates which may eventually affect GNC. Thus, water stress causes a decrease in stomatal conductance which subsequently will increase the $\delta^{13}C$ of photoassimilates and then of grains. Nonetheless, carbon isotopic composition was not used in the models as we focus on spectral information in modelling. The research was structured around 4 questions to investigate which sensors, sensing levels and phenological stages were the most efficient for GNC monitoring in wheat:

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

(i) Which imagery, RGB vs Multispectral, can contribute the most to estimate wheat GNC?
(ii) Which sensing level, UAV vs ground images, is more effective to generate models for GNC monitoring?
(iii) How relevant is phenology for GNC estimation?
(iv) Are models built with ground or UAV images upscalable to equivalent Sentinel-2 images?

This study concludes by testing the most optimal GNC model to estimate field-level GNC on an agro-ecosystem scale.

2. Materials and methods

2.1. Study site and field data

The study area was located at two sites. The experimental plots were located at the research station of Zamadueñas in Valladolid (Spain) and farmer field trials were located in three climatic regions within the community of Navarre in Northern Spain (Fig. 1). Valladolid has annual rainfalls ranging from 400 to 500 mm. Regarding Navarre, it incorporates several agroclimates, with rainfall averaging from 800 to 1 000 mm in the northern areas, 300 to 800 mm in the middle and around 300 mm in the southern areas.

At the Zamadueñas research station a set of 38 post green revolution and advanced wheat lines were grown in experimental plots. The design consisted of three randomized blocks with one replicate per genotype in each block, two plots (one in each corner) were sown to avoid edge effects. The cultivars had support irrigation due to 2016–2017 severe droughts in the region (a total of 129 mm of rainfall plus 60 mm of support irrigation). At the panel, 31 bread wheat (*Triticum aestivum* L.) lines and two durum wheat (*Triticum turgidum* L. subsp. *durum* (Desf) Husn.) cultivars were utilized (Table 1). Five genotypes were no further utilized for GNC estimation in this study due to their unevenness and unsuitability for modelling. The 33 genotypes utilized are described in Table 1. The basal fertilization in the plots consisted in 300 kg/ha of 8–15–15 (NPK) in addition to top-dressing refertilization of 300 kg/ha of Nitrosylsulphuric acid 27% at the phenological stage of stem elongation. Plots were 6 m long and 1.5 m wide and were sown on December 2nd, 2016. At maturity, grains were sampled for further analysis. On July 20th, 2017, the plots were harvested. The experimental plots were visited at two phenological stages, in middle-May 2017, at anthesis, and in the second week of June 2017 at grain filling.

In Navarre, during 2017–2018 and 2018–2019 growing seasons 19 rain-fed famers' fields were monitored. The agricultural fields were growing either Marcopolo (RAGT) or Camargo (Disasem) commercial lines of bread wheat. As reported by the farmers, the analysed fields received a basal fertilization with either superphosphate 45% or pig slurry, top-dressing fertilization was also applied with granulated urea at most fields. In the northern region wheat was sown around the 20th of October, in the middle region around the 28th of October, and in the southern around the 30th of October. The harvest was collected around the middle of July in the North and during the first week of July in the Middle and South.

2.2. Nitrogen content and carbon isotope composition in grains

Three half-square-meter representative areas were manually harvested at each farmers' field in Navarre (Spain), the grains

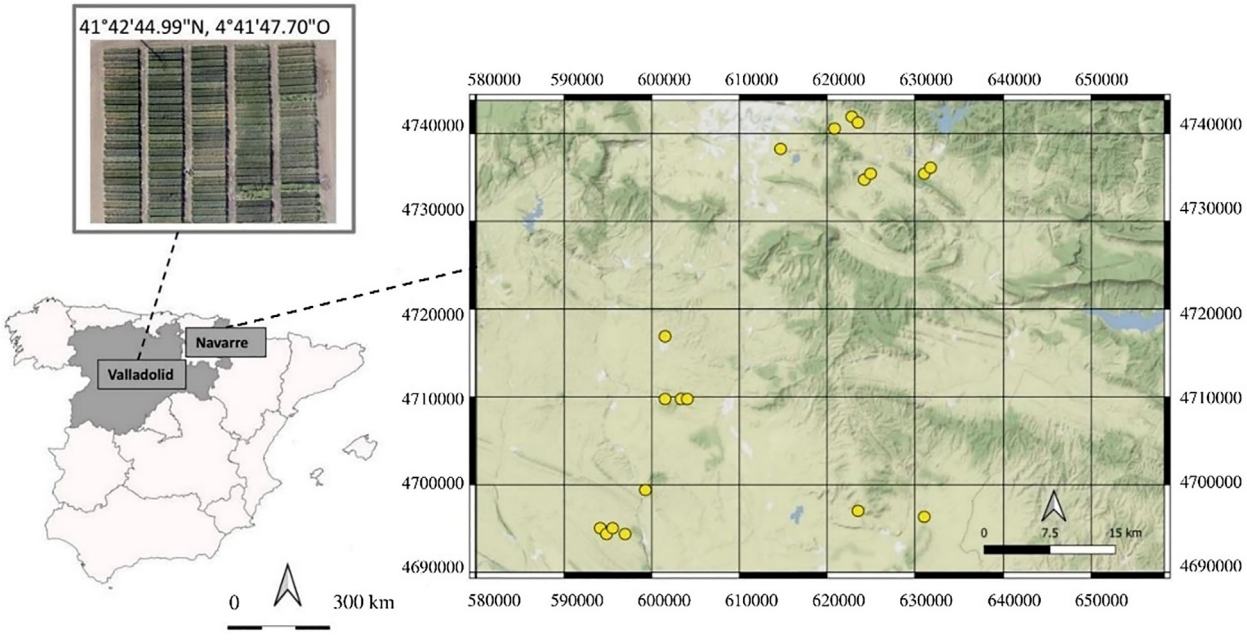


Fig. 1 – Maps of the study sites, Zamadueñas research station in Valladolid, and farmers' fields in Navarre (EPSG: 25,830 (UTM 30N, ETRS89) are indicated with a yellow dot. The map of whole Spain is also shown. The background terrain map is based on OpenStreetMap data.

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

Table 1 – The utilized genotypes to develop the grain nitrogen content estimation models, the breeders of the lines are referenced.

Name	Breeder	Species
Atomo	LG Seeds	*Triticum aestivum*
Bisanzio	MAS Seeds	
Galera	LG Seeds	
Togano	SARL Raoul Rolly	
Eneas	Dafisa	
08THES1262	Batlle	
Albertus	Saatzucht Donau	
Algoritmo	RAGT	
Bologna	Batlle	
Dolly	SARL Raoul Rolly	
Forcalli	KWS	
Ingenio	CC Benoist	
Mecano	Secobra	
Rebelde	Agri-Obtentions	
Rimbaud	Secobra	
Tribat	Batlle	
Chambo	LG Seeds	
Complice	Florimond Desprez	
Cosmic	Lemaire Deffontaines	
Ippon	Florimond Desprez	
Nemo	Secobra	
Oregrain	Florimond Desprez	
PR22R58	Pioneer Hi-Bred	
Soberbio	Caussade	
Soisson	Florimond Desprez	
MH1307	KWS	
MH1341	KWS	
MH1444	KWS	
MH1411	KWS	
Craklin	LG Seeds	
Marcopolo	RAGT	
Mimmo	PRO.SE.ME	*Triticum durum*
Credit	PRO.SE.ME	

were subsequently mixed. The 19 rain-fed famers' fields used in this study ranged from 3.14 to 11.19 ha. At Zamadueñas research station in Valladolid (Spain), one sample was taken for each harvest of the experimental plots. The harvested mature grains were dried at 60 °C for 48 h and grinded (Mixer Mill MM 400; Retsch GmbH, Haan; Germany). After this, 1 mg of each sample was weighed in tin capsules to analyse carbon and nitrogen stable isotope signatures. Total nitrogen content was expressed as the percentage (%) of total nitrogen on dry matter basis. The carbon isotopic composition ($\delta^{13}C$) was expressed following the Eq. (1):

$$\delta^{13}C\ (‰) = [(R_{sample}/R_{standard}) - 1] \times 1000 \quad (1)$$

where R_{sample} is the $^{13}C/^{12}C$ ratio of the sample, while $R_{standard}$ is the molar abundance ratio of the secondary standard calibrated against the primary standard Pee Dee Belemnite ($\delta^{13}C$) [41]. Different secondary standards were used for carbon (IAEA–CH7, IAEA–CH6 and IAEA-600, and USGS 40) isotope analyses. For further information see Rezzouk et al. [42]. Isotopes and elemental analyses were performed employing an elemental analyser operating in a continuous flow mode with a mass spectrometer (Delta C IRMS; ThermoFinnigan, Bremen; Germany), at the Scientific and Technical facilities of the University of Barcelona (Centres Científics i Tecnològics de la Universitat de Barcelona (CCiTUB)).

2.3. *Instruments and indices*

At the ground level, RGB images were taken zenithally at 80 cm above the canopy with a Sony ILCE-QX1 (Sony Europe Limited, Brookland, United Kingdom), which has a resolution of 20.1 megapixels, is equipped with a 23.2 mm × 15.4 mm sensor (type CMOS Exmor HD), and uses a 16 mm focal lens with an exposure time of 1/60 s. Also at ground level, the normalized difference vegetation index (NDVI) was measured at each plot using a GreenSeeker (Trimble, Sunnyvale, CA, USA), which is a hand held spectroradiometer with an active self-illuminated sensor in red (660 ± 10 nm) and near infrared (780 ± 15 nm) [43]. As described by Barmeier and Schmidhalter [44], the measurements were taken at a constant height of 60 cm from above the soil with a perpendicular position to the canopy.

At aerial level, RGB images were taken with a Lumix GX7 (Panasonic, Osaka, Japan) at 50 m, which has a resolution of 16 megapixels with an image sensor size of 17.3 × 13.0 mm (type Live MOS) and uses a 20 mm focal lens with an exposure time of 1/8000 s. The camera was mounted on a manually

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

controlled UAV (Drone Mikrokopter OktoXL 6S12, Moormerland, Germany). Aerial multispectral images were taken with a Tetracam micro-MCA at 50 m a.g.l. with the same UAV used with the RGB camera. The multispectral sensors of the Tetracam micro-MCA consists of 11 bands with 450 ± 40 nm, 550 ± 10 nm, 570 ± 10 nm, 670 ± 10 nm, 700 ± 10 nm, 720 ± 10 nm, 780 ± 10 nm, 840 ± 10 nm, 860 ± 10 nm, 900 ± 20 nm and 950 ± 40 nm of spectral resolution (band centre and FWHM of the filter spectral response curve). Moreover, the camera possesses one sensor dedicated to calibration (Incident Light Sensor, ILS), which provides band-by-band reflectance calibration in real-time, correcting the 11 bands to reduce atmospheric and sun angle effects and provide image by image correction to at-sensor reflectance.

Aerial multispectral images were spatially aligned and radiometrically calibrated using PixelWrench 0.2 version 1.2.2.2 (Tetracam, Chatsworth, CA, USA). Aerial RGB and multispectral image mosaics were reconstructed using Agisoft Photoscan Pro (Agisoft LLC, St. Petersburg, Russia, https://www.agisoft.com) [45]. This software overlaps images (a minimum of 30 images with at least 80% overlapping is required) and removes UAV flight effects to produce accurate orthomosaics. Once the whole experimental field image was resampled, the plots were cropped and processed using the MosaicTool software (developed by S. C. Kefauver and others, for details, https://integrativecropecophysiology.com/software-development/mosaictool/, freely available at https://gitlab.com/sckefauver/MosaicTool/, University of Barcelona, Barcelona, Spain) integrated as a plugin for the open source image analysis platform FIJI (Fiji is Just ImageJ; https://fiji.sc/Fiji) [46].

RGB vegetation indices collected from both ground and aerial platforms were obtained using an updated version of the original Breedpix 2.0 software [47]. From the CIE (Commission Internationale de l'Eclairage; the International Commission on Illumination), CIELab colour space was used to calculate lightness component, a* and b* dimensions; and CIELuv colour space was used to calculate u* and v* coordinates. The b* and v* express the blue to yellow spectrum, while a* and u* represent the green to red spectrum. HIS colour space, referring to the components Hue, Saturation, and Intensity, was also used. Hue is described as the chroma traversing the visible spectrum in the form of an angle between 0° and 360°, where 0° and 360° are decrypted into red, 60° into yellow, 120° into green and 180° into cyan. Derived from the Hue, the indices Green Area (GA) and Greener Area (GGA) were described as the fraction area presented by green pixels in the image, and which Hue ranges from 60° to 180° (GA) and from 80° to 180° (GGA). While GA gives a broader perception to canopy greenness, GGA excludes yellowish green pixels [47,48]. The following RGB indices were also calculated: the triangular green index (TGI) [49], which estimates chlorophyll concentration in canopies; the normalized green–red difference index (NGRDI) [50], which compares the differences between the green and red bands in a calculation similar to NDVI but with less marked differences and less signal saturation; and the crop senescence index (CSI) [51,52], which combines GA and GGA to provide an index of canopies greenness. The instruments and indices calculated are summarized in Table 2.

Regarding multispectral indices, 14 vegetation indices (VIs) were calculated (summarized in Table 2): CCCI (Canopy Chlorophyll Content Index) [53], CCI (Chlorophyll/Carotenoid Index) [54], Cl Green (Chlorophyll Index Green) [55], Cl Red Edge (Chlorophyll Index Red Edge) [55], CVI (Chlorophyll Vegetation Index) [56], EVI (Enhanced Vegetation Index) [57], MCARI (Modified Chlorophyll Absorption Ratio Index) [90], NDVI (Normalized Difference Vegetation Index) [59], OSAVI (Optimized Soil Adjusted Vegetation Index) [60], RDVI (Renormalized Difference Vegetation Index) [61], SAVI (Soil Adjusted Vegetation Index) [62], TCARI (Transformed Chlorophyll Absorption Index) [58], TCARI/OSAVI [58] and TCI (Triangular Chlorophyll Index) [63]. Theses indices were chosen due to its sensitiveness to chlorophyll content and capacities to monitor plants' nitrogen related features. At aerial level, VIs were calculated on Microsoft Excel (2010) with the obtained reflectance from the Tetracam micro-MCA mounted on a UAV.

2.4. *Grain nitrogen content estimation*

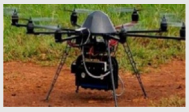
In order to see the GNC gradient in the genotypes grown at the experimental site, a post hoc analyses was calculated on R studio [64] ("agricolae"). After, stepwise multilinear regressions and simple regressions between GNC and VIs were calculated in order to obtain empirical models for GNC estimation. The VIs means were extracted for each field. Preliminary, simple R-Pearson regressions of all the VIs against GNC were calculated with the function "cor.test" on R studio [64] for each phenological stage (anthesis and grain filling), sensing level (ground and aerial) and sensor type (RGB and multispectral). In order to represent these results understandably, they were plotted in a heatmap, the obtained R-Pearson values were added in a matrix and plotted on R studio [64] ("ggplot2").

At aerial and ground level, stepwise multilinear models were calculated with RGB and multispectral indices against GNC with the exception of ground multispectral data, which only had one parameter (NDVI) and thus a simple regression was calculated. In all cases a 70% of each dataset (RGB ground, RGB aerial, multispectral ground and multispectral aerial) was used for training the model while 30% was used for validation. In order to find the most suitable combination of parameters (indices) for the stepwise multilinear regression the library "caret" was used on Rstudio. With the functions "trainControl" and "train" a 10-fold cross validation was used to estimate the average predicting error, root mean square (RMSE), and select the most optimal model. With the best models, the variance inflator factor (VIF) was measured as it is an efficient method to asses collinearity [65–67], which strongly limits stepwise selection method. With the 30% of the dataset that was not used for the training the models were tested.

In order to ease the reading of the models' validation results, a bar chart was plotted on Rstudio ("ggplot2") to show R square, RMSE and the significance of the tested linear models regarding the two phenological stages (anthesis and grain filling), type of sensor (RGB and multispectral images), and sensing level (ground and aerial level). The original graphs of validation are shown in the Appendix B.

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

Table 2 – Summary of sensing levels, sensors type, instruments used, and vegetation indices calculated for each case. The spatial resolution of the instruments is also indicated. The reference (Ref.) is specified too.

Level	Spectral data	Sensor	Index	Description	Ref.
Ground	Multispectral	**GreenSeeker** (60 cm spatial resolution) 	**NDVI**	(B840-B670)/(B840 + B670)	[59]
	RGB	**Sony ILCE-QX1** (0.01 mm/pixel) 	**Intensity**	HIS color space	[48]
			Saturation	HIS color space	[48]
			Lightness	CIElab color space	[48]
			a*	CIElab color space	[48]
			b*	CIElab color space	[48]
			u*	CIEluv color space	[48]
			v*	CIElluv color space	[48]
UAV	RGB	**Lumix GX7** (5 cm/pixel) 	**Hue**	HIS color space	[48]
			GA	% of pixels in hue from 60° to 180°	[48,49]
			GGA	% of pixels in hue from 80° to 180°	[48,49]
			CSI	((GA-GGA)/GA) × 100	[51,52]
			NGRDI	(GREEN-RED)/(GREEN + RED)	[50]
			TGI	((GREEN-0.39) ×(RED-0.61)) × BLUE	[49]
	Multispectral	**Tetracam micro-MCA** (5 cm/pixel) 	**CCCI**	((B840B700)/(B840 + B700))/((B840-B670)/(B840 + B670)	[53]
			CCI	(B550-B570)/(B550 + B670)	[54]
			CI Green	(B840/B560)-1	[55]
			CI red edge	(B840/B700)-1	[55]
			CVI	(B840/B560) × (B670/B560)	[56]
			EVI	(2.5×(B840-B670))/(B840+(6 × B670)-(7.5 × B450) + 1)	[57]
			MCARI	(B700-B670) – 0.2 × (B700 – B550) × (B700/B670)	[90]
Satellite	Multispectral	**Sentinel-2 a + b** (10 and 20 m/pixel) 	**NDVI**	(B840-B670)/(B840 + B670)	[59]
			OSAVI	(1.16) × (B780 – B670)/(B780 + B670 + 0.16)	[60]
			RDVI	(B840-B670)/√(B840 + B670)	[61]
			SAVI	(B840-B670)/(B840 + B670)	[62]
			TCARI	3 × (B700 – B670) – 0.2 × (B700-B550) × (B700/B670)	[58]
			TCARI/OSAVI	TCARI/OSAVI	[58]
			TCI	1.2 × (B700 – B550) – 1.5 × (B670 – B550)×√(B700/B670)	[63]

2.5. *Sentinel-2 imagery*

The specific dates of Sentinel-2 images for 2018 and 2019 were selected as those closest to the most suitable phenological stage for GNC monitoring (Table 3). At regional scale the phenological stage was estimated with growing degree days (GDD) calculations as described by Arnold [68] (Eq. (2)).

Table 3 – Anthesis phenological stage estimated with GDD and Sentinel-2 closest date during 2018 and 2019 seasons.

	North	Middle	South
Anthesis (GDD)	10–05–18 / 08–05–19	17–05–18 / 15–05–19	26–04–18 / 29–04–19
Sentinel-2 (dates)	09–05–18 / 09–05–19	19–05–18 / 14–05–19	24–04–18 / 29–04–19

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

$$GDD = \sum_{Jun}^{Nov} \frac{T_{max}+T_{min}}{2} - T_{base} \quad (2)$$

where GDD is the growing degree days and $\sum_{jun}^{nov}$ indicates the sum throughout the season, November to June, of daily maximum and minimum temperatures (T_{max} and T_{min}) divided by 2, minus the base temperature (T_{base}), which in this case was considered 0 °C.

This data is used following Segarra et al. [69] study for the case of Navarre. The average temperature data, in order to estimate the crop phenological stages through the calculation of GDD in each zone (North, Middle and South), was obtained from 30 openly accessible regional government meteorological stations (https://www.meteo.navarra.es/estaciones), GDD and phenology for the wheat lines was obtained from the genotypes information as described by Goñi [70], corresponding to 1105 accumulated GDD for anthesis as the average for both (Camargo and Marcopolo wheat lines).

In Table 3 the estimated phenological date and the closest Sentinel-2 image is detailed. The Sentinel-2 images were downloaded from Copernicus Open Access Hub (https://scihub.copernicus.eu/) as a 2A product (Bottom of Atmosphere reflectance images) and with a maximum cloud cover threshold of 30%.

Sentinel-2 bands are shown in Table 4, with these bands the following most suitable VIs from the model, Cl red-edge [56] (Eq.(3)), TCARI [63] (Eq.(4)), and EVI [58] (Eq.(5)), were calculated on ArcGIS Pro 2.3.0 with the corresponding Sentinel-2 bands:

$$CI\,red-edge = \frac{(B8)}{(B5)} - 1 \quad (3)$$

$$TCARI = 3 \times (B5 - B4) - 0.2 \times (B5 - B3) \times \frac{B5}{B4} \quad (4)$$

$$EVI = 2.5 \times \frac{(B8 - B4)}{(B8 + (6 \times B4) - (7.5 \times B2) + 1} \quad (5)$$

3. Results

The GNC among the utilized wheat lines ranged between 2.07% and 2.74% (Fig. 2). The highest GNC was in the group formed by the genotypes Dolly, Mimmo, Galera, 08THES1262, Albertus and Rebelde, while the group with the lowest GNC was formed by PR22R58, Nemo, MH1411, Cosmic, Rimbaud, Oregrain, Marcopolo, Mecano, Chambo, Complice, Soberbio, Craklin and Togano (see the appendix Table A1 for the post hoc test and the averaged GNC per wheat line). The genotypes with the highest GNC ranged between 2.74% and 2.60%, while the lowest ranged from 2.26% to 2.07%.

In comparison with grain filling, which presented lower values of R-Pearson, the descriptive results presented in the heat map (Fig. 3) show that the highest R-Pearson values of the simple regressions between GNC and multispectral and RGB indices were observed at the phenological stage of anthesis. At anthesis and aerial level, the multispectral vegetation indices showed a positive direct relationship with GNC. Regarding the RGB indices at aerial level, positive correlations were observed. However, a* and u* indices showed to be more negative with greener vegetation due to the nature of the colour space conversion. Hence, an inverse relationship was observed in this case. Counterintuitively, this was not observed at RGB ground level.

As shown in Fig. 4, the less negative $\delta^{13}C$ values were found, in general terms, in fields located in the southern and middle areas of Navarre, while the most negative values were mainly found in fields in the North.

In Fig. 5, the obtained GNC estimation accuracy (R^2) of the validated models at all the sensing levels and sensor type is shown. At anthesis, the multispectral sensors at UAV level were the ones with the most precise GNC estimation capacities with a significant R^2 of 0.42 (Fig. 6). At grain filling, the UAV-multispectral sensor effectiveness at estimating GNC was reduced to a R^2 of 0.29. At ground level, the validation of the model built with RGB data to estimate GNC was significant at the phenological stage of anthesis with a R^2 of 0.19. None of the validation of the rest of models (UAV RGB and ground multispectral at anthesis; and ground RGB, UAV RGB and ground multispectral at grain filling) showed significant results.

The most suitable GNC estimation model obtained from multispectral imagery at UAV level at the phenological stage of anthesis is summarized in Table 5. The most optimal veg-

Table 4 – Bands and resolution of Sentinel-2 Multispectral Instrument (MSI).

MSI Band	Spatial Resolution(m)	Central Wavelength(nm)
B1: Coastal Aerosol	60	443
B2: Blue	10	490
B3: Green	10	560
B4: Red	10	665
B5: Red-Edge	20	705
B6: Red-Edge	20	740
B7: Red-Edge	20	783
B8: NIR	10	842
B8A: Vegetation RE	20	865
B9: Water Vapour	60	945
B10: SWIR Cirrus	60	1375
B11: SWIR	20	1610
B12: SWIR	20	2190

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

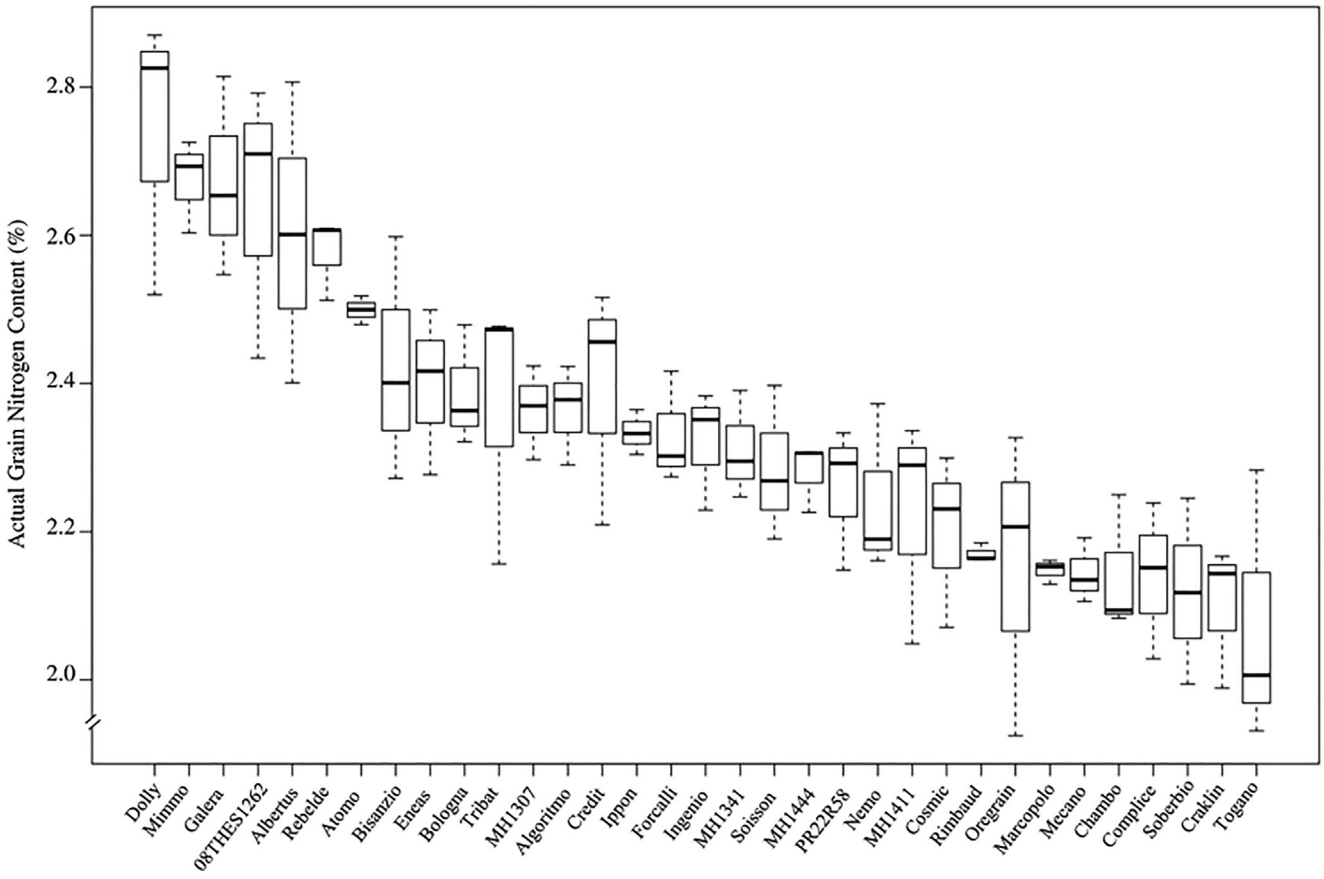


Fig. 2 – Posthoc analyses of the utilized wheat lines at the experimental station of Zamadueñas, the significance letters are shown in the appendix, Table A1.

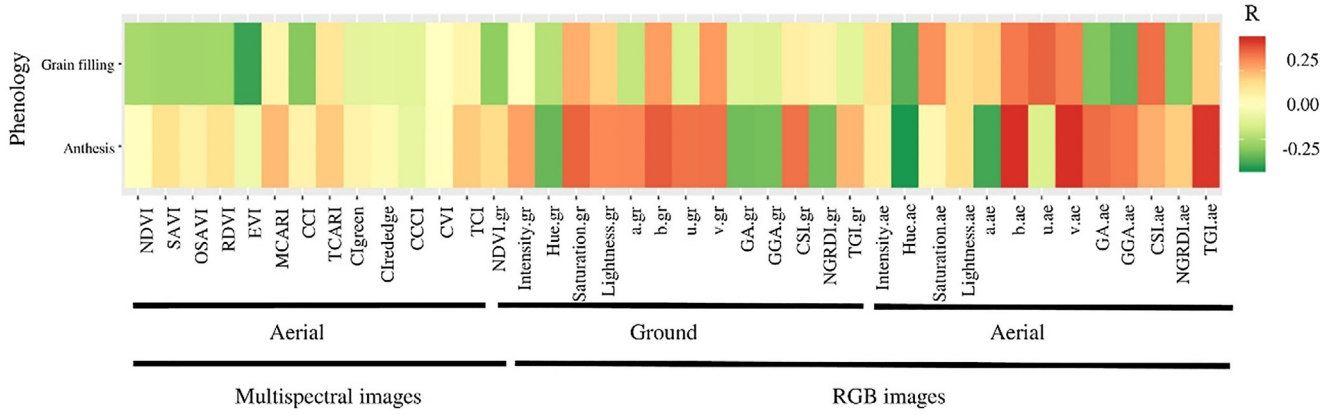


Fig. 3 – Heatmap descriptive representation of R Person correlations between GNC and various multispectral and RGB indices at aerial and ground levels. In the colour scale, red represents the highest positive R value, while green is the lowest negative R value.

etation indices as estimators were TCARI, Cl red edge and EVI. The parameters were significant and had VIF values of 9.0 for TCARI, 9.7 for Cl red edge and 1.6 for EVI. Regarding the goodness of the model the best performing model showed absence of collinearity as all VIF values (9 for TCARI, 9.7 for Cl red edge and 1.6 for EVI) were closer to the collinearity threshold value of 10 [71].

The model built with UAV data that best suited GNC estimation was applied at satellite level with Sentinel-2 imagery. Over the two years the validation of the model with satellite data was significant. The estimation of GNC with satellite data reached an R^2 = 0.40 and RMSE = 0.29% (Fig. 7), alike the one obtained with the UAV multispectral imagery (R^2 = 0.42 and RMSE = 0.18%, Fig. 6).

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

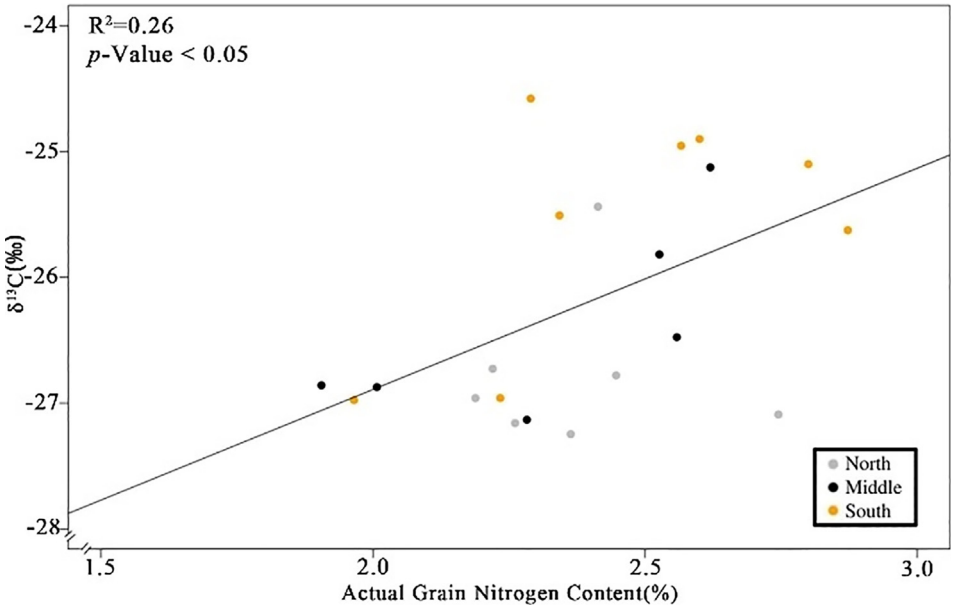


Fig. 4 – Carbon isotopic composition ($\delta^{13}C$, ‰) against grain nitrogen content (%) to assess C isotopic discrimination and characterize hydric conditions across the agroclimates (in North, Middle and South of Navarre, Spain) where the farmers' fields are located.

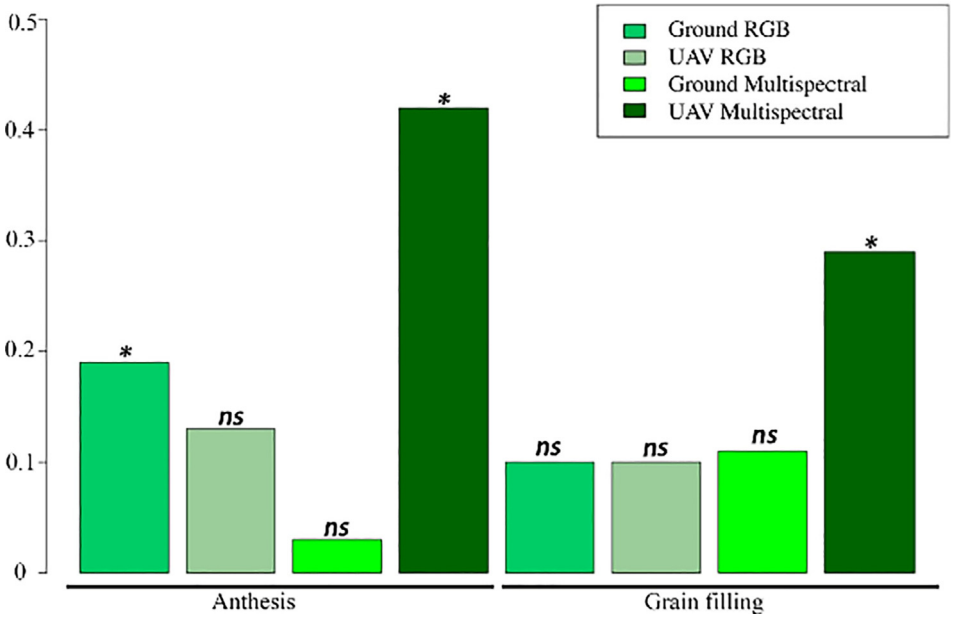


Fig. 5 – Summary of the obtained R^2 at validation for GNC estimation with RGB and multispectral sensors at ground and aerial levels, the * indicates a significance of p-value < 0.05 while ns refers to no significance. The two phenological stages studied are shown (anthesis and grain filling).

4. Discussion

The models developed in this study using UAV multispectral sensors outperformed RGB models (at ground and aerial level) as well as models built with a widely used multispectral ground-level sensor (GreenSeeker). Coming back to the questions one (which imagery, RGB vs Multispectral, can contribute the most to estimate wheat GNC?) and two (which sensing level, UAV vs ground images, is more effective to generate models for GNC monitoring?) the following was observed. When validating the GNC estimation models with multispectral data at UAV level the best results showed (Fig. 5) an R^2 of 0.42 and RMSE of 0.18% at anthesis, which aligns to the range reported by Zhao et al. [38]. No significant correlations were observed with models built with RGB data at ground and aerial level, expect for the low yielding (R^2 = 0.19, RMSE = 0.19%) of the ground RGB model at anthesis (Fig. 5). High spectral resolution in the visible photosyntheti-

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

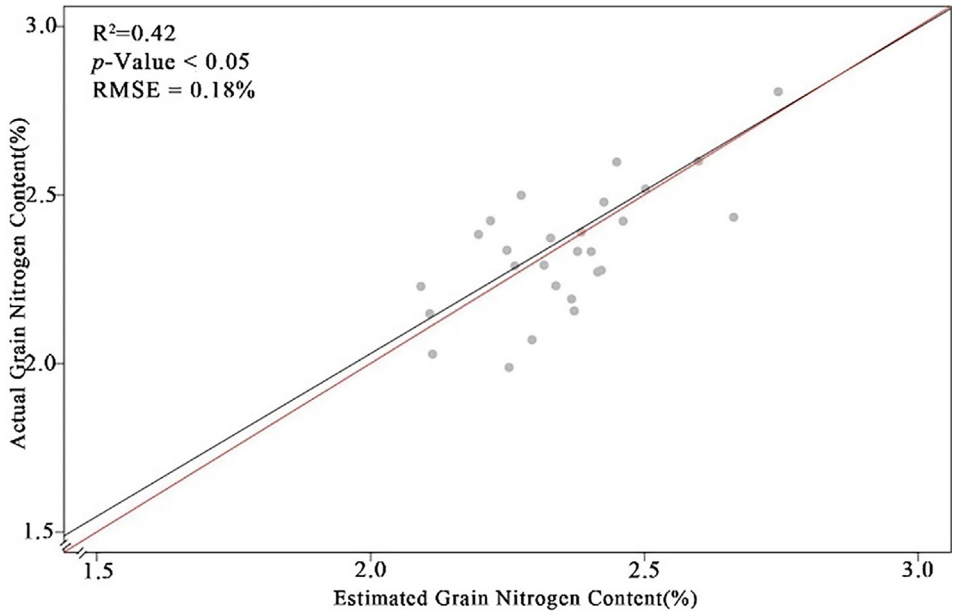


Fig. 6 – Validation of the stepwise multilinear regression models for UAV multilinear data at anthesis as the most suitable model for GNC estimation, 30% of the data set of the experimental plots were used for validation. The equation of the model is shown in Table 5. The R^2, *p*-value and root mean square (RMSE) are shown. The red line indicates 1:1.

Table 5 – Summary of stepwise multilinear results for the UAV multispectral model at anthesis, it shows the equation, the selected vegetation indices and the VIF.

Parameter	Estimate	*p*	VIF
Intercept	−1.09	<0.01	
TCARI	31.50	<0.01	9.0
$Cl_{rededge}$	0.60	<0.01	9.7
EVI	−1.74	<0.01	1.6
Model Summary			
Grain N content = $-1.09 + 31.5 \times TCARI + 0.60 \times Cl_{rededge} - 1.74 \times EVI$			

cally active wavelength regions (RGB) has been shown to be capable of assessing crop nitrogen content [72], nonetheless we observed that it is not assessable when using broad band reflectance measurements provided by commercial RGB cameras, regardless of their high spatial resolution and reliable colour calibration [73]. Hence, we can argue that in this study the model built with multispectral data at UAV level contributed the most to GNC estimation. The multispectral instrument used at the ground level only measures red and near-infrared reflectance (computed as NDVI) and showed no significant correlations. This might be related, first to the lack of relationship between a biomass sensitive index such as NDVI and GNC, which is related to N content in the plant. Moreover, the saturation of NDVI, which is reached in high vegetation density covers such as anthesis, does not contribute to generate an indicator of the plant that could be likely linked to GNC. In addition, Greenseeker has spatial resolution limitations in comparison with UAV sensors (Table 2). These results demonstrate the superior capacities of multiple bands and multispectral information itself, as well as spatial resolution, for the case of GNC estimation.

Due to the nature of the data, empirical models were chosen in contrast with other approaches. Radiative transfer models, for instance, offer an improved understanding of the structure and physiological features of the crop canopy although are very dependent on large datasets and computationally demanding algorithms. Empirical models, such as VIs regressions, sometimes experience limitations, as are more affected by site and sensor conditions, but are more versatile when having reduced datasets and widely used and understood. In this case, 14 VIs were calculated at UAV level using Tetracam multispectral data while only one multispectral index was calculated at ground level (NDVI). Although the Greenseeker is a low-cost and widely used ground multispectral instrument for phenotyping [46,74,75], the spectral resolution it features is insufficient for GNC estimation. On the other hand, this contrasts with the generally more costly and logistically demanding (i.e. a minimum 5 m long-poles or UAV) multispectral cameras. Nonetheless, in the last years multispectral cameras prices have decreased and their use in phenotyping and research-oriented proposes is increasing. Hence, nowadays multispectral cameras mounted on UAVs are common, as the reviewed scientific literature from the last years shows [76], and may further contribute to phenotyping quality related traits.

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

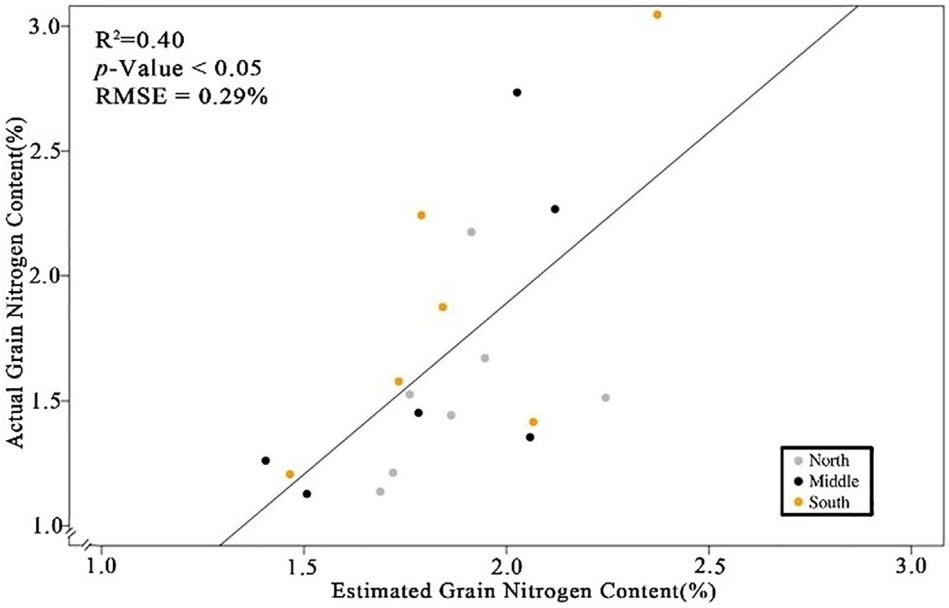


Fig. 7 – Application of the most suitable stepwise multilinear model with Sentinel-2 data at actual farmers' fields in Navarre, Spain, across three agroclimatic regions (North, Middle and South). The R^2, *p*-value and root mean square (RMSE) are shown.

The UAV remotely sensed multispectral data in this study has several advantages over the RGB data. We observed that among the three selected vegetation indices (TCARI, Cl red-edge and EVI) as the most suitable for the best performing stepwise multilinear model (Table 5), EVI uses RGB and NIR bands, while TCARI and $Cl_{red\text{-}edge}$, besides those, use a band in the red-edge part of the spectrum (700 ± 10 nm). The advantages of indices using red-edge bands for determining nitrogen up-taken into grains have been observed in various studies [28,77,78], as the study here presented also corroborates. EVI is an index sensitive to biomass, which has nonetheless been successfully used in plant nitrogen concentration estimations [79], also in the estimation of grain protein content in wheat [32]. EVI has improved features to prevent atmospheric and soil influences, which make it more responsive to canopy variations in comparison with other indices [80]. In this sense, the results suggest that the amount of spectral information available from multi-band multispectral sensors increases GNC estimation effectiveness in comparison with the RGB visible three-band data, red–greenblue, as it expands the sensing spectrum further to red (>670 nm). The red-edge part of the spectrum is highly sensitive to chlorophyll changes and therefore can be used as a nitrogen proxy [20,81]. TCARI has been successfully used at estimating grain protein content in cereals [82], as well as $Cl_{red\text{-}edge}$ at UAV level in the case of wheat with a reported R^2 of 0.50 [30]. We can argue that the spectral resolution (bands in the red-edge) is especially relevant when the sensing aim, nitrogen in grains, is indirectly discernible and highly dependent on chlorophyll-related spectral information.

Regarding the third question (how relevant is phenology for GNC estimation?) we observed that phenology was central for GNC estimation (Figs. 3 and 5). At anthesis, the UAV multispectral GNC estimation model reached an R^2 of 0.42, while at grain filling it decreased to 0.29. This aligns with the conceptualization that wheat nitrogen accumulation in grains start happening at grain filling. Namely, plants nitrogen concentration starts descending at the grain filling stage. Thus, the results here presented suggest that VIs sensitive to nitrogen concentration at anthesis can be an indicator for the final GNC in wheat, as was also observed by Dupont and Altenbach [83]. In wheat, between 60 and 95% of the nitrogen is remobilized from leaves and stems into grains, being the most important sources of GNC [84,85] and thereby of protein in wheat grains. The results of the study show that it is possible to monitor GNC through these vegetation indices sensitive to chlorophyll if plant nitrogen content is sensed around anthesis. This phenological stage corresponds to the maximum nitrogen accumulation [86,87] in wheat plants, before it is remobilized into grains.

At the agro-ecosystem scale, the higher (i.e. less negative) carbon isotope composition (δ13C) values of mature grains in the Middle and Southern parts of Navarre compared with the Northern area (Fig. 4) suggests the occurrence of more drought stress in the central and southern areas of Navarra. This is in agreement with their lower rainfall, higher temperatures, and lower grain yields observed for the fields in this study in Middle and Southern Navarre [69]. In general, grain nitrogen content (GNC) was higher in the Middle and Southern areas than in the Northern area, probably related with a smaller grain size, caused by drought stress during grain filling, in the former areas. In that sense δ13C and GNC across sites and areas were positively correlated (Fig. 4) and served to explain agroclimatic features linked to wheat physiology and GNC.

This study demonstrates that the stepwise multilinear model built at the UAV level could be up-scaled with Sentinel-2 data and cover field-level GNC estimations. Hence,

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

regarding the fourth question (Are models built with ground or UAV images up-scalable to equivalent Sentinel-2 images?), we observed that the applied model over two seasons reached an R^2 coefficient of 0.40 and RMSE of 0.29% (Fig. 7), which is in the same range as the one obtained with the validation at the experimental plots with the UAV multispectral data (R^2 of 0.42 and RMSE of 0.18%, Fig. 6). The RMSE is slightly higher when the model is applied at the field-level; nonetheless, these results align with other studies that concomitantly used UAV obtained spectral data to build models and apply them with Sentinel-2 data [88,89], as well for the case of GNC estimation [38] where several N sensitive stepwise models showed a RMSE range of 0.19–0.53%. Building models with actual field-level to regional data requires vast field work, covering large spatial extensions. On the other hand, building the models at smaller experimental sites eases the process. The study here discussed confirms that the simultaneous use of UAV multispectral cameras and Sentinel-2 data with equivalent spectral bands allows GNC estimation.

This study has been able to define the most optimal approaches for estimating GNC with applications at both experimental plots and agricultural fields. The novel comparison presented here among different sensing platforms and phenological stages, sets the basis for an improved quality-oriented phenotyping and crop monitoring. Often, quality traits have been under-considered in breeding programs [9,10] and the development of models able to separate high GNC phenotypes from others could help identify the most suitable lines in wheat populations exhibiting genetic variability. In this sense, the genotypic range used for this study was 2.07% to 2.74% of GNC, it could be broader (1.2–3.7%) when adding a bigger diversity of genotypes, from modern varieties to landraces. Therefore, the model is specific for a range of GNC in the instance of our study. Despite the achieved R2 is inferior to 0.5, the significance of the models and the results at validation proves its effectiveness in phenotyping at experimental plots and monitoring GNC at actual agricultural fields, while selecting the subset of genotypes with the desired range values of N content. Currently, these models cannot determine GNC with potential applications in precise certification of grain protein content before harvest. For this aim destructive measurements would still be needed. Nonetheless, the obtained models can orient farmers on grain quality after a quick UAV flight before harvest and could also potentially be used in smart farming to spot low and high GNC concentration areas in a field. This information could contribute to optimize the use of resources by prioritizing increased top-dressing fertilization in poor GNC zones, and thus helping to advance to agricultural stability in terms of GNC. In addition, at the agroecosystem level with the use of Sentinel-2 imagery, the current reliability of the models might provide support for developing regional GNC estimation maps which can contribute to understand environmental and soil features affecting GNC, as well as to help farmers and institutions to manage croplands sustainably. The Sentinel-2 spatial resolution makes their direct application as a phenotyping platform in breeding programs unsuitable, nonetheless for the application in crop phenotyping the model developed in this study could be implemented with satellite platforms with higher resolutions.

5. Conclusion

This study demonstrates that UAV multispectral models can provide accurate (R^2 = 0.42 and RMSE = 0.18%) GNC estimations across phenotypes when the sensing period happens at anthesis. Moreover, it corroborates the central use of red-edge bands (700 ± 10 nm) to calculate effective VIs for GNC estimation. The application of UAV-built multispectral models with equivalent Sentinel-2 real data effectively estimated GNC over two seasons at field-level (R^2 = 0.40 and RMSE = 0.29%). RGB true colour images and the widely used ground multispectral phenotyping sensor Greenseeker were not effective in GNC estimation. The models here developed show the potentialities of GNC estimation across phenotypes with UAV multispectral data as it allows to differentiate groups of genotypes regarding their GNC. Moreover, we also demonstrated the potentialities of the application of these models to accessible satellite data (Sentinel-2) for an improved GNC monitoring at field-scale. It takes advantage to the novel Sentinel-2 features with its red-edge bands and high revisit time. To our knowledge, only one other study using UAV-level data (hyperspectral in this case) to build models and monitor GNC at field-level with equivalent Sentinel-2 data has been published [38]. The study here presented shows two main novelties. On the one hand, it demonstrates that multispectral UAV-mounted cameras can produce effective GNC estimation models when having equivalent bands in the red-edge though additional improvements in this field may be reached by including hyperspectral information in GNC estimation in the future. On the other hand, it is the first paper that together reviews remote sensing techniques at various levels to determine the most optimal strategy for both GNC phenotyping and field-level monitoring.

Declaration of Competing Interest

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

Acknowledgements

This study was supported by the projects PID2019-106650RB-C21 (Ministerio de Ciencia e Innovación, MICINN, Spain) and 0011-1365-2018-000213/0011-1365-2018-000150 (Government of Navarre, Spain). J.S. is recipient of a FPI doctoral fellowship (Grant: PRE2020-091907) from MICINN, Spain. J.L.A. acknowledges support from ICREA Academia, Generalitat de Catalunya, Spain. S.C.K. is supported by the Ramon y Cajal RYC-2019-027818-I research fellowship from MICINN, Spain. The processing of satellite images was supported by the European Cooperation in Science and Technology Action CA17134 SENSECO.

Appendix A. Wheat lines differences in GNC

See Table A1.

Table A1 – Average GNC in % for each genotype at the research site of Zamadueñas. The superscript letter indicates differences among the line after post hoc test.

Genotypes	N % in grain	Genotypes	N $_o$ in grain
Dolly	2.74^{a}	MH1341	$2.31^{efghijk}$
Mimmo	2.74^{ab}	Soisson	$2.28^{fghijkl}$
Galera	2.67^{ab}	MH1444	$2.26^{fghijklm}$
08THES1262	2.64^{ab}	PR22R58	$2.24^{fghijklm}$
Albertus	2.60^{abc}	Nemo	$2.25^{fghijklm}$
Rebelde	2.56^{abcd}	MH1411	$2.23^{ghijklm}$
Atomo	2.50^{bcde}	Cosmic	$2.20^{ghijklm}$
Bisanzio	2.42^{cdef}	Rimbaud	2.17^{hijklm}
Eneas	2.40^{defg}	Oregrain	2.15^{ijklm}
Bologna	2.39^{defg}	Marcopolo	2.15^{ijklm}
Tribat	$2.37^{defg\ j}$	Mecano	2.14^{ijklm}
MH1307	2.36^{defgh}	Chambo	2.14^{ijklm}
Algoritmo	2.33^{defgh}	Complice	2.14^{jklm}
Credit	2.33^{defgh}	Soberbio	2.12^{klm}
Ippon	2.33^{efghij}	Craklin	2.10^{lm}
Forcalli	2.33^{efghij}	Togano	2.07^{m}
Ingenio	2.32^{efghij}		

Appendix B. Results and graphs of the calculated models.

See Tables B1 and B2 and Figs. B1–B7.

Table B1 – Summary of stepwise multilinear results for the ground RGB models at anthesis It shows the equation, the selected vegetation indices and the VIF.

Parameter	Estimate	*p*	VIF
Intercept	3.38	<0.01	
Hue	−0.01	<0.01	5.1
v	0.01	<0.01	1.3
GA	−0.6	<0.01	5.2
Model Summary			
Grain N content = 3.38–0.01 × Hue + 0.01 × v − 0.60 × GA			

Table B2 – Summary of stepwise multilinear results for the UAV multispectral models at grain filling. It shows the equation, the selected vegetation indices and the VIF.

Parameter	Estimate	*p*	VIF
Intercept	3.3	<0.01	
OSAVI	–6.1	ns	69.0
RDVI	8.1	<0.01	80.8
EVI	–1.3	<0.01	2.8
Model Summary			
Grain N content = 3.3–6.1 × OSAVI + 8.1 × RDVI − 1.3 × EVI			

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

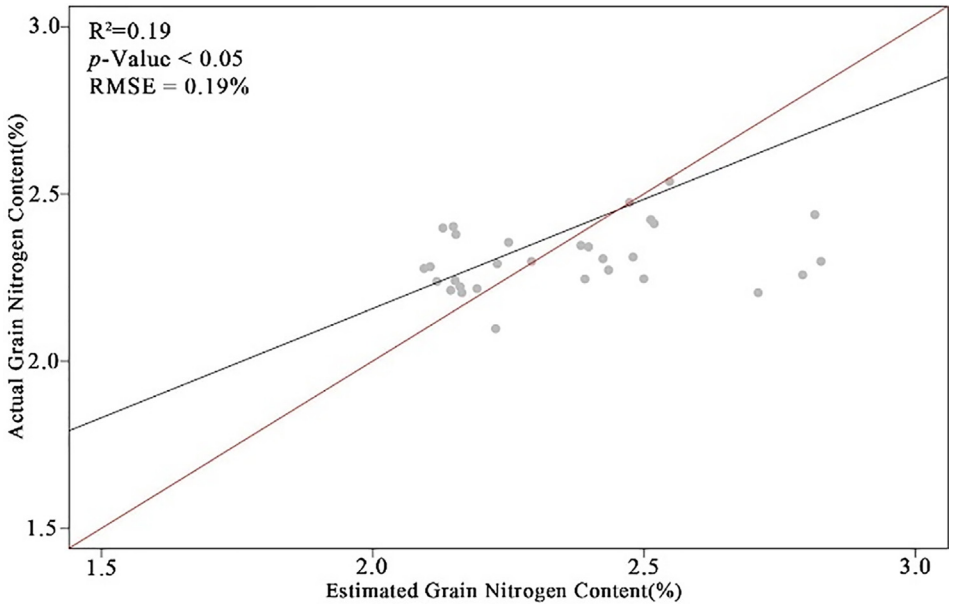


Fig. B1 – Validation of the stepwise multilinear regression models for ground RGB data at anthesis, 30% of the data set of the experimental plots were used for validation. The equation of the model is shown in Table B.1. The R^2, *p*-value and root mean square (RMSE) are shown. The red line indicates 1:1.

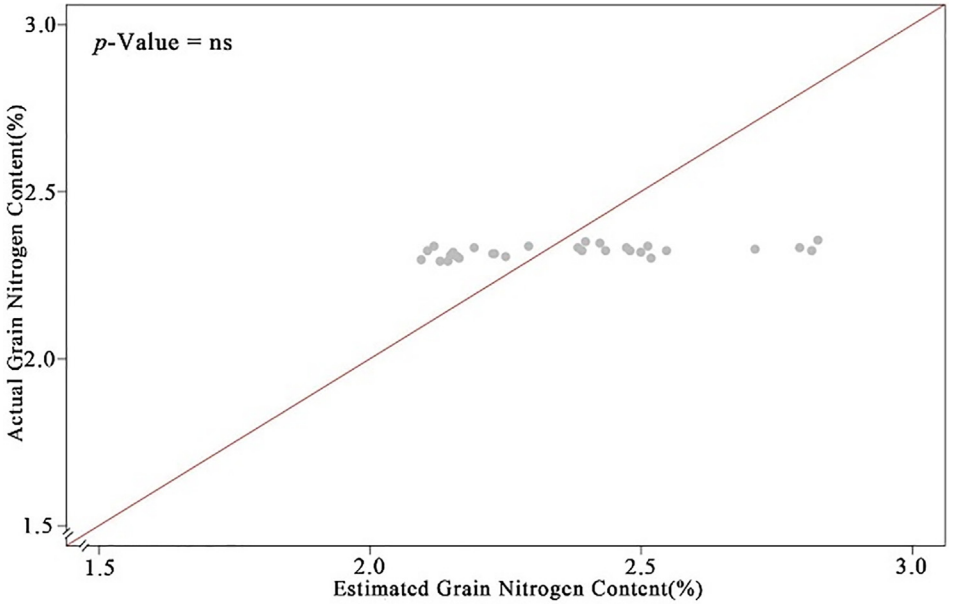


Fig. B2 – Validation of the stepwise multilinear regression models for ground multispectral data at anthesis, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

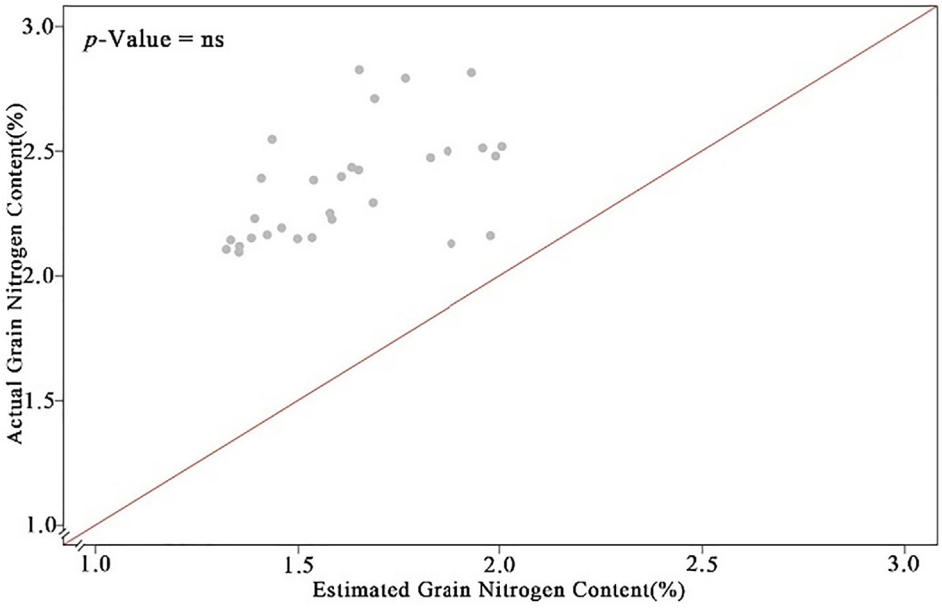


Fig. B3 – Validation of the stepwise multilinear regression models for UAV RGB data at anthesis, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.

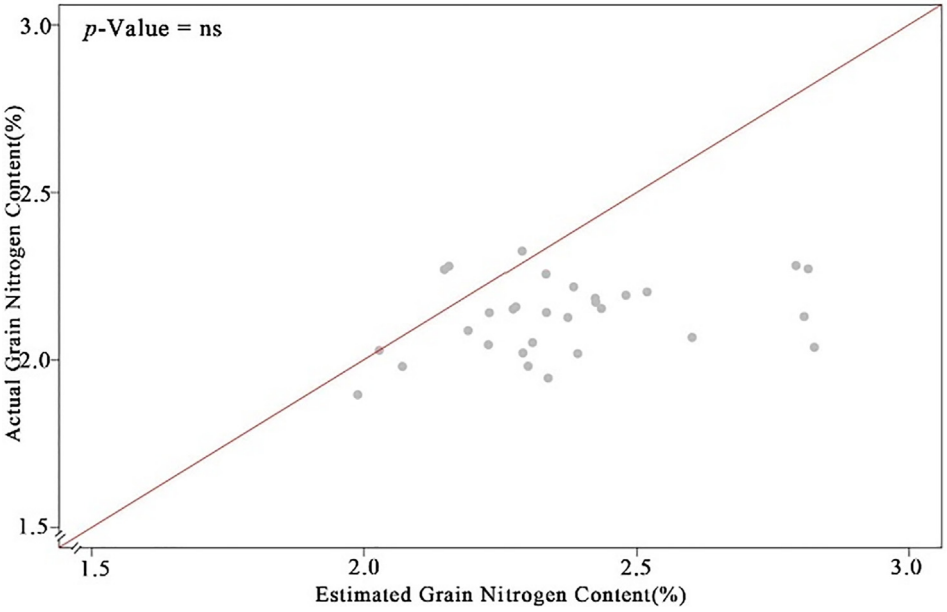


Fig. B4 – Validation of the stepwise multilinear regression models for ground RGB data at grain filling, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

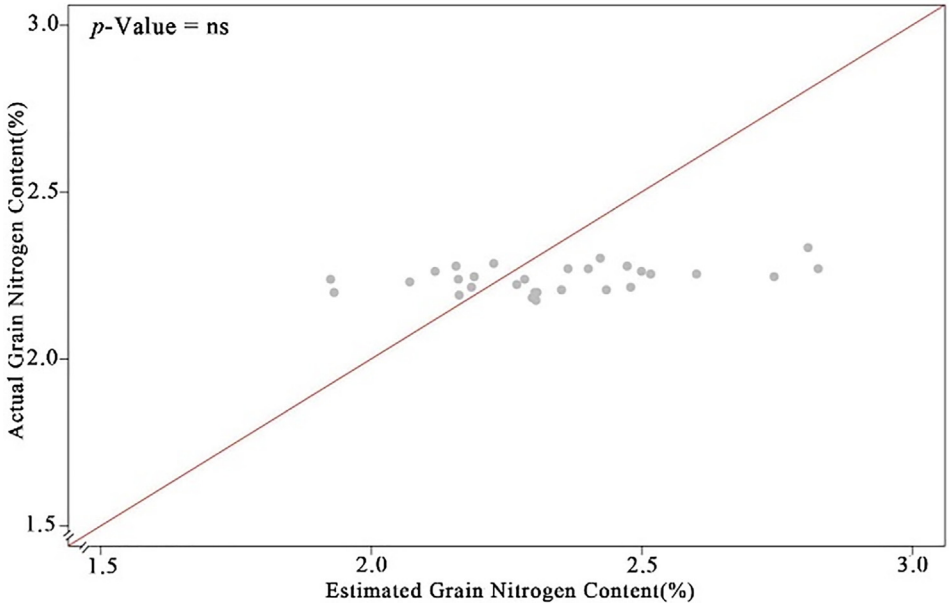


Fig. B5 – Validation of the stepwise multilinear regression models for ground multispectral data at grain filling, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.

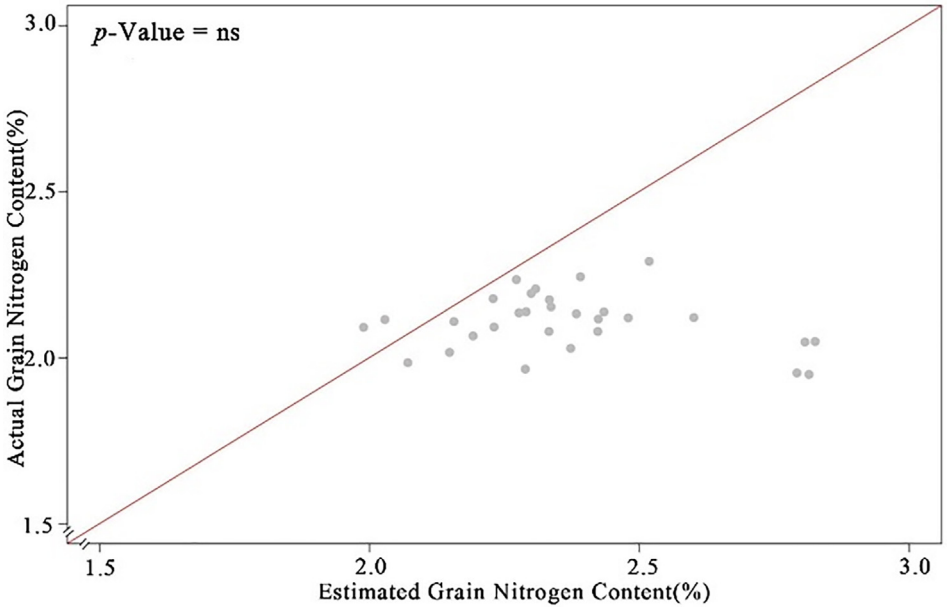


Fig. B6 – Validation of the stepwise multilinear regression models for UAV RGB data at grain filling, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

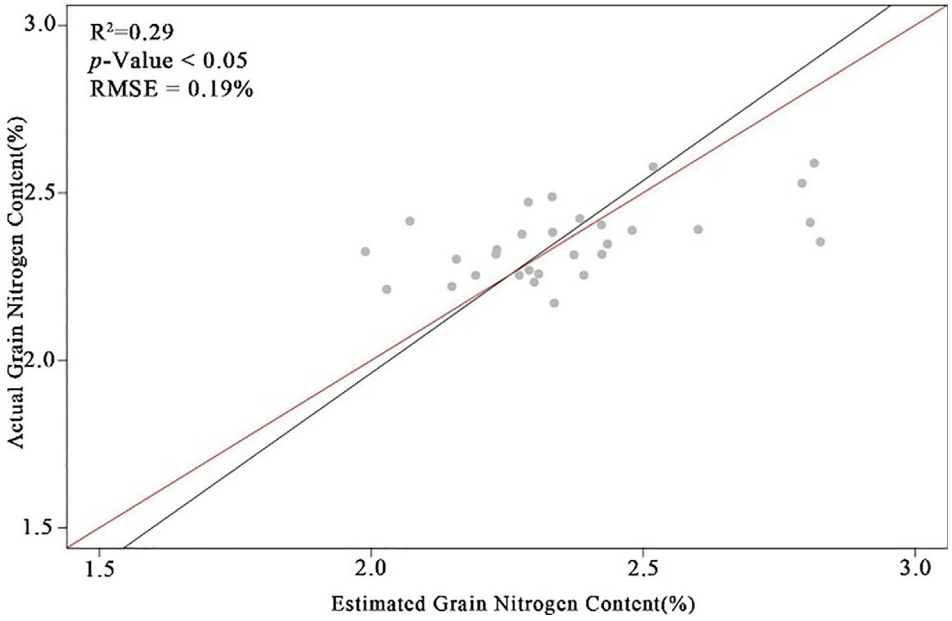


Fig. B7 – Validation of the stepwise multilinear regression models for UAV multispectral data at grain filling, 30% of the data set of the experimental plots were used for validation. The equation of the model is shown in Table B.1. The R^2, *p*-value and root mean square (RMSE) are shown. The red line indicates 1:1.

Appendix C. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.inpa.2022.05.004.

REFERENCES

[1] Assembly G. Transforming our world: the 2030 Agenda for Sustainable Development. New York, United States: United Nations; 2015.
[2] Awika JM. Major cereal grains production and use around the world. In: Awika JM, Piironen V, Bean S, editors. Advances in Cereal Science: Implications to Food Processing and Health Promotion. USA: ACS Symp Ser; 2011. p. 1–13.
[3] Peña RJ. Wheat for Bread and Other Foods. Rome, Italy: Food and Agriculture Organization of the United Nations; 2002.
[4] Esquinas-Alcázar J. Protecting crop genetic diversity for food security: Political, ethical and technical challenges. Nat Rev Genet 2005;6(12):946–53.
[5] Peña RJ. Current and future trends of wheat quality needs. In: Buck HT, Nisi JE, Salomón N, editors. Developments in plant breedingwheat production in stressed environments. Dordrecht: Springer Netherlands; 2007. p. 411–24.
[6] Dinu M, Whittaker A, Pagliai G, Benedettelli S, Sofi F. Ancient wheat species and human health: Biochemical and clinical implications. J Nutr Biochem 2018;52:1–9.
[7] Hazard B, Trafford K, Lovegrove A, Griffiths S, Uauy C, Shewry P. Strategies to improve wheat for human health. Nat Food 2020;1(8):475–80.
[8] Regina A, Guzman C. Starch and starch-associated proteins: impact on wheat grain quality. In: Igrejas G, Ikeda TM, Guzmán C, editors. Wheat quality for improving processing and human health. Berlin, Heidelberg, Germany: Springer; 2020. p. 39–73.
[9] Zhu J, Huang S, Khan K, Brien L. Relationship of protein quantity, quality and dough properties with Chinese steamed bread quality. J Cereal Sci 2001;33(2):205–12.
[10] Pronin D, Börner A, Weber H, Wheat SKA. (Triticum aestivum L.) breeding from 1891 to 2010 contributed to increasing yield and glutenin contents but decreasing protein and gliadin contents. J Agric Food Chem 2020;68(46):13247–56.
[11] Sanchez-Garcia M, Álvaro F, Peremarti A, Martín-Sánchez JA, Royo C. Changes in bread-making quality attributes of bread wheat varieties cultivated in Spain during the 20th century. Eur J Agron 2015;63:79–88.
[12] Vogel KP, Johnson VA, Mattern PJ. Protein and Lysine Content of Grain, Endosperm, and Bran of Wheats from the USDA World Wheat Collection. Crop Sci 1976;16(5):655–60.
[13] Acreche MM, Slafer GA. Variation of grain nitrogen content in relation with grain yield in old and modern Spanish wheats grown under a wide range of agronomic conditions in a Mediterranean region. J Agric Sci 2009;147(6):657–67.
[14] Lloveras J, Lopez A, Ferran J, Espachs S, Solsona J. Bread-Making Wheat and Soil Nitrate as Affected by Nitrogen Fertilization in Irrigated Mediterranean Conditions. Agron J 2001;93(6):1183–90.
[15] Daniel C, Triboí E. Changes in wheat protein aggregation during grain development: Effects of temperatures and water stress. Eur J Agron 2002;16(1):1–12.
[16] Xue LH, Cao WX, Yang LZ. Predicting Grain Yield and Protein Content in Winter Wheat at Different N Supply Levels Using Canopy Reflectance Spectra. Pedosphere 2007;17 (5):646–53.
[17] Serrano L, Peñuelas J, Ustin SL. Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data. Remote Sens Environ 2002;81(2):355–64.
[18] Kokaly RF. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. Remote Sens Environ 2001;75(2):153–61.
[19] Fourty T, Baret F, Jacquemoud S, Schmuck G, Verdebout J. Leaf optical properties with explicit description of its biochemical

composition: Direct and inverse problems. Remote Sens Environ 1996;56(2):104–17.
[20] Guerif M, Houles V, Baret F. Remote sensing and detection of nitrogen status in crops. Application to precise nitrogen fertilization. In: 4th int symp intell inf technol agric ISIITA. Beijin, China; 2007. p. 19–28.
[21] Féret JB, Berger K, de Boissieu F, Malenovský Z. PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents. Remote Sens Environ 2021;252:112173–88.
[22] Ibrahim A, Csúr-Varga A, Jolánkai M, Mansour H, Hamed A. Monitoring some quality attributes of different wheat varieties by infrared technology. Agric Eng Int CIGR J 2018;20 (1):201–10.
[23] Cheng A. Review: Shaping a sustainable food future by rediscovering long-forgotten ancient grains. Plant Sci 2018;269:136–42.
[24] Arzani A, Ashraf M. Cultivated Ancient Wheats (Triticum spp.): A Potential Source of Health-Beneficial Food Products. Compr Rev Food Sci Food Saf 2017;16(3):477–88.
[25] Moges SM, Raun WR, Mullen RW, Freeman KW, Johnson GV, Solie JB. Evaluation of green, red, and near infrared bands for predicting winter wheat biomass, nitrogen uptake, and final grain yield. J Plant Nutr 2004;27(8):1431–41.
[26] Wright DL, Rasmussen VP, Ramsey RD, Baker DJ, Ellsworth JW. Canopy reflectance estimation of wheat nitrogen content for grain protein management. GIScience Remote Sens 2004;41(4):287–300.
[27] Saleem MF, Ma BL, Voldeng H, Wang TC. Nitrogen nutrition on leaf chlorophyll, canopy reflectance, grain protein and grain yield of wheat varieties with contrasting grain protein concentration. J Plant Nutr 2010;33(11):1681–95.
[28] Prey L, Schmidhalter U. Temporal and spectral optimization of vegetation indices for estimating grain nitrogen uptake and late-seasonal nitrogen traits in wheat. Sensors 2019;19 (21):1–27.
[29] Geipel J, Link J, Wirwahn JA, Claupein W. A programmable aerial multispectral camera system for in-season crop biomass and nitrogen content estimation. Agric 2016;6(1):1–19.
[30] Zhou X, Kono Y, Win A, Matsui T, Tanaka TST. Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. Plant Prod Sci 2020;24 (2):137–51.
[31] Quemada M, Pancorbo JL, Alonso-Ayuso M, Gabriel JL., López-Herrera J, Pérez-Martín E. Vegetation indices from remote sensing imagery as proxies for yield and grain N in wheat. In: Precision Agriculture'19 Wageningen Academic Publishers; Montpelier, France; 2019. p. 323–30.
[32] Tan C, Zhou X, Zhang P, Wang Z, Wang D, Guo W, et al. Predicting grain protein content of field-grown winter wheat with satellite images and partial least square algorithm. PLoS ONE 2020;15(3):e0228500.
[33] Zhao C, Liu L, Wang J, Huang W, Song X, Li C. Predicting grain protein content of winter wheat using remote sensing data based on nitrogen status and water stress. Int J Appl Earth Obs Geoinf 2005;7(1):1–9.
[34] Li CJ, Wang JH, Wang Q, Wang DC, Song XY, Wang Y, et al. Estimating Wheat Grain Protein Content Using Multi-Temporal Remote Sensing Data Based on Partial Least Squares Regression. J Integr Agric 2012;11(9):1445–52.
[35] Song X, Wang J, Yang G, Feng H. Winter Wheat Cropland Grain Protein Content Evaluation through Remote Sensing. Intell Autom Soft Comput 2014;20(4):599–609.
[36] Wang L, Tian Y, Yao X, Zhu Y, Cao W. Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. F Crop Res 2014;164:178–88.
[37] Segarra J, Buchaillot ML, Araus JL, Kefauver SC. Remote Sensing for Precision Agriculture : Sentinel-2 Improved Features and Applications. Agronomy 2020;10(5):1–18.
[38] Zhao H, Song X, Yang G, Li Z, Zhang D, Feng H. Monitoring of nitrogen and grain protein content in winter wheat based on Sentinel-2A data. Remote Sens 2019;11(14):1–25.
[39] Reyniers M, Vrindts E. Measuring wheat nitrogen status from space and ground-based platform. Int J Remote Sens 2006;27 (3):549–67.
[40] Du M, Noguchi N, Ito A, Shibuya Y. Correlation analysis of vegetation indices based on multi-temporal satellite images and unmanned aerial vehicle images with wheat protein contents. Eng Agric Environ Food 2019;10:1–12.
[41] Farquhar GD, Ehleringer JR, Hubick KT. Carbon isotope discrimination and photosyntehsis. Annu Rev Plant Biol 1989;40(1):503–37.
[42] Rezzouk FZ, Gracia-Romero A, Kefauver SC, Gutiérrez NA, Aranjuelo I, Serret MD, et al. Remote sensing techniques and stable isotopes as phenotyping tools to assess wheat yield performance: Effects of growing temperature and vernalization. Plant Sci 2020;295:110281–97.
[43] Crain J, Ortiz-Monasterio I, Raun B. Evaluation of a reduced cost active NDVI sensor for crop nutrient management. J Sensors 2012;2012:1–10.
[44] Barmeier G, Schmidhalter U. High-throughput phenotyping of wheat and barley plants grown in single or few rows in small plots using active and passive spectral proximal sensing. Sensors 2016;16(11):1–14.
[45] Bendig J, Bolten A, Bennertz S, Broscheit J, Eichfuss S, Bareth G. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sens 2014;6(11):10395–412.
[46] Gracia-Romero A, Kefauver SC, Fernandez-Gallego JA, Vergara-Díaz O, Nieto-Taladriz MT, Araus JL. UAV and ground image-based phenotyping: A proof of concept with durum wheat. Remote Sens 2019;11(10):1244–69.
[47] Casadesús J, Kaya Y, Bort J, Nachit MM, Araus JL, Amor S, et al. Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. Ann Appl Biol 2007;150(2):227–36.
[48] Gracia-Romero A, Kefauver SC, Vergara-Díaz O, Zaman-Allah MA, Prasanna BM, Cairns JE, et al. Comparative performance of ground vs. Aerially assessed rgb and multispectral indices for early-growth evaluation of maize performance under phosphorus fertilization. Front Plant Sci 2017;8:1–13.
[49] Hunt ER, Doraiswamy PC, McMurtrey JE, Daughtry CST, Perry EM, Akhmedov B. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. Int J Appl Earth Obs Geoinf 2013;21:103–12.
[50] Hunt ER, Cavigelli M, Daughtry CST, Mcmurtrey JE, Walthall CL. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. Precis Agric 2005;6(4):359–78.
[51] Zaman-Allah M, Vergara O, Araus JL, Tarekegne A, Magorokosho C, Zarco-Tejada PJ, et al. Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. Plant Methods 2015;11(1).
[52] Vergara-Diaz O, Kefauver SC, Elazab A, Nieto-Taladriz MT, Araus JL. Grain yield losses in yellow-rusted durum wheat estimated using digital and conventional parameters under field conditions. Crop J 2015;3(3):200–10.
[53] Barnes EM, Clarke TR, Richards SE, Colaizzi PD, Haberland J, Kostrzewski M, et al. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In: Proc fifth int conf precis agric Bloomington, USA; 2000. p. 1–15.
[54] Gamon JA, Huemmrich KF, Wong CYS, Ensminger I, Garrity S, Hollinger DY, et al. A remotely sensed pigment index reveals

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004

photosynthetic phenology in evergreen conifers. PNAS 2016;113(46):13087–92.
[55] Gitelson AA, Viña A, Arkebauer TJ, Rundquist DC, Keydan G, Leavitt B. Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophys Res Lett 2003;30 (5):1248–54.
[56] Vincini M, Frazzi E, D'Alessio P. A broad-band leaf chlorophyll vegetation index at the canopy scale. Precis Agric 2008;9 (5):303–19.
[57] Huete A, Didan K, Miura T, Rodriguez EP, Gao X, Ferreira LG. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens Environ 2002;83 (1–2):195–213.
[58] Haboudane D, Miller JR, Tremblay N, Zarco-Tejada PJ, Dextraze L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens Environ 2002;81(2–3):416–26.
[59] Rouse Lr. JW, Haas R, Schell J, Deering D. Monitoring vegetation systems in the great plains with erts. In: Third earth res tech satellite-1 symposium NASA. Washington D.C., USA; 1974. p. 309–17.
[60] Baret F, Jacquemoud S, Hanocq JF. The soil line concept in remote sensing. Remote Sens Rev 1993;7(1):65–82.
[61] Roujean JL, Breon FM. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sens Environ 1995;51(3):375–84.
[62] Huete A. A soil-adjusted vegetation index (SAVI). Remote Sens Environ 1988;25(3):295–309.
[63] Hunt Jr E, Daughtry CS, Eitel JU, Long DS. Remote sensing leaf chlorophyll content using a visible band index. Agron J 2011;103(4):1090–9.
[64] R Core Team. A language and environment for statistical computing: https://www.r-project.org/index.html. 2020.
[65] Harrell FE. Regression modeling strategies. USA: Springer; 2001. p. 78–9.
[66] Meloun M, Militký J, Hill M, Brereton RG. Crucial problems in regression modelling and their solutions. Analyst 2002;127 (4):433–50.
[67] Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography (Cop) 2013;36(1):27–46.
[68] Arnold C. The determination and significance of the base temperature in a linear heat unit system. Proc Am Soc Hortic Sci 1959;74(1):430–45.
[69] Segarra J, González-Torralba J, Aranjuelo I, Araus JL, Kefauver SC. Estimating wheat grain yield using Sentinel-2 imagery and exploring topographic features and rainfall effects on wheat performance in Navarre. Spain Remote Sens 2020;12 (14):1–24.
[70] Goñi J, Caballero A Nuevas variedades de cereal. Navarra Agrar 2017;242:11–21.
[71] Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate Data Analysis. USA: Macmillan; 1995. p. 133–43.
[72] Hansen PM, Schjoerring JK. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sens Environ 2003;86 (4):542–53.
[73] Buchaillot ML, Gracia-Romero A, Vergara-Diaz O, Zaman-Allah MA, Tarekegne A, Cairns JE, et al. Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques. Sensors 2019;19(8):1815.
[74] Gracia-Romero A, Vergara-Díaz O, Thierfelder C, Cairns JE, Kefauver SC, Araus JL. Phenotyping conservation agriculture management effects on ground and aerial remote sensing assessments of maize hybrids performance in Zimbabwe. Remote Sens 2018;10(2):1–21.
[75] Ali AM, Ibrahim SM, Bijay-Singh. Wheat grain yield and nitrogen uptake prediction using atLeaf and GreenSeeker portable optical sensors at jointing growth stage. Inf Process Agric 2020;7(3):375–83.
[76] Deng L, Mao Z, Li X, Hu Z, Duan F, Yan Y. UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. ISPRS J Photogramm Remote Sens 2018;146:124–36.
[77] Guo B-B, Qi S-L, Heng Y-R, Duan J-Z, Zhang H-Y, Wu Y-P, et al. Remotely assessing leaf N uptake in winter wheat based on canopy hyperspectral red-edge absorption. Eur J Agron 2017;82:113–24.
[78] Prey L, Schmidhalter U. Sensitivity of vegetation indices for estimating vegetative N status in winter wheat. Sensors 2019;19(17):1–17.
[79] Zheng H, Li W, Jiang J, Liu Y, Cheng T, Tian Y, et al. A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. Remote Sens 2018;10(12):2026.
[80] Huete AR, Didan K, Van Leeuwen W. MODIS Vegetation Index (MOD13) algorithm theoretical basis document. USA: University of Arizona; 1999.
[81] Segarra J, Buchaillot ML, Stefani U, Araus JL, Kefauver SC. Sentinel-2 responsiveness to fertilization gradients in wheat at field level in Córdoba Province, Argentina. In: Mediterr middle-east geosci remote sens symp M2GARSS 2020. Tunis, Tunisia; 2020. p. 322–5.
[82] Pettersson CG, Eckersten H. Prediction of grain protein in spring malting barley grown in northern Europe. Eur J Agron 2007;27(2–4):205–14.
[83] Dupont FM, Altenbach SB. Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. J Cereal Sci 2003;38 (2):133–46.
[84] Van Sanford DA, MacKown CT. Variation in nitrogen use efficiency among soft red winter wheat genotypes. Theor Appl Genet 1986;72(2):158–63.
[85] Palta JA, Fillery IR. N application enhances remobilisation and reduces losses of pre-anthesis N in wheat grown on an Duplex soil. Aust J Agric Res 1995;46(3):519–31.
[86] Papakosta DK, Gagianas AA. Nitrogen and Dry Matter Accumulation, Remobilization, and Losses for Mediterranean Wheat during Grain Filling. Agron J 1991;83(5):864–70.
[87] Kong L, Xie Y, Hu L, Feng B, Li S. Remobilization of vegetative nitrogen to developing grain in wheat (Triticum aestivum L.). F Crop Res 2016;196:134–44.
[88] Revill A, Florence A, MacArthur A, Hoad SP, Rees RM, Williams M. The value of Sentinel-2 spectral bands for the assessment of winter wheat growth and development. Remote Sens 2019;11(17):1–18.
[89] Revill A, Florence A, Macarthur A, Hoad S, Rees R, Williams M. Quantifying uncertainty and bridging the scaling gap in the retrieval of leaf area index by coupling sentinel-2 and UAV observations. Remote Sens 2020;12(11):1843–61.
[90] Daughtry CS, Walthall CL, Kim MS, Brown de Colstoun E, McMurtrey III JE. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment 2000(74):229–39.

Please cite this article as: J. Segarra, F. Z. Rezzouk, N. Aparicio et al., Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, Information Processing in Agriculture, https://doi.org/10.1016/j.inpa.2022.05.004