Contents lists available at ScienceDirect



ISPRS Journal of Photogrammetry and Remote Sensing

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



Estimation of grain protein content in commercial bread and durum wheat fields via traits inverted by radiative transfer modelling from Sentinel-2 timeseries

A Longmire^{a,*}, T Poblete^{a,b}, A Hornero^{b,c}, D Chen^a, P.J Zarco-Tejada^{a,b,c}

^a School of Agriculture, Food, and Ecosystem Sciences, Faculty of Science, University of Melbourne, Parkville 3010, Victoria, Australia

^b Department of Infrastructure Engineering, Faculty of Engineering and Information Technology (FEIT), University of Melbourne, Parkville 3010, Victoria, Australia

^c Instituto de Agricultura Sostenible (IAS), Consejo Superior de Investigaciones Científicas (CSIC), Avenida Menéndez Pidal s/n, 14004 Córdoba, Spain

ARTICLE INFO

Keywords: Wheat Grain protein content Stress detection Sentinel-2 time series Radiative transfer model inversion Machine learning

ABSTRACT

Wheat (Triticum spp.) is crucial to food security. Grain protein content (GPC) is key to its nutritional and economic value and is controlled by genetic and agronomic factors, soil properties and weather. GPC prediction from remote sensing could reduce nitrogen (N) losses, help management decisions, and improve profit. However, GPC prediction is complex because multiple plant traits influence GPC and their effects change through the growing season. Traits with known physiological links to GPC, which can be retrieved from imaging spectroscopy, include leaf area index (LAI), chlorophyll (C_{a+b}) , and stress indicators. Further inspection of these and other traits retrieved from satellite data can advance research relevant to precision agriculture. Sentinel-2 (S2) timeseries (TS) were acquired for 6,355 ha of commercial dryland bread (T. aestivum) and durum (T. durum) wheat fields in south-east Australia through two consecutive years with dissimilar rainfall. Wheat growers provided \sim 92,000 GPC data points from harvester-mounted protein monitors. For each, C_{a+b} , leaf dry matter, leaf water content (Cw) and LAI were retrieved from the S2 images by radiative transfer model inversion. A gradient boosted machine learning algorithm was applied to analyse these traits' importance to GPC and to predict GPC in 30% of samples unseen by the algorithm in training. The strongest relationships between predicted and observed GPC ($R^2 = 0.86$, RMSE = 0.56 %), in a model built from five S2 images across a season, were better than those from single-date hyperspectral (HS). In severe water stress, LAI was the main predictor of GPC early in the season, but this switched to Cw later. Trait importance was more evenly distributed in milder conditions. S2 TS had a clear accuracy advantage over single-date S2 and HS, especially in benign conditions, emphasising the potential of S2 TS for large-scale GPC monitoring.

1. Introduction

Wheat (*Triticum* spp.) supplies around 20% of humans' carbohydrate (CHO) and protein intake (FAO, 2022). Grain protein content (GPC; %) dictates the price paid to growers and is influenced by the amount of nitrogen (N) in the canopy at anthesis (Zadoks stage Z65; Giuliani et al., 2011; Masoni et al., 2007; Zadoks et al., 1974), and the N accessible for uptake during grain filling (Gooding et al., 2007). However, total photosynthesis after Z65 controls protein dilution by new assimilates, a long-recognised inverse yield ~ GPC relationship (McNeal et al., 1978). Fertiliser N is a major cost and risky investment for grain growers (Monjardino et al., 2015), and entails a heavy environmental footprint (Galloway et al., 2017). The global urea price increased by > 400 % in

two years to April 2022, (Baffes and Koh, 2022); relief is unlikely as fossil fuels are deeply embedded in fertiliser supply. Timely GPC estimates could improve farmer access to price premiums (Apan et al., 2006; Skerritt et al., 2002), guide N applications aimed at increasing protein and improve farm environmental and economic sustainability.

Leaf area index (LAI) and chlorophyll (C_{a+b}) content strongly influence assimilation (Wolanin et al., 2019), hence protein dilution, while above-ground N correlates with both C_{a+b} (Evans, 1989) and GPC (Feng et al., 2014; Xue et al., 2007; Zhao et al., 2005). C_{a+b} /N concentration and LAI affect both the total protein available for translocation (Masclaux-Daubresse et al., 2010) and, via photosynthetic capacity, the CHO source size. Despite these relationships, C_{a+b} /N may be only moderately correlated with GPC (Longmire et al., 2022; Zhao et al.,

* Corresponding author. E-mail address: alongmire@unimelb.edu.au (A. Longmire).

https://doi.org/10.1016/j.isprsjprs.2023.10.018

Received 3 May 2023; Received in revised form 15 October 2023; Accepted 27 October 2023

0924-2716/© 2023 The Authors. Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).



Fig. 1. Zone 1 (Kaniva; a) and zone 2 (Manangatang; b) commercial wheat fields considered in this study. Subfigure a: Sentinel-2, 2020–09–10 (R = band 8, G = band 4, B = band 3); b; 2019–09–11 (R = band 4, G = band 3, B = band 2).

2005) because other factors, especially water, nutrient and temperature stress, also affect N/CHO sources and sinks, the rate and duration of post-anthesis photosynthesis and the fate of assimilates during grain filling (Asseng et al., 2002; Masclaux-Daubresse et al., 2010). These dynamics complicate GPC estimation relative to that of its contributory variables and yield, as seen in lower estimation accuracies (e.g. Rodrigues et al., 2018; Wright et al., 2004; Zhao et al., 2005).

Radiative transfer models (RTM) elucidate mechanistic relationships between leaf and canopy traits and light interactions (Jacquemoud et al., 1995; Jacquemoud and Baret, 1990). Because they are robust to nonlinearity in these relationships and account for factors including leaf optical properties, soil reflectance and sun position, RTMs are transferable across crop types, phenology and agronomic situations (Clevers and Kooistra, 2012; Dorigo et al., 2007; Jacquemoud et al., 1995). Through inversion, they allow accurate estimation of plant traits relevant to GPC (Bacour et al., 2002; Féret et al., 2008; Jacquemoud et al., 2009, 1995; Li et al., 2015), where structural and physiological traits can be retrieved concurrently (Poblete et al., 2020; Zarco-Tejada et al., 2018). Recent studies have retrieved leaf C_{a+b} , canopy C_{a+b} content (CCC; Clevers et al., 2017; Houlès et al., 2007) and LAI from Sentinel-2 (S2) data by RTM inversion (Delloye et al., 2018; Pan et al., 2019; Upreti et al., 2019; Wolanin et al., 2020), but the work of Longmire et al. (2022) appears to remain unique in using RTM inversions to estimate GPC.

The multispectral instruments (MSI) aboard the S2 satellites offer cost-free images optimised for observing vegetation, with 13 bands at a ground resolution of 10–20 m in the visible, red edge (RE) and near-infrared domains (Drusch et al., 2012). The strong RE focus of the MSI, where it records three bands, is optimised for C_{a+b} , N content and LAI estimation (Frampton et al., 2013; Herrmann et al., 2011). A five-day revisit time facilitates timeseries (TS) observations. Hunt et al. (2019) obtained a substantial improvement in wheat yield estimation on

adding a second S2 TS image, diminishing returns for further images but better performance closer to harvest, as Wolanin and colleagues (2020) saw. Similarly, Wang et al. (2014) reported poor correlation between early-season broadband vegetation indices (VI) and wheat GPC; relationships improved as the season progressed and stacked TS images were best. Cumulative VIs have also improved yield but not GPC estimates (Xue et al., 2007).

Longmire et al. (2022) demonstrated that stress-sensitive hyperspectral (HS) plant traits anthocyanins (Anth), C_{a+b} , and carotenoids (C_{x+c}), the Photochemical Reflectance Indices PRI (Gamon *et al.* 1992), PRI_{m3} and PRI_{m4} (Hernández-Clemente *et al.*, 2011) and solar-induced fluorescence (SIF) are associated with GPC. In water-stressed crops, the thermal crop water stress index (CWSI; Idso, 1982) also showed a strong, positive association with GPC. While the use of S2 data precludes retrieval of stress-related HS and thermal traits, it offers potentially major advantages for GPC estimation at yet larger scales and through TS.

The use of machine- and deep-learning algorithms with RS data is widespread in agriculture, including for disease detection and monitoring (e.g. Adam et al., 2017; Poblete et al., 2020; Zarco-Tejada et al., 2018), weed recognition (Gao et al., 2018), crop- and land use classification (Abdi, 2020; Ji et al., 2018), primary productivity and yield (e.g. Cheng et al., 2022; Gómez et al., 2021; Hunt et al., 2019; Wolanin et al., 2020, 2019), and GPC estimation (Longmire et al., 2022; Tan et al., 2020; Zhou et al., 2021). Zhou et al. (2021) obtained their best GPC predictions with a random forest (RF) and found machine learning (ML) superior to traditional statistical methods. Gradient boosting machines (GBM) are a supervised ML algorithm based on the work of Friedman (2002, 2001) and developed by Chen and Guestrin (2016). The GBM has seen relatively little use in agriculture (van Klompenburg et al., 2020), but performed as well as other algorithms in estimating LAI and CCC in wheat (Upreti et al., 2019). The tree-based GBM algorithm is able to

Location, rainfall, climate and crop areas for Kaniva (Zone 1) and Manangatang (Zone 2) in 2019–20. AAR/GSR = long term average annual / growing season (Apr–Oct) rainfall; rain = annual total (mm).

Zone	Lat.	Lon.	Alt. (m)	AAR	Clim. zone	Year	Rain	GSR	Area (Ha)
1	-36.37°	141.24°	142	449	Cfb	2019	288	238	662
						2020	444	291	858
2	-35.05°	142.88°	55	316	Bsk	2019	194	135	2341
						2020	342	277	2494

assess the relative contribution, termed importance or gain, of input features to target variable estimation (Abdi, 2020; Hunt et al., 2019). This greatly improves the algorithms' interpretability, which is crucial to current research into the dynamic and interacting plant traits influencing GPC.

Some studies estimating wheat GPC from satellites have had moderate success. Wright et al. (2004) found the green NDVI (Gitelson and Merzlyak, 1998) best for GPC (airborne $R^2 = 0.53$, satellite $R^2 = 0.48$). Liu *et al.* (2005) combined synthetic aperture radar with the structure insensitive pigment index (SIPI; Peñuelas et al., 1995), from Landsat data, with $R^2 = 0.56$. Also from Landsat, Zhao et al. (2005) reported relationships between VI_{green} (Gitelson et al., 2002) and GPC ($R^2 = 0.46$), while Feng et al. (2014) estimated GPC at field level in commercial wheat, combining MODIS satellite NDVI from two growth stages and obtaining $R^2 = 0.567$ –0.632 and rRMSE = 0.141–0.144. From S2 data, Zhao et al. (2019) retrieved several plant N indicators with good fidelity, but these predicted GPC with lower skill ($R^2 = 0.428$ –0.467); like others (Wang et al., 2014; Zhao et al., 2005, 2019), this study reported that growth stages $\geq Z65$ offered the best estimations.

A large majority of studies estimating GPC from spectroscopy have used plot experiments (e.g. Raya-Sereno et al., 2021; Walsh et al., 2023; Wang et al., 2004; Zhao et al., 2005, 2019), and to date these appear largely to have used VIs, with no reference to RTM inversions. This research gap demands attention, particularly given the lack of consistency among VIs (Raya-Sereno et al., 2021). Moreover, few studies look at GPC variability within commercial fields; while Rodrigues et al. (2018) and Stoy et al. (2022) do so, considering natural soil variability, only Longmire et al. (2022) sample multiple fields. Finally, the additive combination of information from TS images is rare in the canon, has had limited success and, due to reliance on VIs, is likely poorly transferable between agronomic situations (Feng et al., 2014; Rodrigues et al., 2018; Xue et al., 2007).

To advance precision agriculture, there is a need to improve GPC prediction within fields but across large extents. The current research addresses the gaps identified above by combining satellite TS and RTM inversion to predict GPC within many commercial wheat fields, across regions with very different climatic and soil characteristics, diverse seasons and both bread- and durum cultivars. Leveraging the GBM's flexibility, interpretability and predictive skill, plant trait importances to GPC estimation and model performance are compared between S2 and airborne HS/thermal RS. The effects of bandset reduction and the facultative inclusion of airborne CWSI and/or SIF to S2 traits are tested. In the temporal dimension, trait importance dynamics and skill are comprehensively assessed through seasons, first with TS elements as separate models, then additively stacking them within site-years to form single predictive models.

2. Materials and methods

2.1. Study sites

This study considers 6355 Ha of rainfed commercial hard white bread (cv. Scepter, Vixen, Catapult) and durum (cv. Aurora, Bitalli) wheat crops grown in two areas of the southern Australian wheat belt. Both varieties were sown in late May and mid-June in 2019 and 2020 around Kaniva (zone 1; Fig. 1a), while bread wheat only was sown in similar periods around Manangatang (zone 2; Fig. 1b). Hereafter, the cropping zones are designated cz1 and cz2, with the year appended, e.g., cz1-19. Commercial cropping soils and fields in cz1 are described in Longmire et al. (2022); cz2 soils, Calcarosols in the Australian Soil Classification (Isbell, 2002), vary greatly across tens to hundreds of metres in clay fraction, subsoil constraints including B and Al, calcrete, and hardpans. These influence plant-available water (PAW) and N availability, root growth and harvest outcomes (Nuttall et al., 2003; Sadras et al., 2002). Climate, including Köppen-Geiger classification (Peel et al., 2007), and location details are provided in Table 1. Fertiliser was applied 1–3 times each season in each field, according to grower assessment of conditions, usually as urea, but data relating to these applications were not consistently available.

2.2. Data collection

Harvester-mounted NIR spectrometers (CropScan 3000/3300H, Next Instruments, Sydney, Australia) collected GPC during harvesting, with geolocation (± 0.01 m) by real-time kinetic GPS. These devices use near infrared (NIR) spectroscopy (720–1100 nm) to estimate GPC for \sim 400 ml grain samples temporarily removed from the combine's clean grain elevator. These devices assess GPC with accuracy sufficient to meet legal requirements for weights and measures in the United States and Australia (Clancy and Heiken, n.d.). Readings are taken every 8-10 s, representing GSD \approx 10–25 m parallel to harvester travel and, perpendicularly, equivalent to the swath width (12 m). Over cz1 only, HS and thermal data were collected by sensors on a light aircraft at ~ 2000 m above ground level (AGL) on 2019-10-22 and 2020-10-28. This gave pixels of 1.0 m (HS) and 1.7 m (thermal) ground sampling distance (GSD). HS data were collected with a VNIR E-Series model (Headwall Photonics, Fitchburg, MA, USA), capturing 371 bands from 400 to 1000 nm at 8 nm per pixel, yielding 5.8 nm FWHM with a 25 µm slit. At 12-bit radiometric resolution, the storage rate was 50 frames per second with an exposure time of 18 ms and an 8 mm focal length. The hyperspectral imager was calibrated using an integrating sphere (Labsphere XTH2000C, Labsphere Inc., North Sutton, NH, USA), deriving coefficients at four illumination levels. Thermal images were collected from 7.5 to 14 µm with an A655c camera (Teledyne FLIR LLC, Wilsonville, OR, USA), a scientific-grade instrument radiometrically calibrated by the manufacturer. A further indirect calibration was carried out during flights using ground observations from a handheld infrared thermometer (LaserSight from Optris GmbH, Berlin, Germany), after Calderón et al. (2015). Atmospheric correction of radiance was applied with the Simple Model of Atmospheric Radiative Transfer of Sunshine (SMARTS) model (Gueymard, 1995) using aerosol optical depth (AOD) observed at the time of flight (Micro-Tops II sunphotometer, Solar LIGHT Co., Philadelphia, PA, USA), as done before (Calderón et al., 2015; Poblete et al., 2020; Zarco-Tejada et al., 2018). Orthorectification was performed using Parametric Geocoding and Orthorectification for Airborne Optical Scanner Data (PARGE; ReSe applications GmbH, Wil, Switzerland) using an inertial measurement unit and GPS data from a VN-300 (VectorNav Technologies LLC, Dallas, TX, USA). As for S2 data, Level 1C orthorectified top-of-atmosphere reflectance (Richter et al., 2011) rasters for tiles 54HWE (cz1), 54HXG and 54HYG (cz2) were downloaded from the Copernicus Open Access Hub and atmospherically corrected to surface reflectance with Sen2Cor v. 2.3.1. Bands of GSD =

Cloud-free Sentinel-2 images available in zones 1 and 2 in 2019 and 2020 with associated growing degree days after sowing (GDDAS; °C day) and Zadoks (Z) stage/ name. Entries in bold are those compared against airborne hyperspectral analyses.

	Image		Bread			Durum		
Zone	Year	Date	GDDAS	Z stage	Z name	GDDAS	Z stage	Z name
1	2019	18-Jul	481	15	seedling	684	16	seedling
		23-Jul	537	15		739	17	
		17-Aug	770	17		972	31	stem elong.
		11-Sep	1008	32	stem elong.	1211	37	
		1-Oct	1237	43	booting	1439	52	ear emerg.
		21-Oct	1508	69	anthesis	1710	74	grain fill
	2020	17-Jul	452	14	seedling	560	15	seedling
		1-Aug	581	15		689	17	
		26-Aug	803	31	stem elong.	910	31	stem elong.
		10-Sep	979	32	-	1087	33	-
		10-Oct	1375	51	ear emerg.	1483	67	anthesis
2	2019	17-Jul	904	17	seedling			
		28-Jul	1028	31	stem elong.			
		12-Aug	1169	31	0			
		17-Aug	1220	32				
		27-Aug	1325	32				
		1-Oct	1778	54	ear emerg.			
	2020	17-Jul	719	17	seedling			
		27-Jul	803	17	-			
		26-Aug	1099	32	stem elong.			
		31-Aug	1161	32	-			
		10-Sep	1307	34				
		15-Sep	1385	42	booting			



Fig. 2. Schematic summary of data handling and machine learning processes.

20 m were resampled to 10 m prior to stacking as a multiband raster. All images between 1 July and 31 October in each season were assessed; 23, in which all subject fields were cloud free (Table 2), were retained.

2.3. Data extraction and processing

Potentially erroneous GPC points were discarded: those with unrealistic GPC values, within 20 m of trees, dams, fences and headlands, or

Values and ranges of leaf and canopy traits used for PRO4SAIL (PROSPECT-D + 4SAIL) radiative transfer model inversion and look-up table generation.

Parameter	Abbreviation	Unit	Value/range
Anthocyanins	Anth	µg/cm ²	1–10
Carotenoids	C_{x+c}		1–20
Chlorophyll $a + b$	C _{a+b}		3–70
Dry matter	Cm	g/cm ²	0.001-0.035
Hot spot parameter	h	-	0.01
Leaf area index	LAI	m^2/m^2	1–5
Leaf Inclination Dist. Func.	LIDFa	0	0–90
Mesophyll struct. Coef.	Ν	-	0.5-3.0
Observer angle	tto	0	0
Relative azimuth angle	ψ		0
Solar zenith angle	tts		varied with date
Water content	Cw	g/cm ²	0.001-0.035

in harvester turn/slow travel zones. A Wilcoxon test (Bauer, 1972) was applied across a) sites within year/wheat type combinations and b) wheat types within site/year combinations to test for differences in median GPC. S2 spectra were extracted to the GPC points, then compatible vegetation indices (VIs; Supplementary Table 1) were calculated – and inverted parameters retrieved (see Section 2.4) – for each point. Airborne narrow-band HS indices and inverted parameters, against which multispectral satellite results are compared, are detailed in Longmire et al. (2022); for these, CWSI and SIF (Supplementary Table 1), mean pixel values were calculated per 100 m² regions of interest (ROI). Each ROI had as its centroid a GPC point record, giving geolocation and GSD equivalent to the GPC data and quasi-equivalent to S2 data. SIF was calculated by the Fraunhofer line depth (FLD2) method (Plascyk and Gabriel, 1975). Data-handling processes are summarised in Fig. 2.

2.4. Radiative transfer model inversion

Leaf properties C_{a+b} , C_m and C_w were retrieved with PROSPECT-D

(Féret et al., 2017), while LAI was modelled with 4SAILH (Verhoef et al., 2007), coupled as PRO4SAIL. Simulated spectra were generated by randomly sampling the full range of model parameters, from uniform distributions in appropriate ranges for wheat (Table 3; Camino et al., 2018; Li et al., 2015). The hybrid PRO4SAIL inversions, after Xu et al. (2019), used look-up tables (LUT) of 200,000 simulations, shown sufficient previously (Longmire et al., 2022; Poblete et al., 2021; Xu et al., 2019; Zarco-Tejada et al., 2018). The LUTs were interrogated with support vector machine (SVM) ML algorithms to retrieve each trait independently, using as input the simulated reflectance spectra, convolved to the S2 spectral specifications and the target traits as outputs; such hybrid methods effectively address the ill-posed problem (Verrelst et al., 2015). The SVM models were built in MATLAB (MAT-LAB; Statistics and Machine Learning toolbox and Deep Learning toolbox; Mathworks Inc., Natick, MA, USA) and trained using a radial basis function and SVM hyperparameters optimised during training for each variable. With these trained SVM models, plant traits were inverted from observed reflectance at each GPC point or ROI in each site/image combination, varying solar zenith angle for the changing dates.

2.5. Variance inflation factor analysis

Multicollinearity between potential model input features crop traits was inspected by variance inflation factor analysis (VIF; R package *fsmb*; (Nakazawa, 2022)). Like other recent work, this study used VIF thresholds (t) of 5–10 (Akinwande et al., 2015; Magney et al., 2016; Poblete et al., 2021; Zarco-Tejada et al., 2018). VIF was repeated, with forced inclusion of inverted parameters, for each permutation of site, year and product. None of the 38 VIs calculated was kept at t = 5; those kept at t = 10 were added stepwise to ML models, to assess their contributions to skill. VIs that survived the VIF analysis were: Global Environment Monitoring Index (GEMI; Pinty and Verstraete, 1992), Maccioni Index (Macc; Maccioni et al., 2001), MERIS Terrestrial Chlorophyll Index (MTCI; Dash and Curran, 2004), Transformed Chlorophyll Absorption in Reflectance Index / Optimized Soil Adjusted VI (TCARI/



Fig. 3. Spatial variability in grain protein content (GPC; %) in wheat fields in zones 1 (a) and 2 (b) on Sentinel-2 images (R = band 4, G = band 3, B = band 2).



Fig. 4. Spatial variability in C_{a+b} (µg/cm²; a), C_m (g/cm²; b), C_w (g/cm²; c) and LAI (m²/m²; d) retrieved by radiative transfer model inversion from S2 images of a bread wheat field, 2019. Plots arranged by Zadoks growth stage; n = 4897.

OSAVI; Haboudane et al., 2002) and TCARI₁₆₁₀ (Herrmann et al., 2010). Each of these improved GPC estimation by $R^2 \leq 0.03$ over inverted traits only, so all were discarded. Airborne SIF was linearly independent so it was included where available. Features were also VIF tested along TS and between features within dates. Minor collinearity between inverted traits from close image dates (e.g. < 14 days apart) was disregarded in order to assess feature importance evolution.

2.6. Application of machine learning algorithm to estimate GPC

The study used a GBM to estimate relationships of inverted leaf and canopy traits - input features - with the target variable GPC, focusing on the features' relative importance. Feature importance is a unitless quantity in the range 0-1, expressing the relative gain contributed by each model input feature to estimation of the target variable. In each GBM run, data were randomly split 70%:30% into training and test sets (Hunt et al., 2019; Wu et al., 2021). Stochastic gradient descent (SGD) reduces the likelihood of overfitting by training models on random subsets of observations, introducing randomness (Friedman, 2002); here, models were trained on either 65% or 85% of rows, but all columns were included in every run. Randomised K-fold cross-validation (K = 5) provided further protection against overfitting. In addition to SGD, learning rate, tree depth and minimum node size were varied, with full factorial hyperparameter searching. The GBM was tuned at each run by finding the hyperparameter combination that minimised root mean square error (RMSE) of prediction (GPC %). These procedures mirror previous work (Longmire et al., 2022).

The ML algorithm was first run for each site/year/product combination, using as inputs the plant traits from the S2 image closest to the HS flight date, the last in the TS. To these features, airborne SIF was facultatively added. Each date was analysed in turn, whereby the input feature set contained only the four traits retrieved from the relevant image. Finally, TS were combined such that each permutation of inverted parameter and image date was taken as an individual input feature, disregarding minor collinearity between inverted parameters in the temporal dimension. To facilitate comparison across years, sites and crop types, growing degree days after sowing (GDDAS) were calculated after Asseng et al. (2010), based on daily temperatures and precipitation specific to each location, drawn from the SILO dataset (Jeffrey et al., 2001). Given the large potential for differences in phenological advance between locations and cultivars, and the need to compare these directly, APSIM Next Generation (Holzworth et al., 2018) was used to model phenology in Zadoks stages. For APSIM, met data were from SILO, sowing date was the mean of fields in each year/location, and cultivars were those most planted (cz1 bread, cv. Scepter; cz1 durum, cv. Aurora; cz2-19, cv. Scepter; cz2-20, cv. Kord).

3. Results

3.1. Fields, GPC, retrieved parameters

Growing conditions differed strongly between years at both locations: both total and growing season rainfall were extremely low in 2019 while 2020 was above average (Table 1). A Wilcoxon test (Bauer, 1972) showed significant differences in GPC between all zone/year/product combinations; effect sizes were small to moderate (not shown). Large GPC differences were seen within and between paddocks, often over short distances (Fig. 3).



Fig. 5. Plant traits C_{a+b} , C_m , C_w and LAI inverted from S2 images of a durum wheat field, 2019. Violin plots show distribution within Zadoks (Z) growth stage; white circle = mean, red cross = median; n = 782. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Plant traits for GPC points showed spatial heterogeneity and phenological progression for all site/year/product combinations. Example fields (cz1-19) are plotted to map traits over time for bread (Fig. 4) and to show progression with phenological advance, relationships between traits and density distributions within trait and stage combinations in durum (Fig. 5).

3.2. Feature importance and model performance

3.2.1. End-of-season Sentinel-2 images against hyperspectral images

The following considers first ML models built with traits inverted from S2 images captured as temporally close as possible to airborne HS missions detailed in Longmire et al. (2022). In cz1 bread, Cw was dominant in the very dry conditions of 2019, while in 2020 importance was spread evenly in a tight range (0.26–0.27) between C_{a+b} , C_w and C_m . In each year, LAI was the least important (Fig. 6a). cz1-19 durum had even importance across feature types, with C_w most important and C_{a+b} marginally higher than Cm, and in 2020 durum, feature importance was shared very evenly (Fig. 6b). Airborne SIF was added to S2 inverted features in each wheat type and year. SIF had low importance in cz1-19 bread and did not disturb the feature order relative to the base model, while in cz1-20 importance remained evenly distributed among Cm, Cw and C_{a+b} despite the inclusion of SIF (Fig. 6c). For cz1-19 durum, SIF was marginally more important than Cw but again feature order was otherwise unchanged compared to without SIF, while in cz1-20 durum SIF was most important and other features were approximately equal (Fig. 6d). In cz2–19 (bread wheat only), C_{a+b} was most important, with Cm and Cw intermediate and LAI lowest (not shown). In cz2-20, Cw took >40% of importance over $C_{a+b},$ while C_m and C_w were the lowest.

Relative feature importance was similar between S2 and HS. Across conditions and wheat types, the relative importance of S2 traits was parallel with that of HS features grouped by their type: C_{a+b} (S2) with HS

pigment indicators C_{a+b} , C_{x+c} and PRI; C_m with the HS structural reflectance index EVI; C_w with CWSI and LAI (S2) with HS LAI and LIDF_a (Fig. 6).

Predictive skill for S2 \pm SIF was highest in cz1-19 bread, as it was for HS \pm CWSI previously (Longmire et al., 2022). Overall, S2 models were better predictors than HS models only in cz1-19 bread (Fig. 7a). Adding SIF to S2 traits made only a minor difference in that context but improved GPC estimation relatively more from a lower base in each other site/year/product combination (Fig. 7).

3.2.2. Assessment with individual images in timeseries

Analysis of individual S2 TS components showed that for cz1-19 bread, feature importance was concentrated in LAI until around anthesis (Z60—Z69), when C_w became predominant; neither C_m nor C_{a+b} took importance at any time (Fig. 8). In the more benign 2020, in zone 1 durum wheat in both years and in zone 2 (not shown), feature importance was spread evenly at each stage, changing relatively little and without a discernible pattern through the season.

Model predictive skill in cz1-19 bread wheat was high during Z15 but diminished from Z17 until after anthesis when the final image was captured (Fig. 9). All site/year/product combinations showed better predictive performance in early development than in the mid-season, usually with an increase late in the season.

3.2.3. Stacked timeseries images

Use of all retrieved parameters from all crop stages together revealed patterns consistent with analysis of individual images: high LAI importance early in the cz1-19 bread wheat season, switching to C_w around anthesis, and a relatively even distribution of importance across traits and stages in other crops (Fig. 10).

Predictive skill was compared between single image models and those incorporating all available trait/date combinations; the latter brought large improvements in all site/year/product combinations except cz1-19 bread (Table 4).

4. Discussion

4.1. Growing conditions and protein variability

GPC is a complex variable under genetic, environmental and management control (Zhao et al., 2019). While rainfall can be presumed invariant across fields within a region, and genetics within a field, soil properties vary widely within fields and have large effects on GPC independent of GSR. PAW differences influence GPC via lowered CHO assimilation, hence dilution, especially during grain filling. However, grain count is influenced by earlier conditions, especially around anthesis, is a primary determinant of sink size for both proteins and photosynthate and hence is a strong driver of GPC. Further, excessive early vigour in rainfed crops can dry soil so much that later photosynthesis is restricted, a phenomenon known as 'haying off' in which grain ends with high protein because it fails to fill with CHO (van Herwaarden et al., 1998).

Besides extreme dryness, cz1 saw other weather extremes during critical periods of 2019. From mid-September into early October, frosts (\approx -4 °C) occurred at crucial stages (Z42—69; booting, ear emergence and anthesis) for the durum crop, while bread wheat fields were less affected due to their location higher in the landscape and less susceptible growth stages (Z33—51). Frost around anthesis severely reduces grain count, and durum is more susceptible to both frost and heat damage than bread wheat (Beres et al., 2020; McCallum et al., 2019). Immediately after these frosts, four days had maxima of 35–38 °C, imposing high respiratory loads (Heskel et al., 2016) and with potential to cause permanent damage, reduce total N uptake (van Ittersum et al., 2003) and/or simply exceed the optimum for photosynthesis (Asseng et al., 2011; Lobell and Gourdji, 2012). Heat stress can also severely alter phenology: At T > 34 °C, senescence is accelerated by a factor of three, and sixfold at



Fig. 6. Importance (proportion) of inverted traits to GPC (%) estimation in bread (a, c) and durum wheat (b, d) grown in zone 1, 2019–20. Brown background: S2 inverted traits: C_{a+b} , C_m , C_w and LAI, \pm SIF. Blue background: EVI, PRI, Anth, C_{a+b} , C_{x+c} , LAI and LIDF_a inverted from airborne hyperspectral images. CWSI and SIF were also from airborne data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

T ≥ 36 °C, foreshortening grain filling (Asseng et al., 2011; Porter and Gawith, 1999). Moreover, droughted plants accumulate more heat. These effects reduce yield, especially if cumulative, and should therefore increase GPC (Asseng et al., 2011). Heat stress may therefore increase mean GPC but reduce its variability through a generalised reduction of assimilation. Lower assimilation during grain fill (cz1-19), associated with senescence, would also reduce variability in C_{a+b}, C_m, and SIF, diminishing these as GPC estimators, while likely enhancing the importance of moisture sufficiency measures, as we observed in both CWSI and C_w. These conditions likely affected both GPC and the difficulty of its prediction, differentially between bread and durum wheat.

In cz2, where rainfall is lower and crops are grown on former dunefields, phenomena influencing PAW and N, including both osmotic and physical impediments to root growth, arise from the dune-swale morphology and can be severe (Nuttall et al., 2003; Sadras et al., 2002). Though high-resolution soil data were unavailable, the influence of the cz2 dune-swale systems is seen in GPC (Fig. 3b); the dune effect on biomass can also be seen in the S2 RGB areas of this figure. In these areas, height within the dunes can be strongly discernible at harvest, including complete reversal of yield response between dune and swale between wet and dry years; mid-slope areas, which are relatively unaffected by soil variability, can occupy a high proportion of fields



Fig. 7. Mean skill in GPC prediction of models comprising S2 traits C_{a+b} , C_m , C_w and LAI, \pm airborne SIF and models comprising airborne EVI, PRI, Anth, C_{a+b} , C_{x+c} , LAI, LIDF_a and SIF, \pm CWSI from commercial bread and durum wheat fields in zone 1, 2019–20.

(Armstrong et al., 2009; Hoffmann et al., 2016). The severe soil variability, known yield response differences, and the proportion of fields affected, appear also to be reflected both in less definitive feature importance dynamics, and lower GPC estimation skill, as seen in cz2. This is supported by Rab et al. (2009): a large majority of the \sim 80 Ha they studied across four years in the same region showed great yield variability.

4.2. Plant trait contributions to protein estimation

To be important for GPC determination, a plant trait needs to vary across the crop in some physiological relationship with GPC. This biophysical reality is seen in traits' relative importance to GPC prediction across and between years. For single-image analyses of cz1 bread and durum wheat, and regarding both relative trait importance and predictive skill, our results mirrored those from airborne HS data (Fig. 6, Fig. 7). As no VI was both linearly independent and a strong contributor to model skill, it is concluded that S2 data, with retrieval as shown, provide information adequate to estimate GPC. This accords with the findings of Wolanin and colleagues (2019) in estimating complex traits via RTM inversion and ML. While it contrasts with work based on HS imaging (Longmire et al. 2022), where several narrow-band indices complemented inverted parameters, it is as expected for data of lower spectral and spatial resolution. In lieu of CWSI, excluded as collinear, Cw assumed high importance in the very dry conditions of cz1-19 (Fig. 5a, c), suggesting that PROSAIL Cw retrieval can replace thermal observations. Under dry conditions, Cw apparently retains more variability and

predictive power than other traits, especially C_{a+b} and C_m , whose low importance can be understood as complementarity but also highlights their low variability in such conditions. Physiologically, this is reasonable because they were established earlier in the season when water and nutrients were not limiting. C_{a+b} , C_m , and LAI contribute to the pools of both N and CHO available for translocation, a process impeded by water stress. Hence low water stress should increase these factors' influence on GPC, while conversely under high stress it is less than that of the ongoing photosynthetic rate. C_{a+b} and C_m importances are low throughout cz1-19 and vary least through all crops and seasons.

In contrast to VIs and CWSI, and despite its close links to C_{a+b} and C_w, SIF was linearly independent of all inverted parameters and where available was included in models. This independence, and the substantial extra skill it conferred to our predictions, are important findings in themselves. SIF proxies instantaneous photosynthesis, improving estimates of other complex physiological quantities (Camino et al., 2019); it follows that, as a measure of assimilation and hence protein dilution, it improves GPC estimation. Substantial improvements seen with SIF inclusion, especially where base model accuracy was low (Fig. 7), suggest that TS SIF could substantially improve GPC prediction, though perhaps with a strong role limited to grain filling and to benign moisture conditions. Indeed, like those of C_{a+b} and C_m , SIF contribution was minor in very dry conditions, as found by others (Cai et al., 2019; Sloat et al., 2021). Further, the relatively minor skill improvement on adding SIF in 2019 durum may relate to weather damage not picked up in the SIF signal but crucial to GPC. In mild conditions, the combined relationships of C_{a+b} , C_m and SIF with GPC were insufficient to offset the strong $C_w \sim$



Fig. 8. Relative importance (proportion; sum = 1 within each year/stage) of C_{a+b} , C_m , C_w and LAI inverted from timeseries S2 images and used as separate feature sets for GPC estimation in commercial bread (a) and durum wheat (b) in zone 1 in 2019 and 2020. Image capture dates by Zadoks stage from top down.



Fig. 9. Performance metrics (R^2 and RMSE) for GPC estimation from plant traits inverted from single S2 images of bread and durum wheat fields in cz1, 2019–2020. Metrics are shown as a function of growing degree days after sowing (GDDAS) at image capture.

stomatal conductance \sim photosynthesis \sim GPC dynamic in drought. LAI had low importance in single-image S2 analyses, as it and other structural components did in our HS models.

Structure played a far greater role in TS analyses because it encompasses the early part of seasons. For example, when TS images were considered separately for cz1-19 bread, LAI took 60–80% of total importance at each stage up to Z65, whereafter the emphasis switched to C_w (Fig. 8a). When all dates were pooled, Z15 LAI took > 40% of importance, and C_w at anthesis took > 30% (Fig. 10a). This sudden

change was seen only under drought, but a gradual switch from LAI to C_w was seen in cz1-19 durum also, despite a likely reduced crop water demand after frosting and noise from weather damage. In other situations, the distribution of importance between dates and traits was quite seen whether image dates were separate or pooled, but always included substantial contributions from LAI. This, and the decline of LAI from a mid-season peak when all date/feature combinations were pooled, shows the high sensitivity of GPC to above-ground biomass, as a source of both proteins and CHO. However, PAW variability is lower early in



Fig. 10. Relative importance (proportion; sum = 1 within each year) of C_{a+b} , C_m , C_w and LAI inverted from timeseries S2 images and used together as a single training feature set for GPC estimation in commercial bread (a) and durum wheat (b) in zone 1, 2019–2020. Features are arranged by Zadoks growth stage at image capture from top down, within plant trait category.

Model performance (\mathbb{R}^2 and RMSE) for GPC estimation from single, late-season S2 images (base) and TS models from 5 to 6 S2 images across the same season, captured over commercial bread and durum wheat fields in cropping zones 1 and 2.

			base model		timeseries model	
zone	year	wheat type	R^2	RMSE	R ²	RMSE
CZ1	2019	bread	0.82	0.65	0.86	0.56
		durum	0.38	0.93	0.55	0.81
	2020	bread	0.33	0.80	0.52	0.67
		durum	0.16	0.85	0.43	0.70
CZ2	2019	bread	0.37	1.97	0.55	1.68
	2020	bread	0.14	0.98	0.31	0.90

the season partly because crops have had less time – and biomass – to withdraw water. The growth of C_w importance through seasons likely reflects increasingly variable soil moisture as plants differentially accumulate biomass, hence capacity to dry the soil, and soil properties exert more influence due to drying.

4.3. Model predictive skill

GPC predictions from single S2 images late in seasons were substantially less accurate than those from HS data (Longmire et al., 2022), except in very dry conditions (Fig. 7). The lower S2 spectral resolution explains this: indicators linked to GPC through their detection of mild stress, inverted Anth and C_{x+c} , and the PRI, cannot be calculated from S2 data so are absent from those analyses, but made substantial contributions to HS predictions. Here the advantages of S2 TS become clear, whereby TS prediction metrics were as good, and sometimes better than those from HS (Fig. 7). The improvement over single-date S2 estimates was also large, especially in cz2, where they came off a low base (Table 4). These gains also come from the incorporation of early-season structural information and demonstrate that S2 TS can be used to predict GPC even in very difficult conditions. In all situations, predictive performance was better during early development than in the mid-season, confirming both that emergence and early vigour are important to GPC outcomes and that our ML approach is sensitive to the same biological reality.

Cai et al. (2019) assert that the optimal timing for Australian wheat yield prediction accuracy is before October; this may hold for GPC prediction also, given that the two quantities co-vary in opposition. Others contend that Z65 is effective (Tan et al., 2020; Zhao et al., 2019) and that the addition of extra S2 TS data after June gave little improvement (yield; England; Hunt et al., 2019). This study shows that despite the many factors that can intervene between potential and reality, observations of specific inverted parameters as early as Z15 can contribute to GPC prediction. It also confirms that later images bring higher accuracy, and that TS stacking further improves performance. A lack of late-season images leaves an unfortunate gap, notably in the milder, wetter 2020 when no image was available after Z50/1400 GDDAS in either zone but also in cz2-19. Performance may improve further with later cloud-free images, but the likelihood of finding these would not be substantially higher in any similarly mild season. Estimation of GPC is considerably more complex than estimation of intermediate quantities (Zhao et al., 2019). Nevertheless our accuracies (Table 4) are comparable to and often improve on other recent results from satellite RS: Zhao et al. (2019) recorded 0.428 $\leq R^2 \leq$ 0.467 for wheat GPC from S2, while Tan et al. (2020) achieved best metrics of R^2 = 0.81 and RMSE = 0.54 % from Landsat.

5. Conclusions

Plant and canopy traits, retrieved by radiative transfer modelling from Sentinel-2 image timeseries, were assessed for their contribution to, and ability to predict, wheat grain protein content in commercial fields and under diverse soil and weather. Using equivalent modelling methods from single images, GPC estimation accuracy was generally lower when based on S2 than on airborne HS traits. However, in very dry conditions, our best model using a single S2 image and made with PRO4SAIL-inverted C_{a+b} , C_m , C_w and LAI, outperformed that built with HS inputs and CWSI. Adding timeseries S2 inverted traits substantially improved all models over single-image versions; improvement was very strong in benign conditions and compensated for the accuracy reduction caused by switching from HS to S2. The best predictive performance was achieved by stacking retrieved parameters from all dates as inputs to a single model ($R^2 = 0.86$, RMSE = 0.56 %). The order and relative importance of S2 plant traits were similar to airborne HS: S2 importance was dominated by C_w in drought but evenly spread between structural and physiological features in benign conditions. The results obtained suggest potential applications in precision agriculture. Nevertheless, many refinements are possible: GPC and RS data from a wider range of seasonal and agronomic conditions, spaceborne SIF and ground-based sources such as soil moisture assessments should be tested for their impact on model skill.

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

The authors gratefully acknowledge Jonathan Dyer, Ryan O'Sullivan and Hayden Walters, grain growers and Simon Craig, grower/agronomist for access to GPC data. Andrew Longmire received PhD funding from the Grains Research and Development Corporation (GRDC) Australia. The GRDC had no role in designing or conducting the study.

Author Contributions

AL conceived the project, did the experimental design, negotiated access to GPC and other data at ground level, did the machine learning and data analyses, and wrote the manuscript. TP processed the image collections and performed the model inversions using radiative transfer models. AH obtained and pre-processed the Sentinel-2 data. AL, PZ and TP contributed to data interpretation, manuscript drafting and designed objectives and discussed results and conclusions. PZ supervised the work and led the airborne hyperspectral campaigns. DC supervised the work.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2023.10.018.

References

- Abdi, A.M., 2020. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. GIScience & Remote Sensing 57, 1–20. https://doi.org/10.1080/15481603.2019.1650447.
- Adam, E., Deng, H., Odindi, J., Abdel-Rahman, E.M., Mutanga, O., 2017. Detecting the Early Stage of Phaeosphaeria Leaf Spot Infestations in Maize Crop Using *In Situ* Hyperspectral Data and Guided Regularized Random Forest Algorithm. J. Spectrosc. 2017, 1. https://doi.org/10.1155/2017/6961387.
- Akinwande, M.O., Dikko, H.G., Samson, A., 2015. Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. Open J. Stat. 05, 754. https://doi.org/10.4236/ojs.2015.57075.
- Apan, A., Kelly, R., Phinn, S., Strong, W., Lester, D., Butler, D., Robson, A., 2006. Predicting grain protein content in wheat using hyperspectral sensing of in-season crop canopies and partial least squares regression. Int. J. Geoinform. 2, 93–108.
- Armstrong, R.D., Fitzpatrick, J., Rab, M.A., Abuzar, M., Fisher, P.D., O'Leary, G.J., 2009. Advances in precision agriculture in south-eastern Australia. III. Interactions between soil properties and water use help explain spatial variability of crop production in the Victorian Mallee. Crop Pasture Sci. 60, 870–884. https://doi.org/ 10.1071/CP08349.
- Asseng, S., Bar-Tal, A., Bowden, J.W., Keating, B.A., Van Herwaarden, A., Palta, J.A., Huth, N.I., Probert, M.E., 2002. Simulation of grain protein content with APSIM-Nwheat. Eur. J. Agron. 16, 25–42. https://doi.org/10.1016/S1161-0301(01)00116-2.
- Asseng, S., Foster, I., Turner, N.C., 2011. The impact of temperature variability on wheat yields. Glob. Chang. Biol. 17, 997–1012. https://doi.org/10.1111/j.1365-2486.2010.02262.x.
- Bacour, C., Jacquemoud, S., Leroy, M., Hautecœur, O., Weiss, M., Prévot, L., Bruguier, N., Chauki, H., 2002. Reliability of the estimation of vegetation characteristics by

inversion of three canopy reflectance models on airborne POLDER data. Agronomie 22, 555–565. https://doi.org/10.1051/agro:2002039.

- Baffes, J., Koh, W.C., 2022. Fertilizer prices expected to remain higher for longer. World Bank Blogs accessed 2022-05-26.
- Bauer, D.F., 1972. Constructing Confidence Sets Using Rank Statistics. J. Am. Stat. Assoc. 67, 687–690. https://doi.org/10.1080/01621459.1972.10481279.
- Beres, B.L., Rahmani, E., Clarke, J.M., Grassini, P., Pozniak, C.J., Geddes, C.M., Porker, K. D., May, W.E., Ransom, J.K., 2020. A Systematic Review of Durum Wheat: Enhancing Production Systems by Exploring Genotype, Environment, and Management (G × E × M) Synergies. Front. Plant Sci. 11 https://doi.org/10.3389/ fpls.2020.568657.
- Cai, Y., Guan, K., Lobell, D., Potgieter, A.B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., Peng, B., 2019. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. Agric. For. Meteorol. 274, 144–159. https://doi.org/10.1016/j.agrformet.2019.03.010.
- Calderón, R., Navas-Cortés, J.A., Zarco-Tejada, P.J., 2015. Early Detection and Quantification of Verticillium Wilt in Olive Using Hyperspectral and Thermal Imagery over Large Areas. Remote Sens. (Basel) 7, 5584–5610. https://doi.org/ 10.3390/rs70505584.
- Camino, C., González-Dugo, V., Hernández, P., Sillero, J.C., Zarco-Tejada, P.J., 2018. Improved nitrogen retrievals with airborne-derived fluorescence and plant traits quantified from VNIR-SWIR hyperspectral imagery in the context of precision agriculture. Int. J. Appl. Earth Obs. Geoinf. 70, 105–117. https://doi.org/10.1016/j. jag.2018.04.013.
- Camino, C., Gonzalez-Dugo, V., Hernandez, P., Zarco-Tejada, P.J., 2019. Radiative transfer Vcmax estimation from hyperspectral imagery and SIF retrievals to assess photosynthetic performance in rainfed and irrigated plant phenotyping trials. Remote Sens. Environ. 231, 111186 https://doi.org/10.1016/j.rse.2019.05.005.
- Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16. Association for Computing Machinery, New York, NY, USA, pp. 785–794. Doi: 10.1145/2939672.2939785.
- Cheng, E., Zhang, B., Peng, D., Zhong, L., Yu, L., Liu, Y., Xiao, C., Li, C., Li, X., Chen, Y., Ye, H., Wang, H., Yu, R., Hu, J., Yang, S., 2022. Wheat yield estimation using remote sensing data based on machine learning approaches. Front. Plant Sci. 13 https://doi. org/10.3389/fpls.2022.1090970.
- Clancy, P., Heiken, D., n.d. Montana Field Trials for On Combine NIR Analyzer. URL https://cropscanag.com/wp-content/uploads/2019/02/2015_Montana_Field_Tria ls_for_On_Combine_NIR_Analyzer.pdf.
- Clevers, J.G.P.W., Kooistra, L., 2012. Using Hyperspectral Remote Sensing Data for Retrieving Canopy Chlorophyll and Nitrogen Content. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5, 574–583. https://doi.org/10.1109/JSTARS.2011.2176468.
- Clevers, J.G.P.W., Kooistra, L., Van den Brande, M.M.M., 2017. Using Sentinel-2 Data for Retrieving LAI and Leaf and Canopy Chlorophyll Content of a Potato Crop. Remote Sens. (Basel) 9, 405. https://doi.org/10.3390/rs9050405.
 Dash, J., Curran, P.J., 2004. The MERIS terrestrial chlorophyll index. Int. J. Remote Sens.
- Dash, J., Curran, P.J., 2004. The MERIS terrestrial chlorophyll index. Int. J. Remote Sens. 25, 5403–5413. https://doi.org/10.1080/0143116042000274015.
- Delloye, C., Weiss, M., Defourny, P., 2018. Retrieval of the canopy chlorophyll content from Sentinel-2 spectral bands to estimate nitrogen uptake in intensive winter wheat cropping systems. Remote Sens. Environ. 216, 245–261. https://doi.org/10.1016/j. rse.2018.06.037.
- Dorigo, W.A., Zurita-Milla, R., de Wit, A.J.W., Brazile, J., Singh, R., Schaepman, M.E., 2007. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. International Journal of Applied Earth Observation and Geoinformation, Advances in airborne electromagnetics and remote sensing of agro-ecosystems 9, 165–193. https://doi.org/10.1016/j.jag.2006.05.003.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing of Environment, The Sentinel Missions - New Opportunities for Science 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.
- Evans, J.R., 1989. Photosynthesis and Nitrogen Relationships in Leaves of C Plants. Oecologia 78, 9–19.
- FAO, 2022. FAOSTAT (production statistics), Crops and livestock products. Food and Agriculture Organization of the United Nations Statistics Division, Rome, Italy.
- Feng, M., Xiao, L., Zhang, M., Yang, W., Ding, G., 2014. Integrating Remote Sensing and GIS for Prediction of Winter Wheat (*Triticum aestivum*) Protein Contents in Linfen (Shanxi). China. PLOS ONE 9, e80989.
- Féret, J.-B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidel, L.P.R., Ustin, S. L., le Maire, G., Jacquemoud, S., 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. Remote Sens. Environ. 112, 3030–3043. https://doi.org/10.1016/j.rse.2008.02.012.
- Féret, J.-B., Gitelson, A.A., Noble, S.D., Jacquemoud, S., 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sens. Environ. 193, 204–215. https://doi.org/10.1016/j.rse.2017.03.004.
- Frampton, W.J., Dash, J., Watmough, G., Milton, E.J., 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS J. Photogramm. Remote Sens. 82, 83–92. https://doi.org/10.1016/j. isprsjprs.2013.04.007.
- Friedman, J.H., 2001. Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics 29, 1189–1232.
- Friedman, J.H., 2002. Stochastic gradient boosting. Computational Statistics & Data Analysis, Nonlinear Methods and Data Mining 38, 367–378. https://doi.org/ 10.1016/S0167-9473(01)00065-2.

- Galloway, J.N., Leach, A.M., Erisman, J.W., Bleeker, A., 2017. Nitrogen: the historical progression from ignorance to knowledge, with a view to future solutions. Soil Res. 55, 417–424. https://doi.org/10.1071/SR16334.
- Gao, J., Nuyttens, D., Lootens, P., He, Y., Pieters, J.G., 2018. Recognising weeds in a maize crop using a random forest machine-learning algorithm and near-infrared snapshot mosaic hyperspectral imagery. Biosyst. Eng. 170, 39–50. https://doi.org/ 10.1016/j.biosystemseng.2018.03.006.
- Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. Remote Sens. Environ. 80, 76–87. https:// doi.org/10.1016/S0034-4257(01)00289-9.
- Gitelson, A.A., Merzlyak, M.N., 1998. Remote sensing of chlorophyll concentration in higher plant leaves. Advances in Space Research, Synergistic Use of Multisensor Data for Land Processes 22, 689–692. https://doi.org/10.1016/S0273-1177(97)01133-2.
- Giuliani, M.M., Giuzio, L., Caro, A.D., Flagella, Z., 2011. Relationships between Nitrogen Utilization and Grain Technological Quality in Durum Wheat: I. Nitrogen Translocation and Nitrogen Use Efficiency for Protein. Agron. J. 103, 1487–1494. https://doi.org/10.2134/agronj2011.0153.
- Gómez, D., Salvador, P., Sanz, J., Časanova, J.L., 2021. Modelling wheat yield with antecedent information, satellite and climate data using machine learning methods in Mexico. Agric. For. Meteorol. 300, 108317 https://doi.org/10.1016/j. agrformet 2020 108317
- Gooding, M.J., Gregory, P.J., Ford, K.E., Ruske, R.E., 2007. Recovery of nitrogen from different sources following applications to winter wheat at and after anthesis. Field Crop Res 100, 143–154. https://doi.org/10.1016/j.fcr.2006.06.002.
- Gueymard, C., 1995. Simple Model for the Atmospheric Radiative Transfer of Sunshine (SMARTS2) Algorithms and performance assessment. University of Central Florida, Florida Solar Energy Center.
- Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens. Environ. 81, 416–426. https:// doi.org/10.1016/S0034-4257(02)00018-4.
- Herrmann, I., Karnieli, A., Bonfil, D.J., Cohen, Y., Alchanatis, V., 2010. SWIR-based spectral indices for assessing nitrogen content in potato fields. Int. J. Remote Sens. 31, 5127–5143. https://doi.org/10.1080/01431160903283892.
- Herrmann, I., Pimstein, A., Karnieli, A., Cohen, Y., Alchanatis, V., Bonfil, D.J., 2011. LAI assessment of wheat and potato crops by VENµS and Sentinel-2 bands. Remote Sens. Environ. 115, 2141–2151. https://doi.org/10.1016/j.rse.2011.04.018.
- Heskel, M.A., O'Sullivan, O.S., Reich, P.B., Tjoelker, M.G., Weerasinghe, L.K., Penillard, A., Egerton, J.J.G., Creek, D., Bloomfield, K.J., Xiang, J., Sinca, F., Stangl, Z.R., Martinez-de la Torre, A., Griffin, K.L., Huntingford, C., Hurry, V., Meir, P., Turnbull, M.H., Atkin, O.K., 2016. Convergence in the temperature response of leaf respiration across biomes and plant functional types. Proceedings of the National Academy of Sciences 113, 3832–3837. Doi: 10.1073/pnas.1520282113.
- Hoffmann, M.P., Llewellyn, R.S., Davoren, C.W., Whitbread, A.M., 2016. Assessing the Potential for Zone-Specific Management of Cereals in Low-Rainfall South-Eastern Australia: Combining On-Farm Results and Simulation Analysis. J Agro Crop Sci n/an/a. https://doi.org/10.1111/jac.12159.
- Holzworth, D., Huth, N.I., Fainges, J., Brown, H., Zurcher, E., Cichota, R., Verrall, S., Herrmann, N.I., Zheng, B., Snow, V., 2018. APSIM Next Generation: Overcoming challenges in modernising a farming systems model. Environ. Model. Softw. 103, 43–51. https://doi.org/10.1016/j.envsoft.2018.02.002.
- Houlès, V., Guérif, M., Mary, B., 2007. Elaboration of a nitrogen nutrition indicator for winter wheat based on leaf area index and chlorophyll content for making nitrogen recommendations. Eur. J. Agron. 27, 1–11. https://doi.org/10.1016/j. eia.2006.10.001.
- Hunt, M.L., Blackburn, G.A., Carrasco, L., Redhead, J.W., Rowland, C.S., 2019. High resolution wheat yield mapping using Sentinel-2. Remote Sens. Environ. 233, 111410 https://doi.org/10.1016/j.rse.2019.111410.
- Isbell, R., 2002. The Australian soil classification. CSIRO publishing, Melbourne, Australia.
- Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sens. Environ. 34, 75–91. https://doi.org/10.1016/0034-4257(90)90100-Z.
- Jacquemoud, S., Baret, F., Andrieu, B., Danson, F.M., Jaggard, K., 1995. Extraction of vegetation biophysical parameters by inversion of the PROSPECT + SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. Remote Sens. Environ. 52, 163–172. https://doi.org/10.1016/0034-4257(95)00018-V.
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L., 2009. PROSPECT+SAIL models: A review of use for vegetation characterization. Remote Sensing of Environment, Imaging Spectroscopy Special Issue 113, S56–S66. https://doi.org/10.1016/j.rse.2008.01.026.
- Jeffrey, S.J., Carter, J.O., Moodie, K.B., Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environ. Model. Softw. 16, 309–330. https://doi.org/10.1016/S1364-8152(01)00008-1.
- Ji, S., Zhang, C., Xu, A., Shi, Y., Duan, Y., 2018. 3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images. Remote Sens. (Basel) 10, 75. https://doi.org/10.3390/rs10010075.
- Li, Z., Jin, X., Wang, J., Yang, G., Nie, C., Xu, X., Feng, H., 2015. Estimating winter wheat (*Triticum aestivum*) LAI and leaf chlorophyll content from canopy reflectance data by integrating agronomic prior knowledge with the PROSAIL model. Int. J. Remote Sens. 36, 2634–2653. https://doi.org/10.1080/01431161.2015.1041176.
- Lobell, D.B., Gourdji, S.M., 2012. The Influence of Climate Change on Global Crop Productivity. Plant Physiol. 160, 1686–1697. https://doi.org/10.1104/ pp.112.208298.
- Longmire, A.R., Poblete, T., Hunt, J.R., Chen, D., Zarco-Tejada, P.J., 2022. Assessment of crop traits retrieved from airborne hyperspectral and thermal remote sensing

imagery to predict wheat grain protein content. ISPRS J. Photogramm. Remote Sens. 193, 284–298. https://doi.org/10.1016/i.isprsiprs.2022.09.015.

- Maccioni, A., Agati, G., Mazzinghi, P., 2001. New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra.
 J. Photochem. Photobiol. B Biol. 61, 52–61. https://doi.org/10.1016/S1011-1344 (01)00145-2.
- Magney, T.S., Eitel, J.U.H., Huggins, D.R., Vierling, L.A., 2016. Proximal NDVI derived phenology improves in-season predictions of wheat quantity and quality. Agric. For. Meteorol. 217, 46–60. https://doi.org/10.1016/j.agrformet.2015.11.009.
- Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., Suzuki, A., 2010. Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. Ann. Bot. 105, 1141–1157. https://doi.org/10.1093/aob/mcq028.
- Masoni, A., Ercoli, L., Mariotti, M., Arduini, I., 2007. Post-anthesis accumulation and remobilization of dry matter, nitrogen and phosphorus in durum wheat as affected by soil type. Eur. J. Agron. 26, 179–186. https://doi.org/10.1016/j.eja.2006.09.006.
- McCallum, M., Peirce, C., Porker, K., 2019. What drives the yield gap between durum and bread wheat?, in: Proceedings of the 19th Australian Agronomy Conference. Australian Society of Agronomy, Wagga Wagga, N.S.W., p. 4.
- McNeal, F.H., McGuire, C.F., Berg, M.A., 1978. Recurrent Selection for Grain Protein Content in Spring Wheat. Crop Sci. 18, 779–782. https://doi.org/10.2135/ cropsci1978.0011183X001800050022x.
- Monjardino, M., McBeath, T., Ouzman, J., Llewellyn, R., Jones, B., 2015. Farmer riskaversion limits closure of yield and profit gaps: A study of nitrogen management in the southern Australian wheatbelt. Agr. Syst. 137, 108–118. https://doi.org/ 10.1016/j.agsv.2015.04.006.

Nakazawa, M., 2022. Functions for Medical Statistics Book with some Demographic Data. R package "fsmb.".

- Nuttall, J.G., Armstrong, R.D., Connor, D.J., Matassa, V.J., 2003. Interrelationships between edaphic factors potentially limiting cereal growth on alkaline soils in northwestern Victoria. Aust. J. Soil Res. 41, 277. https://doi.org/10.1071/SR02022.
- Pan, H., Chen, Z., Ren, J., Li, H., Wu, S., 2019. Modeling Winter Wheat Leaf Area Index and Canopy Water Content With Three Different Approaches Using Sentinel-2 Multispectral Instrument Data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 482–492. https://doi.org/10.1109/JSTARS.2018.2855564.
- Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci. 11, 1633–1644. https://doi. org/10.5194/hess-11-1633-2007.
- Peñuelas, J., Filella, I., Gamon, J.A., 1995. Assessment of photosynthetic radiation-use efficiency with spectral reflectance. New Phytol. 131, 291–296. https://doi.org/ 10.1111/j.1469-8137.1995.tb03064.x.
- Pinty, B., Verstraete, M.M., 1992. GEMI: a non-linear index to monitor global vegetation from satellites. Vegetatio 101, 15–20. https://doi.org/10.1007/BF00031911.
- Plascyk, J.A., Gabriel, F.C., 1975. The Fraunhofer Line Discriminator MKII-An Airborne Instrument for Precise and Standardized Ecological Luminescence Measurement. IEEE Trans. Instrum. Meas. 24, 306–313. https://doi.org/10.1109/ TIM.1975.4314448.
- Poblete, T., Camino, C., Beck, P.S.A., Hornero, A., Kattenborn, T., Saponari, M., Boscia, D., Navas-Cortes, J.A., Zarco-Tejada, P.J., 2020. Detection of *Xylella fastidiosa* infection symptoms with airborne multispectral and thermal imagery: Assessing bandset reduction performance from hyperspectral analysis. ISPRS J. Photogramm. Remote Sens. 162, 27–40. https://doi.org/10.1016/j.isprsjprs.2020.02.010.
- Poblete, T., Navas-Cortes, J.A., Camino, C., Calderon, R., Hornero, A., Gonzalez-Dugo, V., Landa, B.B., Zarco-Tejada, P.J., 2021. Discriminating *Xylella fastidiosa* from Verticillium dahliae infections in olive trees using thermal- and hyperspectral-based plant traits. ISPRS J. Photogramm. Remote Sens. 179, 133–144. https://doi.org/ 10.1016/j.isprsjprs.2021.07.014.
- Porter, J.R., Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. Eur. J. Agron. 10, 23–36. https://doi.org/10.1016/S1161-0301(98)00047-
- Rab, M.A., Fisher, P.D., Armstrong, R.D., Abuzar, M., Robinson, N.J., Chandra, S., 2009. Advances in precision agriculture in south-eastern Australia. IV. Spatial variability in plant-available water capacity of soil and its relationship with yield in site-specific management zones. Crop Pasture Sci. 60, 885–900. https://doi.org/10.1071/ CP08350.
- Raya-Sereno, M.D., Ortiz-Monasterio, J.I., Alonso-Ayuso, M., Rodrigues, F.A., Rodríguez, A.A., González-Perez, L., Quemada, M., 2021. High-Resolution Airborne Hyperspectral Imagery for Assessing Yield, Biomass, Grain N Concentration, and N Output in Spring Wheat. Remote Sens. (Basel) 13, 1373. https://doi.org/10.3390/ rs13071373.
- Rodrigues, F., Blasch, G., Defourny, P., Ortiz-Monasterio, J., Schulthess, U., Zarco-Tejada, P., Taylor, J., Gérard, B., Rodrigues, F.A., Blasch, G., Defourny, P., Ortiz-Monasterio, J.I., Schulthess, U., Zarco-Tejada, P.J., Taylor, J.A., Gérard, B., 2018. Multi-Temporal and Spectral Analysis of High-Resolution Hyperspectral Airborne Imagery for Precision Agriculture: Assessment of Wheat Grain Yield and Grain Protein Content. Remote Sens. (Basel) 10, 930. https://doi.org/10.3390/ rs10060930.
- Sadras, V., Roget, D., O'Leary, G., 2002. On-farm assessment of environmental and management constraints to wheat yield and efficiency in the use of rainfall in the Mallee. Aust. J. Agr. Res. 53, 587–598. https://doi.org/10.1071/AR01150.
- Skerritt, J.H., Adams, M.L., Cook, S.E., Naglis, G., 2002. Within-field variation in wheat quality: implications for precision agricultural management. Aust. J. Agr. Res. 53, 1229–1242. https://doi.org/10.1071/ar01204.
- Sloat, L.L., Lin, M., Butler, E.E., Johnson, D., Holbrook, N.M., Huybers, P.J., Lee, J.-E., Mueller, N.D., 2021. Evaluating the benefits of chlorophyll fluorescence for in-

A. Longmire et al.

season crop productivity forecasting. Remote Sens. Environ. 260, 112478 https://doi.org/10.1016/j.rse.2021.112478.

- Stoy, P.C., Khan, A.M., Wipf, A., Silverman, N., Powell, S.L., 2022. The spatial variability of NDVI within a wheat field: Information content and implications for yield and grain protein monitoring. PLoS One 17, e0265243.
- Tan, C., Zhou, X., Zhang, P., Wang, Z., Wang, D., Guo, W., Yun, F., 2020. Predicting grain protein content of field-grown winter wheat with satellite images and partial least square algorithm. PLoS One 15, e0228500.
- Upreti, D., Huang, W., Kong, W., Pascucci, S., Pignatti, S., Zhou, X., Ye, H., Casa, R., 2019. A Comparison of Hybrid Machine Learning Algorithms for the Retrieval of Wheat Biophysical Variables from Sentinel-2. Remote Sens. (Basel) 11, 481. https:// doi.org/10.3390/rs11050481.
- van Herwaarden, A.F., Angus, J.F., Richards, R.A., Farquhar, G.D., 1998. "Haying-off", the negative grain yield response of dryland wheat to nitrogen fertiliser II. Carbohydrate and protein dynamics. Aust. J. Agr. Res. 49, 1083–1094. https://doi. org/10.1071/a97040.
- van Ittersum, M.K., Howden, S.M., Asseng, S., 2003. Sensitivity of productivity and deep drainage of wheat cropping systems in a Mediterranean environment to changes in CO2, temperature and precipitation. Agr Ecosyst Environ 97, 255–273. https://doi. org/10.1016/S0167-8809(03)00114-2.
- van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: A systematic literature review. Comput. Electron. Agric. 177, 105709 https://doi.org/10.1016/j.compag.2020.105709.
- Verhoef, W., Jia, L., Xiao, Q., Su, Z., 2007. Unified Optical-Thermal Four-Stream Radiative Transfer Theory for Homogeneous Vegetation Canopies. IEEE Trans. Geosci. Remote Sens. 45, 1808–1822. https://doi.org/10.1109/TGRS.2007.895844.
- Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G. P.W., Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. ISPRS J. Photogramm. Remote Sens. 108, 273–290. https://doi.org/10.1016/j.isprsjprs.2015.05.005.
- Walsh, O.S., Marshall, J.M., Nambi, E., Jackson, C.A., Ansah, E.O., Lamichhane, R., McClintick-Chess, J., Bautista, F., 2023. Wheat Yield and Protein Estimation with Handheld and Unmanned Aerial Vehicle-Mounted Sensors. Agronomy 13, 207. https://doi.org/10.3390/agronomy13010207.
- Wang, L., Tian, Y., Yao, X., Zhu, Y., Cao, W., 2014. Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. Field Crop Res 164, 178–188. https://doi.org/10.1016/j.fcr.2014.05.001.
- Wang, Z.J., Wang, J.H., Liu, L.Y., Huang, W.J., Zhao, C.J., Wang, C.Z., 2004. Prediction of grain protein content in winter wheat (*Triticum aestivum* L.) using plant pigment ratio (PPR). Field Crop Res 90, 311–321. https://doi.org/10.1016/j.fcr.2004.04.004.
- Wolanin, A., Camps-Valls, G., Gómez-Chova, L., Mateo-García, G., van der Tol, C., Zhang, Y., Guanter, L., 2019. Estimating crop primary productivity with Sentinel-2

and Landsat 8 using machine learning methods trained with radiative transfer simulations. Remote Sens. Environ. 225, 441–457. https://doi.org/10.1016/j. rse.2019.03.002.

- Wolanin, A., Mateo-García, G., Camps-Valls, G., Gómez-Chova, L., Meroni, M., Duveiller, G., Liangzhi, Y., Guanter, L., 2020. Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. Environ. Res. Lett. 15, 024019 https://doi.org/10.1088/1748-9326/ab68ac.
- Wright, D.L., Rasmussen, V.P., Ramsey, R.D., Baker, D.J., Ellsworth, J.W., 2004. Canopy Reflectance Estimation of Wheat Nitrogen Content for Grain Protein Management. GISci. Remote Sens. 41, 287–300. https://doi.org/10.2747/1548-1603.41.4.287.
- Wu, D., Hur, K., Xiao, Z., 2021. A GAN-Enhanced Ensemble Model for Energy Consumption Forecasting in Large Commercial Buildings. IEEE Access 9, 158820–158830. https://doi.org/10.1109/ACCESS.2021.3131185.
- Xu, X.Q., Lu, J.S., Zhang, N., Yang, T.C., He, J.Y., Yao, X., Cheng, T., Zhu, Y., Cao, W.X., Tian, Y.C., 2019. Inversion of rice canopy chlorophyll content and leaf area index based on coupling of radiative transfer and Bayesian network models. ISPRS J. Photogramm. Remote Sens. 150, 185–196. https://doi.org/10.1016/j. isprsjprs.2019.02.013.
- Xue, L.-H., Cao, W.-X., Yang, L.-Z., 2007. Predicting Grain Yield and Protein Content in Winter Wheat at Different N Supply Levels Using Canopy Reflectance Spectra. Pedosphere 17, 646–653. https://doi.org/10.1016/S1002-0160(07)60077-0.
- Zadoks, J.C., Chang, T.T., Konzak, C.F., 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.
- Zarco-Tejada, P.J., Camino, C., Beck, P.S.A., Calderon, R., Hornero, A., Hernández-Clemente, R., Kattenborn, T., Montes-Borrego, M., Susca, L., Morelli, M., Gonzalez-Dugo, V., North, P.R.J., Landa, B.B., Boscia, D., Saponari, M., Navas-Cortes, J.A., 2018. Previsual symptoms of *Xylella fastidiosa* infection revealed in spectral planttrait alterations. Nat. Plants 4, 432. https://doi.org/10.1038/s41477-018-0189-7.
- Zhao, C., Liu, L., Wang, J., Huang, W., Song, X., Li, C., 2005. Predicting grain protein content of winter wheat using remote sensing data based on nitrogen status and water stress. Int. J. Appl. Earth Obs. Geoinf. 7, 1–9. https://doi.org/10.1016/j. jag.2004.10.002.
- Zhao, H., Song, X., Yang, G., Li, Z., Zhang, D., Feng, H., 2019. Monitoring of Nitrogen and Grain Protein Content in Winter Wheat Based on Sentinel-2A Data. Remote Sens. (Basel) 11, 1724. https://doi.org/10.3390/rs11141724.
- Zhou, X., Kono, Y., Win, A., Matsui, T., Tanaka, T.S.T., 2021. 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. 24, 137–151. https://doi.org/10.1080/1343943X.2020.1819165.