Assessment of crop traits retrieved from airborne hyperspectral and thermal remote sensing imagery to predict wheat grain protein content

A.R. Longmire a, T. Poblete b, J.R. Hunt a,c, D. Chen a, P.J. Zarco-Tejada a,b,d,*

a School of Agriculture and Food, Faculty of Veterinary and Agricultural Sciences (FVAS), University of Melbourne, Melbourne 3010, Victoria, Australia
b Department of Infrastructure Engineering, Faculty of Engineering and Information Technology (FEIT), University of Melbourne, Melbourne 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:
Hyperspectral
Radiative transfer
Stress detection
Grain protein content
Machine learning
Thermal

ABSTRACT

Wheat (Triticum spp.) is among the world’s most widely grown crops and receives large quantities of nitrogen (N) fertiliser. Grain protein content (GPC) is influenced by genetic, agronomic and weather conditions affecting crops’ physiological status and stress levels; accurate GPC prediction has potential to reduce N losses and improve profit. Success in GPC estimation from remotely sensed plant traits has been limited. For progress to be made, it is necessary to robustly identify imaging spectroscopy-based physiological traits most closely associated with GPC in both experimental and commercial contexts. We present results from piloted hyperspectral flights and ground campaigns at two dryland field experiments with divergent water supply and wide-ranging N treatments, and from two years’ flights over 17 commercial fields planted to either bread (T. aestivum) or durum (T. durum) wheat, in the southern Australian wheat belt. Imagery was acquired with airborne hyperspectral and thermal sensors, with spatial resolutions of approx. 0.3 m and 0.5 m for experimental plots and 1 m/1.7 m in commercial fields. Leaf clip measurements, leaf and grain samples were collected and, in commercial fields, ~40,000 records obtained from harvester-mounted protein monitors. Crop Water Stress Index (CWSI), solar-induced fluorescence (SIF), reflectance indices and PRO4SAIL radiative transfer model inverted parameters were retrieved for each plot and GPC record location. The photochemical reflectance index (PRI) related to xanthophyll pigments was consistently associated with GPC at both leaf and canopy scale in the plots and transect. In the commercial crops, a gradient boosted machine learning algorithm (GBM) ranked CWSI as the strongest indicator of GPC under severe water stress, while SIF, PRI and inverted biochemical constituents anthocyanins and carotenoids were consistently important under more benign conditions. Structural parameters inverted from the hyperspectral reflectance imagery were not prominent except under severe drought. We attained statistically significant results estimating GPC in unseen samples, with best relationships between predicted and observed GPC of $r^2 = 0.80$ in a model built with thermal and physiological traits obtained from the hyperspectral and thermal imagery.

1. Introduction

Combined, bread (Triticum aestivum L.) and durum wheat (T. durum Desf.) provide >20% of humans’ carbohydrate (CHO) and protein needs (Shiferaw et al., 2013). Bread wheat receives ~17% of global nitrogen (N) fertiliser (Heffer and Prudhomme, 2020), but in Australia, a major wheat producer, only ~40% of fertiliser N is assimilated in the year of (N) fertiliser (Heffer and Prudhomme, 2020). This is a major expense for farmers (Monjardino et al., 2015), and has severe environmental costs (Sutton et al., 2011). Grain protein content (GPC) determines the economic and nutritional value of grain and the rheological properties of flour. While most of the N and protein ultimately translocated to the grain is already in the plant before anthesis (Giuliani et al., 2011; Lopez-Bellido et al., 2004; Masoni et al., 2007), GPC is also influenced by the amount of soil N plants can extract, especially during grain filling (Gooding et al., 2007; Jamieson and Semenov, 2000; Ottman et al., 2000). Post-anthesis, CHO are accumulated while photosynthesis proceeds, but drought or heat may depress assimilation, reducing protein dilution (Gooding et al., 2007; Jamieson and Semenov, 2000; Ottman et al., 2000).
2007). Further, early season N oversupply can exacerbate drought stress if vigorous early growth exhausts soil moisture (Angus and Fischer, 1991; van Herwaarden et al., 1998). Pre-harvest knowledge of GPC could allow forecasting for grain segregation or blending and hence access to price premiums (Apan et al., 2006), while GPC monitoring at harvest opens opportunities for analysis (Whelan et al., 2009) so far little explored. Piloted light aircraft can carry multiple sensors to capture thermal and hyperspectral images concurrently. This allows efficient operation over individual fields of up to 300 Ha and up to 6500 Ha in equivalent operation conditions on a single day; UAV payloads are insufficient for such large-scale operations particularly when advanced hyperspectral imagers and several instruments are used concurrently.

Vegetation indices (VI) from remote sensing (RS) have often been used to estimate wheat yield. The normalised difference VI (NDVI; Rouse et al., 1974), the most-used index in agriculture (Herrmann Bellido et al., 2004; Prey and Schmidhalter, 2019; Raya-Sereno et al., 2005), is unreliable where other factors intervene between biomass and yield, and lacks transferability. Other VIs have been used in yield estimation, especially those based in the red-edge which can account for chlorophyll concentration as a proxy for crop physiological status.

Grain quality assessment, in particular GPC and grain N content (GNC), is less advanced than yield assessment, but hyperspectral RS is promising as it quantifies variables with physiological links to grain quality. GPC estimations have been based on satellite (Feng et al., 2014; Wang et al., 2014; Wright et al., 2004; Zhao et al., 2005) and airborne imagery (Jensen et al., 2007; Prey and Schmidhalter, 2019; Raya-Sereno et al., 2021; Rodrigues et al., 2018), mostly in experimental plots. This typically induces GPC variability but limits the scope of conclusions by omission of commercial settings. Where tested over multiple seasons, estimates of wheat GPC, and the specific predictors most suited to doing so, have lacked consistency, but for both yield and GPC, observations around early grain filling (Zadoks et al., 1974); development stage Z65—73 appear optimal (Apan et al., 2006; Jensen et al., 2007; Lopez-Bellido et al., 2004; Prey and Schmidhalter, 2019; Raya-Sereno et al., 2021). To the best of our knowledge, the objective of assessing wheat GPC from hyperspectral RS data at commercial scale has been addressed only in a single-year study with ~200 sampling points in one 86 ha field (Rodrigues et al., 2018). We build on previous work (Rodrigues et al., 2018) and make progress focusing on the understanding of specific spectrally derived plant traits such as biochemical constituents and structural traits to evaluate their importance across experimental and commercial fields, over two years and including advanced traits such as thermal indicators of water stress and solar-induced fluorescence.

Empirical VI-based models predominate for estimating biophysical parameters including pigment concentrations, leaf N and structural aspects related to crop condition and GPC from hyperspectral data (Ferwerda et al., 2005; Herrmann et al., 2010). Narrow band hyperspectral reflectance indices (NBHI) have been used successfully in chlorophyll (C_a) and N estimation in wheat (Li et al., 2014; Wang et al., 2012); some are designed to disentangle leaf biochemistry from structural variables and water status. Many successful VIs incorporate the red edge for its strong correlations with C_a (Clevers and Gitelson, 2013; Haboudane et al., 2008; Prey and Schmidhalter, 2019). These include the normalised difference red edge index (NDRl) (Gitelson et al., 1994a, 1994b), used to retrieve C_a in wheat (Li et al., 2015). Devised to overcome canopy effects, the Transformed Chlorophyll Absorption in Reflectance Index (TCARI; Haboudane et al., 2002) normalised by the Optimised Soil-Adjusted Vegetation Index (OSAVI; Rondeaux et al., 1996)) estimated C_a (Gonzalez-Dugo et al., 2015) and total N (Klem et al., 2018) in wheat. (Zhao et al., 2020) predicted yield in commercial wheat using two indices focused in the red edge: the chlorophyll index (CI; Gitelson et al., 2003), and OSAVI. Raya-Sereno et al. (2021) found that RE indices consistently correlated better than other indices with GPC, and that the NDRE was a relatively stable indicator of GNC over four years. Similarly, Li et al. (2020) obtained the best of moderate results estimating GPC with the double-peak canopy N index (DCNI; Chen et al., 2010) and CI. In contrast, in a study estimating GPC from airborne multispectral data, greater importance was attributed to the red and NIR domains, than to the RE (Zhou et al., 2021). Rodrigues et al. (2018) found that combinations of green- and RE bands performed best.

Indices also target anthocyanins (Anth) and carotenoids (C_a+c), whose stress responses suggest them as indicators of GPC. Anth are upregulated under water stress (Chalker-Scott, 2002; Naing and Kim, 2021), can be high in senescing leaves (Sims and Gamon, 2002), and post-anthesis drought increases their accumulation in grain (Li et al., 2018). C_a+c are produced in response to photooxidative, water and heat stress (Groth et al., 2020; Janeczko et al., 2018). The photochemical reflectance index (PRI; Gamon et al., 1992)) is sensitive to C_a+c and reacts to instantaneous changes in water stress and photosynthetic rate (Feng et al., 2017; Magney et al., 2016). PRI is a pre-visual indicator of water stress and recovery and is proposed as an alternative to thermal RS for water stress detection (Kobayashi et al., 2021; Suarez et al., 2008). PRI has been used to improve yield estimates and discriminate water and disease stress in wheat (Feng et al., 2017; Magney et al., 2014). Alongside PRI, the carotenoid index CAR (Zarco-Tejada et al., 2013b) has shown moderately strong relationships with yield across rainfed and irrigated bread and durum wheat (Gonzalez-Dugo et al., 2015). Despite the links between stress and GPC and use of CAR and PRI in stress detection, neither index seems to have been tested as a predictor of GPC in wheat.

As an alternative to VIs, whose empirical relationships with crop performance measures may vary across seasons and locations, parameters retrieved by radiative transfer models (RTM) are more robust across location, phenological stage and crop type (Clevers and Kooistra, 2012; Dorigo et al., 2007; Jacquemoud et al., 1995). By linking leaf and canopy models, biochemical and structural parameters at each level can be estimated concurrently. Quantities directly associated with yield and/or GPC, including C_a, and LAI, and others that change more dynamically with stress such as Anth and C_a+c, have been quantified by RTM inversion in wheat under water/N stress (Botha et al., 2010; Camino et al., 2018). Moreover, retrieval accuracies for PROSPECT and SAIL have been shown since early in RTM development (Bacour et al., 2002; Fétet et al., 2008; Jacquemoud et al., 2009, 1995; Li et al., 2015; Ustin et al., 2009).

Solar-induced fluorescence (SIF) is a proxy for the functional status of vegetation, instantaneous photosynthetic rate and assimilation (Genty et al., 1989; Meroni et al., 2009; Mohammed et al., 2019). SIF shows diurnal and seasonal variations in photosynthetic rate and is key to stress diagnosis (Poblete et al., 2020; Zarco-Tejada et al., 2018, 2016, 2013a). In bread and durum wheat, SIF has been combined with C_a and leaf structural traits to diagnose N deficiency and water stress (Camino et al., 2018). SIF variability should therefore parallel relative CHO availability during grain filling in wheat. While C_a+c SIF, and PRI are sensitive to short-term changes induced by stress (Gamon et al., 1992; Penuelas et al., 1994), C_a+c and structural measures related to LAI reflect more temporally stable canopy traits with influence on GPC.

Reduced stomatal conductance lowers leaf cooling and assimilation, making thermal data useful for tracking the effects of water stress (Gonzalez-Dugo et al., 2015; Grant et al., 2007; Idso, 1982). High temperatures and drought during grain filling correlate with higher GPC (Camino et al., 2015). (GNC) is less advanced than yield assessment, but hyperspectral RS is promising for such large-scale operations particularly when advanced hyperspectral imagers and several instruments are used concurrently.

Gradient boosting/boosted machines (GBM) are a machine learning algorithm based on work by Friedman (2001). With good predictive skill, robustness to multicollinearity and the ability to deal with large
numbers of both input features and observations in a computationally efficient way (Ruan et al., 2022), GBMs assess input features’ importance in terms of the gain they bring to a model (Grinberg et al., 2020). GBM use with RS data has included estimation of leaf N (Yang et al., 2021) and yield (Cao et al., 2020; Ruan et al., 2022) in wheat, but GPC estimation to date has been based solely on genotype data (Grinberg et al., 2020). To advance GPC estimation from RS, the variables most closely associated with it should be identified and their stability across cropping situations assessed; tree-based algorithms such as GBM and RF have the advantages both of offering feature importance assessment, the primary focus of our work, and of being among the most accurate ML algorithms (Abdi, 2020; Ruan et al., 2022; Zhang et al., 2020).

2. Materials and methods

2.1. Study sites

We consider experimental plots and commercial crops, grown under rainfed conditions at three locations in the southern Australian wheat belt with a variety of climatic regimes and soils. The plot trials were located at Birchip (site 1; 35.969°S, 142.822°E; altitude 102 m; average annual rainfall (AAR) = 353 mm) and Yarrawonga (site 2; 36.050°S, 145.983°E; altitude 129 m; AAR = 470 mm; Fig. 1) and were planted in randomised complete block designs with hard white bread wheat (cv. Scepter) for N fertiliser treatment trials in 2019 (Fig. 1). Site 1 takes the Köppen-Geiger climate classification Bsk and site 2, Cfa (Peel et al., 2007). Soils are classified as calcarosol and sodosols, respectively, based on the Australian Soil Classification (ASC; Isbell, 2002). Site 1 was sown on 2019-05-16 and N rates were adjusted according to in-season rainfall to target yield modelled with the Yield Prophet® decision support tool (Hochman et al., 2009; Hunt et al., 2006) and applied on 2019-08-10 at Z31. This produced small increments in N rates, some of which were grouped for analysis (Table 1). Site 2 was sown on 2019-05-09 and fertiliser was applied in equal doses at Z23 and Z31. All plots were approx. 12 × 2 m, treatments were replicated four times and agronomic procedures were equivalent.

The commercial crops included hard white bread and durum wheat cultivars in 17 fields across 30 km of latitude around Kaniva (36.37°S, 141.24°E; altitude 142 m; AAR = 451 mm; Fig. 2a). There, the dominant soils are vertosols, but both sodosols and chromosols are also common; the Köppen-Geiger climate classification for Kaniva is Cfb. Seven fields totalling 815 ha of bread and durum wheat were sown between 2019-05-15 and 2019-06-03; similar crops were sown from 2019-05-12–2019-06-05 across 10 fields (1039 ha). Fertiliser was applied 1–3 times each season, usually as urea. Dualex and SpectraPen leaf clip observations and soil samples were taken in a transect across one field, M01 (36.30°S 141.35°E), concurrent with the 2020 flight. Absent differential fertiliser treatments in the crops, K-means clustering (Hartigan and Wong, 1979) by GPC was used to divide the transect data into three levels for analysis.

2.2. Ground data collection and laboratory processes

Dualex (FORCE-A, Orsay, France) and SpectraPen (PSI, Drasov, Czech Republic) measurements were made on the adaxial surface of

Table 1
<table>
<thead>
<tr>
<th>Location</th>
<th>Soil N (mg kg⁻¹)</th>
<th>Total fertiliser (kg N/ha)</th>
<th>Treatment (aggregated)</th>
<th>Plots (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>46.8</td>
<td>0</td>
<td>B0</td>
<td>4</td>
</tr>
<tr>
<td>Birchip</td>
<td>30</td>
<td>B1</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>98</td>
<td>B2</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>162</td>
<td>B3</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>167</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>171</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Site 2</td>
<td>46.1</td>
<td>0</td>
<td>Y1</td>
<td>4</td>
</tr>
<tr>
<td>Yarrawonga</td>
<td>46</td>
<td>Y2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>92</td>
<td>Y3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>138</td>
<td>Y4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>184</td>
<td>Y5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

141.24°E; altitude 142 m; AAR = 451 mm; Fig. 2a). There, the dominant soils are vertosols, but both sodosols and chromosols are also common; the Köppen-Geiger climate classification for Kaniva is Cfb. Seven fields totalling 815 ha of bread and durum wheat were sown between 2019-05-15 and 2019-06-03; similar crops were sown from 2019-05-12–2019-06-05 across 10 fields (1039 ha). Fertiliser was applied 1–3 times each season, usually as urea. Dualex and SpectraPen leaf clip observations and soil samples were taken in a transect across one field, M01 (36.30°S 141.35°E), concurrent with the 2020 flight. Absent differential fertiliser treatments in the crops, K-means clustering (Hartigan and Wong, 1979) by GPC was used to divide the transect data into three levels for analysis.

2.2. Ground data collection and laboratory processes

Dualex (FORCE-A, Orsay, France) and SpectraPen (PSI, Drasov, Czech Republic) measurements were made on the adaxial surface of
upper sun-adapted leaves central in plots. Ten measurements per plot were taken with each instrument, quasi-concurrently with flights and in equivalent light and meteorological conditions. Approximately 80 g of entire flag or other upper, sun-exposed leaves were cut from the central area of plots, sealed in plastic, refrigerated in transit then kept at 20 °C until processing. Ten subsamples were taken from the central area of these cut leaves with leaf tissue punches of known diameter so that areal N content could be calculated; these discs were weighed before and after drying. Leaf tissue was dried at 65 °C for 48 h, then ground to a uniform powder in a ball mill and analysed by Dumas combustion for total N (mass %). Plots were machine harvested and the grain assessed for protein content by near infrared (NIR) spectroscopy (CropScan 3000B Grain Analyser, Next Instruments, Sydney Australia). Ambient temperature, barometric pressure and incoming shortwave and longwave radiation were recorded on the ground using a portable weather station (model WXT510, Vaisala, Helsinki, Finland). During commercial harvesting, NIR spectrometers (CropScan 3000/3300H, Next Instruments), mounted on combine harvesters, collected GPC with location (±0.01 m) from real-time kinetic GPS. Dualex and SpectraPen measurements were also made during a transect across one field at Miram (M01) in 2020. Three soil samples from the top 15 cm were taken with a hand auger and mixed to represent each location, which was marked with a hand-held GPS (Garmin, Olathe, Kansas, USA). Samples were oven-dried at 65 °C for 48 h and mineral N extracted as per Rayment and Lyons (2010) on a Skalar San++ SFA (FlowAccess V 3.2).

2.3. Airborne data collection

Airborne hyperspectral and thermal images were collected by sensors flown on a light aircraft over plots at site 1 on 2019–10-03 (1409 days after sowing (DDAS)), site 2 on 2019-10-09 (1559 DDAS), and commercial fields near Kaniva on 2019-10-22 (bread wheat 1514 DDAS, durum 1736 DDAS) and 2020-10-28 (bread 1592 DDAS, durum 1742 DDAS). Thermal time, and ambient temperature for CWSI calculation, were based on data from the Australian Bureau of Meteorology’s BoM recording station #78015 (Bureau of Meteorology, 2021). The hyperspectral data were collected in the visible and near infrared (VNIR) domains with a hyperspectral VNIR sensor (VNIR E-Series model; Headwall Photonics, Fitchburg, MA, USA), capturing 371 bands from 400 to 1001 nm at 8 nm per pixel, yielding 7 nm FWHM with a 25 µm slit. At 12-bit radiometric resolution, storage rate was 50 frames per second with exposure time of 18 ms and an 8 mm focal length. Radiometric and spectral calibration was completed in the laboratory prior to flights. 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 time of flight with a Micro-Tops II sunphotometer (Solar LIGHT Co., Philadelphia, PA, USA). This method has previously been implemented for hyperspectral data (Calderón et al., 2015; Poblete et al., 2020; Zarco-Tejada et al., 2018). Orthorectification was performed using Parametric Geocoding & Orthorectification for Airborne Optical Scanner Data (PARGe; ReSe applications GmbH, Wil, Switzerland) using the IMU and GPS flight data obtained from a VN-300 (VectorNav Technologies LLC, Dallas, TX, USA). Thermal images were collected in the 7.5–14 µm region with an A655c camera (FLIR systems, Wilsonville, Oregon, USA). Over commercial fields, images were acquired at ~2000 m above ground level (AGL), yielding pixels of 1.0 m (hyperspectral) and 1.7 m (thermal) ground sampling distance (GSD). Images acquired at 350 m and 400 m AGL gave GSD = 0.2 m at site 2, respectively. Radiance (R; Fig. 2b) and thermal (canopy temperature; T_c; Fig. 2c) values were aggregated to pixel mean per plot. Mean L and R spectra by fertiliser treatment level are shown for both plot sites in Fig. 3.

The crop water stress index (CWSI) was calculated according to Idso et al. (1981), normalising canopy temperature (T_c) with air temperature (T_a) and vapour pressure deficit (VPD).
CWSI = \frac{(T_c - T_a) - (T_c - T_a)_{UL}}{(T_c - T_a)_{UL} - (T_c - T_a)_{LL}} \quad (1)

The lower limit \((T_c - T_a)_{UL}\) represents the canopy/air temperature differential in a canopy transpiring at its maximum potential rate for a given VPD while the upper limit \((T_c - T_a)_{UL}\) represents the same for a canopy in which the transpiration flux is zero. The lower limit used here was defined by Idso (1982) and adopted by Gonzalez-Dugo et al. (2015).

\[(T_c - T_a)_{LL} = -3.25 \cdot \text{VPD} + 3.38 \quad (2)\]

2.4. Data processing

After their collection, the datasets were combined at each level’s primary study unit: plots, transect waypoints and regions of interest (ROI). Dualex data from the plots and transect were screened for outliers ≥2 standard deviations from the plot or waypoint mean, and for erroneous SpectraPen spectra, then aggregated to mean values and spectra per study unit. Data analysis for commercial fields was based on the geolocated GPC records collected during harvest. A 100 m2 ROI was established around each GPC point by buffering to a 5 m radius, then drawing bounding geometry for each point, and the GPC value adopted for the ROI (Fig 2d). The 5 m radius was chosen so that ROI width was less than the harvester swath width (12 m) and to increase spatial independence between ROIs. Areas within 20 m of perimeter fences, dams, trees and cloud shadow, and all ROIs intersecting these, were excluded (Fig. 2b–d). A Wilcoxon test (Bauer, 1972) was applied to assess the significance and effect size of differences between GPC observations across years and wheat types. Mean L and R spectra and Tc values were calculated from image pixels contained within each ROI; for L and R, n ≈ 100 pixels, Tc n ≈ 36 pixels. NBHI, inverted leaf and canopy parameters, solar-induced fluorescence (SIF) and CWSI were retrieved from airborne R, L, and Tc, respectively for all plots and ROIs. NBHI were also calculated from SpectraPen R at plot and transect scales. Where transect waypoints intersected with ROIs, airborne GPC, NBHI, inverted parameters, SIF and CWSI values retrieved for the ROI were assigned to waypoints. The NBHI relevant to later procedures are detailed in Table 2. Spatial analysis was done in QGIS (QGIS Development Team, 2020) and R (R Core Team, 2020).

Analogously to Poblete et al. (2021), inverted leaf and canopy parameters were retrieved with the PROSPECT-D (Féret et al., 2017) and 4SAILH (Verhoef et al., 2007) radiative transfer models (RTM), linked as PRO4SAIL. PROSPECT-D was used to retrieve the leaf pigments \(C_{a+b}\) and Anth, while 4SAILH was used to estimate canopy structural traits LAI and leaf inclination distribution function (LIDFa) from mean plot and ROI spectra. RTM parameters were randomly sampled from uniform distributions in the ranges given in Table 3 to construct a look-up table (LUT) of 200,000 reflectance spectra simulations with their

Fig. 3. Mean radiance (W/sr m⁻² nm⁻¹) and reflectance spectra captured by airborne hyperspectral sensors at Birchip (site 1; a, b) and Yarrawonga (site 2; c, d) by treatment.
associated leaf and canopy values. To compare synthetic and airborne spectra, the LUT was interrogated using support vector machine (SVM) algorithms run in MATLAB (MATLAB; Statistics and Machine Learning, Deep Learning and Parallel Computing toolboxes; Mathworks Inc., Natick, MA, USA), applying a radial basis function. Hyperparameters were optimised for each target variable during training, and inversions took simulated reflectance spectra as SVM inputs and plant/canopy characters as outputs.

Fluorescence was retrieved using bands inside and outside the O$_2$–A Fraunhofer line (FLD2; Plascyk and Gabriel, 1975) and based on irradiance simulated with the SMARTS software (Guymard, 1995), then convolved to the FWHM and spectral sampling interval (SSI) of the hyperspectral sensor. The inside band (762 nm) is the local minimum incoming irradiance at the relevant SSI while the outside band (750 nm) is on the shoulder of the O$_2$–A line.

To evaluate collinearity between inputs and reduce dimensionality, each year’s data were assessed by variance inflation factor analysis (VIF; R package vif; (Nakazawa, 2022)) with threshold $\delta = 5$ (Akimwande et al., 2015). Multicollinearity features were first excluded among NBH; those surviving (VIF $\leq 5$) then underwent a second VIP analysis with the inverted parameters, CWSI and SIF. Variables thus chosen were kept only if they improved the ML model; these final input features were categorised into three layers: physiological (Anth, $c_{a,b}$, $c_{b,c}$, SIF, PRI); structural (EVI$_{GSPc}$ hereafter ‘EVI’, LAI, LIDFs); and thermal (CWSI).

2.5. Algorithms for GPC assessment

A gradient boosted machine (GBM) machine learning (ML) algorithm was used to estimate, through supervised learning, relationships between leaf and canopy traits and the target variable GPC. Our primary objective was to assess input features’ relative importance to GPC estimation (Chen and Guestrin, 2016). Feature importance, and the performance of models containing them, was assessed stepwise: physiological + structural + thermal (Model 1); physiological + structural (Model 2); physiological (Model 2), as proposed by Friedman and Meulman (2003).

The GBM used decision trees as its base learners, sub-models with weak predictive skill which learn from their predecessors’ error in estimating the target variable to iteratively improve the estimate (Chen and Guestrin, 2016). Datasets from each site were randomly split 70:30 into training and test sets at each model run and passed to a linear function. Four model hyperparameters were varied across either three or four levels: learning rate, tree depth, minimum node size and stochastic gradient descent (SGD), and this space was searched by full factorial sampling. In SGD, each iteration runs on a randomly sampled subset of rows, introducing noise and improving robustness to overfitting. Randomised K-fold ($K = 5$) cross-validation (CV) was implemented, also to reduce overfitting. The combination of hyperparameters that minimised root mean square error (RMSE) was adopted as the top model for each training set, then applied to predict GPC in the relevant test set. Data analysis and ML were done in R (R Core Team, 2020) using the packages xgbost (Chen et al., 2021) for gradient boosting and caret (Kuhn, 2020) for model tuning.

3. Results

3.1. Plot experiments

The 2019 growing season rainfall (GSR) at experimental site 1 was 65% of long-term AAR, but with good soil moisture from 2018, while at site 2 GSR was 54% of AAR with low starting soil moisture. A strong gradient in GPC was seen at both sites, parallel with N dosing; GPC saturated under high N and was higher overall at the droughted site 2 (Figs. 4 and 5b).

Several leaf-level indicators were associated with higher GPC along the N treatment gradient at both sites 1 and 2. The NPCI, VOG1, ZMI, R$_{920}$/R$_{729}$ and PRI$_{23}$ (Hernández-Clemente et al., 2011) increased with fertiliser N rate (Fig. 5a, b) although trends with GPC were less distinct at site 2. At image level, structural indices NDVI and EVI increased in

Table 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Abbreviation</th>
<th>Value/Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorophyll $a + b$ content [μg/cm$^2$]</td>
<td>$c_{a+b}$</td>
<td>3–70</td>
</tr>
<tr>
<td>Carotenoid content [μg/cm$^2$]</td>
<td>$c_{ab}$</td>
<td>1–20</td>
</tr>
<tr>
<td>Anthocyanin content [μg/cm$^2$]</td>
<td>Anth</td>
<td>1–10</td>
</tr>
<tr>
<td>Dry matter content [g/cm$^2$]</td>
<td>$c_d$</td>
<td>0.001–0.035</td>
</tr>
<tr>
<td>Water content [g/cm$^2$]</td>
<td>$c_w$</td>
<td>0.001–0.035</td>
</tr>
<tr>
<td>Mesophyll struct. Coef.</td>
<td>N</td>
<td>0.5–3.0</td>
</tr>
<tr>
<td>Leaf area index [m$^2$/m$^2$]</td>
<td>LAI</td>
<td>1–5</td>
</tr>
<tr>
<td>Leaf Inclination Dist. Func. ['']</td>
<td>LIDFs</td>
<td>0–90</td>
</tr>
<tr>
<td>Hot spot parameter</td>
<td>$h$</td>
<td>0.01</td>
</tr>
<tr>
<td>Observer angle [deg.]</td>
<td>$\theta_o$</td>
<td>0</td>
</tr>
<tr>
<td>Sun zenith angle [deg.]</td>
<td>$\theta_s$</td>
<td>47.7</td>
</tr>
<tr>
<td>Relative azimuth angle [deg.]</td>
<td>$\psi$</td>
<td>0</td>
</tr>
</tbody>
</table>
line with N treatment and GPC but saturated after the first treatment level at site 2 and after the second at site 1. The CI, PRI\textsubscript{m4} and, as at leaf level, R\textsubscript{920}/R\textsubscript{729} trended with GPC. Agreement between PRI\textsubscript{m3} (leaf) and PRI\textsubscript{m4} (canopy) was close, and leaf-level readings were lower at both sites (Fig. 5a–d). Each PRI\textsubscript{m*} tracked as well with GPC as with leaf N from optical and, where conducted, destructive sampling (Fig. 5e, f). Associations between image-level indicators and GPC were weaker at site 2 (Fig. 5c, d). Dualex observations were similar across sites: C\textsubscript{a}+b and the Nitrogen Balance Index (NBI) had positive associations with GPC, with higher absolute readings at site 1, while Anth were inverted with respect to GPC. Leaf N from destructive sampling generally increased with N treatment at both sites, but like GPC declined in treatment Y4.

Among parameters retrieved by RTM inversion, C\textsubscript{a}+b trended with GPC at both sites, though with saturation at higher values, particularly at site 2 where C\textsubscript{a}+b concentration was lower (Fig. 6a, b). Inverted C\textsubscript{a}+b was significantly correlated with Dualex C\textsubscript{a}+b (r\textsuperscript{2} = 0.61, p < 0.0001). C\textsubscript{k+c} showed a minor trend parallel to GPC at site 2 but saturated after one N treatment and was lower range than at site 1, where no trend was seen. Anth trended higher with N treatment, and were in higher concentration overall at site 1, but this relationship was inverted at site 2. LAI trended higher with N treatment at site 1, while at site 2 the opposite was seen, with saturation, and LAI values were very high. No alignment between SIF observations and treatment was seen at site 1, while SIF was in a substantially higher range at site 2. The following relationships at site 1 were statistically significant: Anth (R = 0.61, p < 0.01), C\textsubscript{a}+b (R = 0.63, p < 0.01) and CWSI (R = −0.52, p < 0.05). At site 2 these were: Anth (R = −0.62, p < 0.01) and LAI (R = −0.55, p < 0.05).

3.2. Field transect

The range of GPC at M01 was intermediate between sites 1 and 2, the crop had not senesced as much as at site 2 and retained soil moisture during our campaigns. At leaf level, VOG1, ZMI, R\textsubscript{920}/R\textsubscript{729} (not shown) and PRI\textsubscript{m4} (canopy) was close, and leaf-level readings were lower at both sites (Fig. 5a–d). Each PRI\textsubscript{m*} tracked as well with GPC as with leaf N from optical and, where conducted, destructive sampling (Fig. 5e, f). Associations between image-level indicators and GPC were weaker at site 2 (Fig. 5c, d). Dualex observations were similar across sites: C\textsubscript{a}+b and the Nitrogen Balance Index (NBI) had positive associations with GPC, with higher absolute readings at site 1, while Anth were inverted with respect to GPC. Leaf N from destructive sampling generally increased with N treatment at both sites, but like GPC declined in treatment Y4.

Among parameters retrieved by RTM inversion, C\textsubscript{a}+b trended with GPC at both sites, though with saturation at higher values, particularly at site 2 where C\textsubscript{a}+b concentration was lower (Fig. 6a, b). Inverted C\textsubscript{a}+b was significantly correlated with Dualex C\textsubscript{a}+b (r\textsuperscript{2} = 0.61, p < 0.0001). C\textsubscript{k+c} showed a minor trend parallel to GPC at site 2 but saturated after one N treatment and was lower range than at site 1, where no trend was seen. Anth trended higher with N treatment, and were in higher concentration overall at site 1, but this relationship was inverted at site 2. LAI trended higher with N treatment at site 1, while at site 2 the opposite was seen, with saturation, and LAI values were very high. No alignment between SIF observations and treatment was seen at site 1, while SIF was in a substantially higher range at site 2. The following relationships at site 1 were statistically significant: Anth (R = 0.61, p < 0.01), C\textsubscript{a}+b (R = 0.63, p < 0.01) and CWSI (R = −0.52, p < 0.05). At site 2 these were: Anth (R = −0.62, p < 0.01) and LAI (R = −0.55, p < 0.05).

3.2. Field transect

The range of GPC at M01 was intermediate between sites 1 and 2, the crop had not senesced as much as at site 2 and retained soil moisture during our campaigns. At leaf level, VOG1, ZMI, R\textsubscript{920}/R\textsubscript{729} (not shown) and PRI\textsubscript{m4} (canopy) was close, and leaf-level readings were lower at both sites (Fig. 5a–d). Each PRI\textsubscript{m*} tracked as well with GPC as with leaf N from optical and, where conducted, destructive sampling (Fig. 5e, f). Associations between image-level indicators and GPC were weaker at site 2 (Fig. 5c, d). Dualex observations were similar across sites: C\textsubscript{a}+b and the Nitrogen Balance Index (NBI) had positive associations with GPC, with higher absolute readings at site 1, while Anth were inverted with respect to GPC. Leaf N from destructive sampling generally increased with N treatment at both sites, but like GPC declined in treatment Y4.

Among parameters retrieved by RTM inversion, C\textsubscript{a}+b trended with GPC at both sites, though with saturation at higher values, particularly at site 2 where C\textsubscript{a}+b concentration was lower (Fig. 6a, b). Inverted C\textsubscript{a}+b was significantly correlated with Dualex C\textsubscript{a}+b (r\textsuperscript{2} = 0.61, p < 0.0001). C\textsubscript{k+c} showed a minor trend parallel to GPC at site 2 but saturated after one N treatment and was lower range than at site 1, where no trend was seen. Anth trended higher with N treatment, and were in higher concentration overall at site 1, but this relationship was inverted at site 2. LAI trended higher with N treatment at site 1, while at site 2 the opposite was seen, with saturation, and LAI values were very high. No alignment between SIF observations and treatment was seen at site 1, while SIF was in a substantially higher range at site 2. The following relationships at site 1 were statistically significant: Anth (R = 0.61, p < 0.01), C\textsubscript{a}+b (R = 0.63, p < 0.01) and CWSI (R = −0.52, p < 0.05). At site 2 these were: Anth (R = −0.62, p < 0.01) and LAI (R = −0.55, p < 0.05).

3.3. Commercial fields

A large difference in growing conditions, especially low rainfall in 2019 (280 mm) compared to 2020 (443 mm), also affected the commercial fields (Bureau of Meteorology 2021). Rainfall from December 2018 until sowing in early May 2019 was also very low (87 mm) compared to the subsequent equivalent period (164 mm). Weather conditions changed suddenly around anthesis 2019, whereby frost was recorded on 9 October, and daily maxima of >35° C were recorded in mid-late October. Flights over the 2019 bread wheat crop took place 1514 DDAS in a year in the lowest decile of long-term AAR: the crop was under severe water stress. Remote sensing of bread wheat was done 1592 DDAS in 2020, a year of rainfall at the long-term AAR. Data capture for durum wheat was 1736 and 1742 DDAS in the respective years. Such conditions, especially the moisture contrast between soil types and years, can have large effects on grain protein.

Mean commercial bread wheat GPC was higher in 2019 (mean =
11.6, SD = 1.52) than 2020 (mean = 11.3, SD = 1.05; Wilcoxon’s p < 0.0001, effect size r = 0.489). Mean durum wheat GPC was higher in 2020 (mean = 12.7, SD = 0.94) than in 2019 (mean = 11.9, SD = 1.18; Wilcoxon’s p < 0.0001, r = 0.360). GPC was significantly higher in durum than bread wheat in each year also (2019: p < 0.0001, r = 0.112; 2020: p < 0.0001, r = 0.564). Wide GPC variability was also seen between and within fields of each product type (Fig. 8). In a subset of our data, PRO4SAIL C\(_a\)+\(_b\) and CI were strongly correlated (\(r^2 = 0.86, p < 0.0001\)).

Two reflectance indices, EVI and PRI, were both a) selected after VIF and b) robust across the two years of the study; these were retained as ML input features, supporting our finding of their correlations with GPC (Section 3.1). The relative importance of features to GPC estimation was quantified for input layers comprising: (i) physiological indicators Anth, C\(_a\)+\(_b\), and NBI in Dualex proprietary units; leaf N in mg N cm\(^{-2}\).

---

Fig. 5. Reflectance indices at leaf (a, d) and canopy level (b, e) and ground-truth indicators (c, f) by N fertiliser treatment at sites 1 (upper) and 2 (lower). At site 1, n = 20 for ground observations (a) and n = 36 for airborne indices (c); at site 2, n = 19. Anth, C\(_a\)+\(_b\), and NBI in Dualex proprietary units; leaf N in mg N cm\(^{-2}\).

Fig. 6. Chlorophyll a + b, Carotenes, Anthocyanins and LAI (C\(_a\)+\(_b\), C\(_x\)+\(_c\), Anth; μg cm\(^{-2}\)), retrieved by physical model inversion, plus solar-induced fluorescence (SIF; mW m\(^{-2}\) nm sr) and crop water stress index (CWSI) from hyperspectral and thermal data at sites 1 (n = 36) and 2 (n = 19).
3.4. Model predictive skill

The skill of each model in predicting GPC in an unseen 30% of observations was assessed for each combination of year/product/input layer by coefficient of determination ($r^2$) and relative root mean square error ($rRMSE$; GPC (%); Table 4). In every case, more information gave better predictions: Model 1, with all three layers of input data, outperformed Models 2 and 3. The best GPC prediction was seen in the 2019 bread wheat crop ($r^2 = 0.80$, $rRMSE = 0.62$), when CWSI contributed a mean 69% of total predictive power. However, this year and crop also had the highest $r^2$ when models 2 or 3 were used. In less water-stressed scenarios, the contribution of CWSI relative to total model skill was lower, confirming the tested physiological quantities as important indicators of GPC. Added to the physiological layer, the structural layer, comprising LAI, LIDFa and EVI, also increased model skill by between 11 and 21% despite their low importance rankings. This was higher than the respective CWSI contribution in any year/product/model combination.

4. Discussion

4.1. Plot experiments

Few studies have focused on GPC estimation using airborne hyperspectral remote sensing. Our objective was to identify traits related to harvest GPC in bread and durum wheat in experimental and commercial settings. High GPC is common in conditions unfavourable for CHO production and translocation, observations verified through our focus on stress indicators. At site 2, N supply at the upper treatment levels manifestly exceeded plant requirements given the low starting soil moisture and rainfall. This N excess is shown by the high LAI, which was extreme for Australian wheat (Waldner et al., 2019) and far exceeded LAI at site 1 (Fig. 6). Similarly, high GPC and declining LAI and leaf N, especially at high fertiliser rates, are signs of excessive N at site 2 but not site 1 (Figs. 5 and 6). While a greater GPC response to fertiliser is expected when grain filling is water limited (Angus and Fischer, 1991; Holford et al., 1992), and is seen in the comparison of our two sites, the site 2 GPC response saturated at heavy N applications, and at the highest level was below the next highest N treatment. Site 2 was also more advanced at the time of flights, and little extractable soil moisture remained: at the highest N treatment levels, high biomass worsened water stress, likely further restricting N uptake. Evidence for this is seen in declining leaf N and GPC in treatments Y3 and Y4, not seen in the more moderate moisture and rainfall. This N excess is shown by the high LAI, which was extreme for Australian wheat (Waldner et al., 2019) and far exceeded LAI at site 1 (Figs. 5 and 6). While a greater GPC response to fertiliser is expected when grain filling is water limited (Angus and Fischer, 1991; Holford et al., 1992), and is seen in the comparison of our two sites, the site 2 GPC response saturated at heavy N applications, and at the highest level was below the next highest N treatment. Site 2 was also more advanced at the time of flights, and little extractable soil moisture remained: at the highest N treatment levels, high biomass worsened water stress, likely further restricting N uptake. Evidence for this is seen in declining leaf N and GPC in treatments Y3 and Y4, not seen in the more moderate moisture and calibrated N dosing of site 1. The close alignment of CI, VOG1 and ZMI with $C_{a+b}$ and NBI at leaf level, inverted canopy $C_{a+b}$, and GPC supports findings of these indices’ links with leaf N (Ustin et al., 2009; Vogelmann et al., 1993; Zarco-Tejada et al., 2001), and GPC (Li et al., 2020). These associations were more pronounced at site 1 than at site 2. The VOG1 index correlated strongly with leaf N at site 1 and with $C_{a+b}$, NBI and GPC at both sites. This, and close associations between NPCI and leaf N, support other findings in wheat (Ranjan et al., 2012), although NPCI was less closely associated with GPC. Further, higher ranges of CI, VOG1, ZMI and leaf- and canopy-level $C_{a+b}$ at site 1, but higher GPC at site 2, suggest more advanced translocation at site 2, driven by senescence. This latter hypothesis is also supported by the reduction in LAI along the N gradient at site 2.

While most traits and indices discussed above are stable over days to weeks, the PRI family of indices can change on much shorter timescales; however, given the stable weather, here PRI likely shows stable stresses. Especially at site 2, evidence for the physiological link between PRI and photosynthesis was seen in the similarity between PRI and GPC, via a structural indicators lowest. Anth and $C_{a+b}$ ranked highly in 2019 but declined in 2020, while PRI and LIDFa had higher importance in 2020. In Model 3 (Fig. 9f), SIF retained top rank across years while $C_{a+b}$ was lower in 2020 than 2019 and PRI higher.
clear inverse relationship between airborne SIF and GPC, and in lower overall SIF. Our findings accord with the inverse PRI \sim SIF relationship seen by Magney et al. (2016), Penuelas et al. (1994) and Suarez et al. (2008).

Although no conclusive CWSI trend was seen at either site, its higher range at site 2 aligned with that site’s lower SIF: SIF also declined at higher N treatments, likely due to greater drought stress. Seen together, high GPC, distinct alignments in the PRI \sim GPC and CWSI \sim SIF responses show strongly constrained assimilation at site 2 but not at site 1. This was also seen in the C_{x+c} response, which despite saturation was also associated with GPC. However C_{x+c} were also parallel with leaf N and C_{a+b}, so their response may simply reflect N nutrition, as seen by Shah et al. (2017) in wheat. At image capture, C_{x+c} were lower overall at site 2, likely due to prior remobilisation to the grain. Low C_{a+b} values at site 2, compared with site 1 and other work (Hamblin et al., 2014), also suggest that C_{a+b} declined before C_{x+c} (Gitelson and Merzlyak, 1994a, 1994b). Given Anth upregulation under stress, one may expect high Anth to correspond to high GPC, but this was seen only in canopy Anth at site 1, which aligned with both leaf N and GPC. Alternatively, these observations may show a positive Anth response to N fertiliser (Yamuangmorn et al., 2018). At site 2, similar Anth responses at leaf and canopy levels, and their inversion with respect to N treatment and GPC, suggest that like C_{a+b} and C_{x+c}, Anth translocation had begun before data collection, especially in the high N treatments. Assessed on C_{x+c} and Anth, and if these compounds’ concentrations are taken as a function of stress, site 1 appears more stressed. But C_{x+c} and Anth also respond positively to high N supply, and this suggests more accessible soil N at site 1 due to higher soil moisture. At site 2, strong declines in Anth at high N appear to confirm lower soil N accessibility at high N treatments and/or earlier senescence. NDVI and EVI are inferior to the indices discussed above as stress indicators, as seen previously (Gamon et al., 1992; Penuelas et al., 1994) and showed little association with GPC. The consistency of the R_{020/0729} index with C_{a+b} NBI, and GPC suggests it could be applied in GPC estimation, due to its component bands’ sensitivity to water, structure and C_{a+b}.

4.2. Field transect

As at site 1, inverted canopy Anth increased with GPC in the field transect. Also as at both plot sites, C_{x+c} was parallel with GPC, and with soil N in M01, supporting an effect of N nutrition, stress or both on GPC. The association of PRI with lower assimilation is evident at M01 through its alignment with GPC across the transect. Given that wheat in field M01 was at least as advanced in phenology during RS campaigns as the plot sites were, its substantially lower CWSI likely shows a better match of N supply to soil moisture than at either plot site, reflecting the contrasting experimental and commercial objectives.

4.3. Commercial fields

The influence of heterogeneous soil moisture and N availability is strong and operates through mechanisms discernible in the leaf and canopy traits we retrieved. Our division of input features into thermal, structural and physiological layers, and their sequential removal from the model, allowed us to assess each input’s contribution to GPC estimation, and each layer’s influence on predictive skill. In each situation, especially high water stress, CWSI was an important indicator of GPC via its relationship with assimilation, but physiological features also showed commonalities across moisture conditions. In severe stress, in 2019 bread wheat, physiological components contributed little. With stomatal conductance and photosynthesis universally depressed, the physiological links of C_{x+c} SIF, PRI with GPC lacked power and were deemphasised as predictive features for ML. In these conditions, C_{x+c} was highest among low importances for physiological indicators and was joined by Anth and C_{a+b} on removal of thermal and structural layers. When CWSI was excluded as a model input for severely stressed crops, LAI importance was high, confirming the greater relative influence of structure in drought than in benign conditions. The displacement of EVI
by LAI in drought suggests redundancy between them; each may indicate canopy variability established prior to the onset of water stress and both are less affected by drought than are physiological traits.

In more moderate conditions, soil moisture heterogeneity drives variability in photosynthetic rate and hence physiological indicators across and between fields. This greater variability allows physiological features to convey more information, giving higher importance in GPC prediction. For example, at moderate stress in the 2019 durum crop, SIF approached CWSI in importance and all remaining physiological features were close together below SIF, none individually prominent but each more informative than structural traits. The order of physiological features changed little on removal of thermal and structural layers, showing their robustness and utility as GPC predictors.

**Fig. 9.** Relative importance (proportion) of input features to a gradient boosting machine estimating grain protein content (GPC; %) in bread (left) and durum wheat (right) in commercial fields near Kaniva, Australia. Three models are shown: physiological + structural + CWSI (model 1; a, d); physiological + structural (model 2; b, e); physiological (model 3; c, f). Each sub-figure represents the 2019 (left; bread n = 7213, durum n = 5030) and 2020 seasons (right; bread n = 11060, durum n = 17310). Error bars show standard deviation of the mean proportional importance over 80 runs.

**Table 4** Predictive skill ($r^2$, rRMSE; %) for Model 1, built with physiological + structural + CWSI layers, Model 2 (physiological + structural) and Model 3 (physiological only) across bread and durum wheat. Each model/product/year combination was run 80 times.

<table>
<thead>
<tr>
<th>Model</th>
<th>Bread wheat</th>
<th></th>
<th></th>
<th>Durum wheat</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r^2$</td>
<td>rRMSE</td>
<td>$r^2$</td>
<td>rRMSE</td>
<td>$r^2$</td>
<td>rRMSE</td>
<td>$r^2$</td>
</tr>
<tr>
<td>1</td>
<td>0.8</td>
<td>0.62</td>
<td>0.54</td>
<td>0.66</td>
<td>0.54</td>
<td>0.81</td>
<td>0.49</td>
</tr>
<tr>
<td>2</td>
<td>0.7</td>
<td>0.76</td>
<td>0.5</td>
<td>0.69</td>
<td>0.49</td>
<td>0.84</td>
<td>0.46</td>
</tr>
<tr>
<td>3</td>
<td>0.57</td>
<td>0.91</td>
<td>0.39</td>
<td>0.76</td>
<td>0.43</td>
<td>0.9</td>
<td>0.37</td>
</tr>
</tbody>
</table>

By LAI in drought suggests redundancy between them; each may indicate canopy variability established prior to the onset of water stress and both are less affected by drought than are physiological traits.

In more moderate conditions, soil moisture heterogeneity drives variability in photosynthetic rate and hence physiological indicators across and between fields. This greater variability allows physiological features to convey more information, giving higher importance in GPC prediction. For example, at moderate stress in the 2019 durum crop, SIF approached CWSI in importance and all remaining physiological features were close together below SIF, none individually prominent but each more informative than structural traits. The order of physiological features changed little on removal of thermal and structural layers, showing their robustness and utility as GPC predictors in such conditions. Meanwhile in 2020, when water stress was lowest, importance was more evenly spread between physiological, thermal and structural components and physiological indicators $C_{\text{arb}}$ and PRI reached higher importance than CWSI. In these conditions, $C_{\text{arb}}$ was the best indicator overall in bread wheat, and in durum was below only CWSI, SIF and PRI,
likely showing a water stress differential between the two wheat types.

Results at moderate and low stress confirm that $c_{x,1,c}$, Anth, PRI and SIF are sensitive to water stress that is yet insufficient to strongly diminish photosynthesis. Moreover, each of these consistently had equal or higher importance than $c_{x,1,b}$ in low stress, showing that short-term stress indicators also indicate GPC. Overall, physiological features were stable across conditions, and consistently supplied more information than structural features. This concurs with our own findings at plot scale and those of others that Anth, $c_{x,1,c}$ and SIF, or combinations thereof, were crucial in stress diagnosis (Poblete et al., 2021; Suárez et al., 2008; Zarco-Tejada et al., 2018). The PRI ~ SIF alignment also mirrors our plot and transect studies and comments PRI as a proxy for photosynthesis in wheat (Magney et al., 2016). Anth importance was moderate in all conditions and stable between wheat types in low stress; this shows a differential Anth response under variable water stress and was supported at plot scale and in the literature (Chalker-Scott, 2002; Li et al., 2018). Unstable importance in EVI, LAI and LIDF limits their value for GPC estimation, except in extreme stress, but together they contribute to model performance and generalisability. In benign conditions, LAI and LIDF generally show low importance.

4.4. Model predictive skill

Though a minor focus of our work, model skill improved on some previous results for wheat GPC. Using leave-one-out (LOO) validation, Overgaard et al. (2013) obtained $r^2 = 0.16–0.68$, while Hansen et al. (2002) achieved $r^2 = 0.56$ and Apan et al. (2006) $r^2 = 0.92$. Li et al. (2020) obtained $r^2 = 0.13–0.85$, testing on 33% of their observations, and Zhou et al. (2021) also realised $r^2 = 0.55–0.63$ in an unseen 31% of their full (n = 327) dataset. Rodrigues et al. (2018) obtained $r^2 = 0.21$. Like Overgaard et al. (2013), Li et al. (2020) and Zhou et al. (2021), we tested on a substantial unseen hold-out of observations across all fields, a more robust model proof than the LOO methods often seen. We also tested prediction with a field-wise LOO method, such that each field’s data were successively used as the unseen test set for a model calibrated on the rest. Successful demonstration of this is important for many potential applications of our methods. When zero data from the LOO field were included in training, predictive skill was very poor (not shown). We then reduced stepwise from 70% to 10% the availability of training data from the LOO field; predictive skill declined, but not dramatically, at each step down to 10%.

That our physiological layer predicted GPC with acceptable accuracy without thermal or structural features attests to its coverage of GPC-relevant traits, from instantaneous SIF to the relatively stable $c_{x,1,b}$. These results agree with findings that SIF and $c_{x,1,b}$ were far better than structural measures in estimating wheat leaf N (Camino et al., 2018). Because the indicators we use to estimate GPC are proxies of the water and nutrient stresses present in the region of study, our methods will probably work in regions with similar water and nutrient stress levels. In regions which are not water- or nutrient-limited, plant traits other than those described in this paper will be sensitive to GPC. Each of these proposals should be tested; our methods should also be tested with RS data of lower resolution, from spaceborne sensors, and in diverse seasonal, soil, cultivar, and agronomic conditions. It would also be valuable to test timeseries data captured within and across seasons. Further inputs such as year-to-date rainfall, soil or agronomic data may improve model predictive skill, as may training on multi-year databases. Further investigation of field-wise LOO is also needed to assess model transferability to unseen paddocks.

5. Conclusions

This study identified the most important hyperspectral-based plant traits related to grain protein content in rainfed wheat under variable stress levels. In experimental plots, two variants of the PRI index related to the xanthophyll pigment cycle showed consistent trends very similar to GPC along the induced N nutrition gradient, and in this respect performed better than any other spectral trait. In commercial crops, we implemented a gradient boosted machine to investigate relationships between input features and GPC. The thermal CWSI indicator of canopy transpiration contributed strongly to the model under water stress conditions, while Anth, $c_{x,1,c}$, PRI and SIF consistently showed high importance in GPC estimation under more benign conditions. Structural indicators such as LAI or its proxy NDVI contributed significantly less. We obtained promising results using gradient boosted machine learning to estimate GPC from hyperspectral and thermal images. Results yielded $r^2 = 0.80$ with RMSE = 0.62% between predicted and observed GPC using a model built with thermal and physiological traits quantified by radiative transfer modelling methods.

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 the Birchip Cropping Group, Riverine Plains Incorporated and Foundation for Arable Research Australia for their provision and management of the plot trials, collection of GPC data and support of field operations. We also thank Jonathan Dyer, grain grower at Kaniva for his enthusiastic participation, advice and for access to GPC data. A. Gracia-Romero and A. Belwalkar are acknowledged for their technical support during the field and airborne campaigns. Andrew Longmire received PhD funding from the Grains Research and Development Corporation Australia. We also thank three anonymous reviewers for their thoughtful contributions, which improved this manuscript.

Author contributions

AL conceived the project, did the experimental design, collected field data at ground level, did the machine learning and data analyses, and wrote the manuscript. TP performed model inversions and supervised the machine learning algorithms. JH led the plot experiment at Birchip. AL, PZ and TP contributed to the data interpretation, and to manuscript drafting. PZ supervised the remote sensing work and led the airborne hyperspectral campaigns. DC supervised the work. JH, PZ and TP reviewed and edited the manuscript.

References
