Integration of APSIM and PROSAIL models to develop more precise radiometric estimation of crop traits using deep learning

Qiaomin Chen[1, 2, *], Bangyou Zheng[2], Tong Chen[3], Scott Chapman[1,*]

[1] School of Agriculture and Food Sciences, The University of Queensland, St Lucia, 4067, QLD, Australia

[2] CSIRO Agriculture and Food, Queensland Biosciences Precinct 306 Carmody Road, St Lucia, 4067, QLD, Australia

[3] School of Information Technology and Electrical Engineering, The University of Queensland, St Lucia, 4067, QLD, Australia

[*]Corresponding Author: qiaomin.chen@uq.edu.au or scott.chapman@uq.edu.au

Abstract

A major challenge for the estimation of crop traits (biophysical variables) from canopy reflectance is the creation of a high-quality training dataset. This can be addressed by using radiative transfer models (RTMs) to generate training dataset representing 'real-world' data in situations with varying crop types and growth status as well as various observation configurations. However, this approach can lead to "ill-posed" problems related to assumptions in the sampling strategy and due to uncertainty in the model, resulting in unsatisfactory inversion results for retrieval of target variables. In order to address this problem, this research investigates a practical way to generate higher quality 'synthetic' training data by integrating a crop growth model (CGM, in this case APSIM) with an RTM (in this case PROSAIL). This allows control of uncertainties of the RTM by imposing biological constraints on distribution and co-distribution of related variables. Subsequently, the method was theoretically validated on two types of synthetic dataset generated by PROSAIL or the coupling of APSIM and PROSAIL through comparing estimation precision for leaf area index (*LAI*), leaf chlorophyll content (*Cab*), leaf dry matter (*Cm*) and leaf water content (*Cw*). Additionally, the capabilities of current deep learning techniques using high spectral resolution hyperspectral data were investigated. The main findings include: (1) Feedforward neural network (FFNN) provided with appropriate configuration is a promising technique to retrieve crop traits from input features consisting of 1 nm-wide hyperspectral bands across 400-2500 nm range and observation configuration (solar and viewing angles), leading to a precise joint estimation for *LAI* (RMSE=0.061 m^2 m^{-2}), *Cab* (RMSE=1.42 µg cm^{-2}), *Cm* (RMSE=0.000176 g cm^{-2}) and *Cw* (RMSE=0.000319 g cm^{-2}); (2) For the aim of model simplification, a narrower range in 400-1100 nm without observation configuration in input of FFNN model provided less precise estimation for *LAI* (RMSE=0.087 m^2 m^{-2}), *Cab* (RMSE=1.92 µg cm^{-2}), *Cm* (RMSE=0.000299 g cm^{-2}) and *Cw* (RMSE=0.001271 g cm^{-2}); (3) The introduction of biological constraints in

training datasets improved FFNN model performance in both average precision and stability, resulting in a much accurate estimation for *LAI* (RMSE=0.006 m^2 m^{-2}), *Cab* (RMSE=0.45 µg cm^{-2}), *Cm* (RMSE=0.000039 g cm^{-2}) and *Cw* (RMSE=0.000072 g cm^{-2}), and this improvement could be further increased by enriching sample diversity in training dataset.

Keywork: model integration; neural network; hyperspectral data; variable retrieval

1. Introduction

Since the 1960s, using satellite imagery to measure reflectance of surfaces at a scale of tens of metres has been utilised to monitor vegetation health and to attempt to estimate and forecast changes in vegetation cover and condition (Thenkabail et al., 2019). More recently, these imagery methods have been deployed in more proximal sensors (planes, drones, vehicles) that allow analysis of vegetation at higher resolutions (to sub-centimetre scales) in a research field that is sometimes referred to as 'high-throughput phenotyping' (HTP) (Chapman et al., 2018). The aim is to indirectly retrieve crop traits such as water and chlorophyll content, with estimation of integrative traits like leaf area index (LAI) gaining the most attention (e.g. Bacour et al., 2002; Jay et al., 2019; Shibayama and Watanabe, 2007; Xu et al., 2019; Yu et al., 2017). HTP methods based on sensor and imaging technologies can rapidly measure a large number of crop traits across time and space in a cost- and labour-efficient way, which can benefit applications in precision agriculture and plant breeding.

Existing retrieval methods can be classified into two major categories depending on where the source of the training data for establishing relationship between target crop trait and spectral signal (canopy reflectance and its derived variables such as vegetation index): (1) statistical methods use observation data collected from practical experiments to build relationship; (2) physical methods either directly use established cause-effect relationship expressed in radiative transfer models (RTMs) or use simulation (synthetic) data generated by these models to rebuild

relationship. Compared to statistical methods, the main advantage of physical methods is to allow construction of a training dataset that represents the entire range of possible situations varying in crop types and growth status as well as observation configurations (Baret and Buis, 2008; Dorigo et al., 2007) and consequently provides an easier way to develop more general relationships unrestricted to situations for variable retrieval. There is an increasing interest of the application of 'model inversion methods' to RTMs, among of which PROSAIL is the most popular one and has been widely used for variable retrieval (Berger et al., 2018). Within a RTM such as PROSAIL, canopy reflectance across 400-2500 nm range as model output is regulated by input variables including leaf properties, canopy architecture, soil background and observation geometry (viewing and illumination conditions) via radiation absorption and scattering (Jacquemoud et al., 2009). Theoretically, only crop traits presented as input variables in RTMs could be retrieved from model inversion; however, by treating these retrieved variables as intermediate mediums, model inversion can be extended to broader applications, i.e., estimation of leaf properties at canopy level (Campos-Taberner et al., 2018), phenology prediction (J. Xu et al., 2019), land quality evaluation (Wu et al., 2019) and stress detection (Xia et al., 2019).

Model inversion methods are generally subdivided into three sub-categories: numerical optimization approach (e.g., Bacour et al., 2002; Eon et al., 2019; Lunagaria and Patel, 2019), look-up table approach (e.g., Weiss et al., 2000; X. Xu et al., 2019; Zhu et al., 2019), machine learning approach including use of neural networks (e.g., Bacour et al., 2006; García-Haro et al., 2018; Upreti et al., 2019). As summarized in reviews of variable estimation from remote sensing data, different methods have advantages and limitations with no obvious global solution (Baret and Buis, 2008; Dorigo et al., 2007; Verrelst et al., 2015). Although neural networks did not outperform the other approaches for variable retrieval in previous studies (e.g., Combal et al., 2003; Dhakar et al., 2019; Upreti et al., 2019), this may have been due to the low quality of

the training dataset without correction using prior knowledge, and/or insufficient utilization of spectral data and/or limitations of the selected neural network algorithms. Compared with other methods, neural networks are theoretically superior in inverting models with massive input variables (such as using hundreds of hyperspectral bands as input to infer canopy variables) and are computationally efficient once the network is fully trained. Thanks to recent developments in deep learning techniques, in this research, we attempted to explore the use of deep learning approach in model inversion for canopy variable retrieval by optimizing network architecture (hyperparameter tuning) and improving training data quality.

Although model inversion methods provide a reasonable way for estimating variables from remote sensing data, none of them can avoid the "ill-posed" problem, namely, the same model output may result from different combination of model input variables. Essentially, the problem is caused by the model uncertainty which results from its simplification of the structure and biochemistry of a canopy, so that more than one state situations of a canopy could result in exactly the same reflectance profile. In practical applications, this problem is aggravated by the poor input parameter selection, i.e. not accounting for bio-physical limitations in the combinations of parameters physically existing in the real-world. However, this problem can be alleviated by using prior knowledge to strengthen constraints on individual variables or between variables. The simplest way is to define the lower and upper values between which the target trait can be retrieved from based on prior information, for example, field measurement data was used for defining input parameter range in the study of Lunagaria and Patel (2019). M. Xu et al. (2019) indirectly introduced constraints between leaf chlorophyll content and LAI by establishing a 2-dimensional matrix-based relationship between leaf chlorophyll content and two vegetation indices (VIs) for VI-based look-up table inversion, which resulted in a better estimation precision than using individual VI. In addition, the utilization of multi-angular observation data in numerical optimization inversion was reported to improve estimation

111 precision of LAI and leaf chlorophyll content (Roosjen et al., 2018) by reducing possible
112 solutions of variable combination via solutions interception.

113 Linking crop growth model (CGM) to RTM provides a more straightforward solution to address
114 this "ill-posed" problems by directly constraining the sets of RTM input parameters that
115 contribute to canopy reflectance. Theoretically, such an integration method can be implemented
116 in two ways to improve variable retrieval from remote sensing data. The first way is to
117 calibrate/parameterize CGM using canopy reflectance and then using the calibrated CGM to
118 predict target crop traits. Such an application mode can directly retrieve those variables not
119 included in RTMs such as crop yield and also provide variable estimation across the whole
120 growth season (e.g., Guo et al., 2019; Huang et al., 2019; Thorp et al., 2012). The other way is
121 to convert CGM output variables into input variables of RTM and then apply model inversion
122 method on these constrained input variables and corresponding canopy reflectance. This
123 approach has been rarely discussed or explored.

124 This research focused on estimation of leaf area index (*LAI*), leaf chlorophyll content (*Cab*),
125 leaf dry weight (*Cm*) and leaf water content (*Cw*) of wheat in four locations across Australia
126 wheatbelt. The overall objective was to investigate inversion procedures based on a deep
127 learning approach (feedforward neural network, FFNN) for crop trait estimation, with a special
128 focus on alleviating the "ill-posed" problem in model inversion through linking a CGM
129 (APSIM) and a RTM (PROSAIL) to generate a higher quality training dataset. Firstly, a
130 baseline FFNN model was established to evaluate the use of FFNN for crop trait retrieval.
131 Secondly, this baseline model was used to explore possibility of reduction of hyperspectral
132 bands and effect of observation configuration (solar and viewing angles) for the aim of model
133 simplification. Finally, this simplified model was trained using different datasets generated by
134 PROSAIL or the coupling of APSIM and PROSAIL to investigate the function of model
135 integration by comparing performance of trained FFNNs.

2. Methods

2.1 Overview

This research contains several key steps as shown in Figure 1. APSIM and PROSAIL models were integrated by passing variables from the former to the latter based on variable transformation relationships. A defined wheat growing space (characterized by genotype, environment and management) and observation conditions (determined by local latitude, day of year and day time) were set up to run APSIM and PROSAIL for simulation of crop traits and canopy reflectance, which resulted in two types of synthetic datasets. The first dataset (PROSAIL dataset) uses the ranges of the input parameters converted from APSIM outputs but allows PROSAIL to be run using samples from full parameter space for any combination of inputs. The second dataset (APSIM-PROSAIL dataset) directly uses input data converted from APSIM outputs to explore a sub-space of input parameters (i.e. limited by the APSIM biology) to run PROSAIL. According to research objectives, these synthetic datasets were reconstructed and used for FFNN training and evaluation in order to explore the possibility of hyperspectral bands reduction, the effect of observation configuration (solar and viewing angles), and the effect of limiting the PROSAIL input parameters to the sub-space as determined by the APSIM.

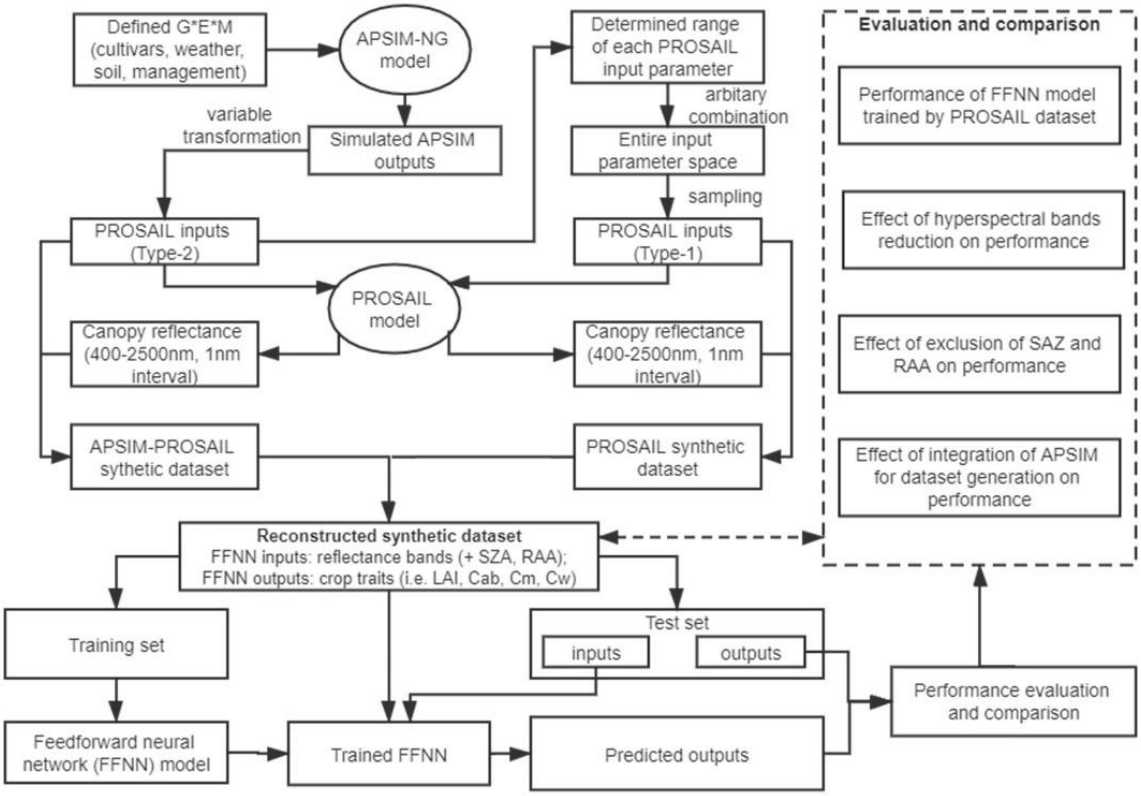


Figure 1 Research flowmap. APSIM-NG denotes Agricultural Production Systems sIMulator (APSIM) Next Generation, which is a crop model. PROSAIL is a radiative transfer model, coupling a leaf optical property model (PROSPECT-D) and a canopy bidirectional reflectance model (4SAIL).

2.2 Models and their integration

Agricultural Production System Simulator (APSIM) Next Generation (https://www.apsim.info/apsim-next-generation/) is the new version of APSIM, which is simpler and faster than the classic version (i.e. 7.10., D. Holzworth et al., 2018). APSIM Next Generation or APSIM is driven by major processes in crop physiology and interactions with environment factors and management practices, and widely is used to simulate dynamics of many crop traits (e.g. leaf area index, dry weight of organ parts (i.e. grain, leaf, spike, stem and root)) during growth season at daily scale. APSIM has been validated in many regions around the world (Holzworth et al., 2014).

PROSAIL is the combination of PROSPECT (a leaf optical property model) and SAIL (a canopy bidirectional reflectance model). PROSAIL links the spectral variation of canopy reflectance (mainly related to leaf biochemical contents) with its directional variation (primarily related to canopy architecture and soil/vegetation contrast), which is key to simultaneously estimate canopy biophysical/structural variables such as leaf chlorophyll content and LAI (Jacquemoud et al., 2009). The current version of the PROSAIL model couples PROSPECT-D (see Féret et al., 2017) and 4SAIL (see Berger et al., 2018) and can be downloaded from http://teledetection.ipgp.jussieu.fr/prosail/. Both input and output variables of this PROSAIL model are presented in Table 1. The 14 input parameters can be divided into four categories: leaf properties (*N*, *Cw*, *Cm*, *Cab*, *Car*, *Cant*, *Cbrown*), background soil properties (*rsoil*), canopy architecture (*LAI*, *LIDF*, *hspot*) and solar-object-sensor observation geometry (*SZA*, *VZA*, *RAA*). PROSAIL can output directional canopy reflectance, which is also represented using canopy reflectance or model output without explicit specification in the following sections. For further details about PROSPECT and SAIL model, refer to the original papers (Féret et al., 2017; Jacquemoud and Baret, 1990; Verhoef, 1998, 1984).

Table 1 Description of input parameters and output of PROSAIL (PROSPECT-D + 4SAIL)

Variable	Unit	Description
Input		
N	unitless	Leaf mesophyll structure parameter, relates to the cellular arrangement within the leaf.
C_w	g cm^{-2} or cm	Leaf water content (g cm^{-2}) or leaf equivalent water thickness (cm)
C_m	g cm^{-2}	Leaf dry matter content per leaf area
C_{ab}	µg cm^{-2}	Leaf chlorophyll-a and -b content per leaf area
C_{ar}	µg cm^{-2}	Leaf carotenoid content per leaf area
C_{ant}	µg cm^{-2}	Leaf anthocyanins content per leaf area
C_{brown}	unitless	Leaf brown pigment concentration

Variable	Unit	Description
r_{soil}	unitless	Reflectance of soil as a libertarian surface. It is usually adjusted by a soil brightness factor: asoil (to be multiplied with single r_{soil} spectrum) or psoil (scaling factor between the two model-implemented r_{soil} spectra - wet versus dry).
LAI	$m^2\ m^{-2}$	Leaf area index
LIDF	-	Leaf inclination distribution function. There are two methods provided in 4SAIL to calculate LIDF: use two parameters LIDFa and LIDFb or single parameter – ALA (average leaf angle, degree)
hspot	$m\ m^{-1}$	Hot spot size parameter. It is primarily designed to correct the canopy reflection regarding bidirectional effects. Hot spot effect is the case that a spot displays maximum reflectivity and appears brighter than surroundings because of no visible shadows at the hot spot position where the sensor is in direct alignment between the sun and the ground target.
SZA	degree	Solar zenith angle (from vertical)
VZA	degree	Viewing (or observing) zenith angle (from vertical)
RAA	degree	Relative azimuth angle. It is equal to viewing azimuth angle minus solar azimuth angle.
Output		
resv	unitless	Directional reflectance of canopy

The coupling of APSIM and PROSAIL is realized by passing output variables of APSIM to PROSAIL as input variables. This permits the coupling model to estimate canopy reflectance from 400 to 2500 nm in 1 nm interval at defined observation conditions (determined by latitude, day of year, and day time) given that required parameters are specified. The transformation of variables is based on a series of equations (Table 2) and more details could be found in Section 1 of supplementary materials.

190 *Table 2 Variable transformation from APSIM output to PROSAIL input*

APSIM output variable	Transformation formula	PROSAIL input variable
LAI_{Total}, *LDW*	$N=(0.9*SLA+0.025)/(SLA-0.1)$ where $SLA=10*LAI_{Total}/LDW$	*N*
Z_s, LAI_{Total}, LAI_{Dead}	$C_w = \begin{cases} -0.000196\ Z_s + 0.0298 & (f_{dead} = 0) \\ 0.0223 \exp(-1.90\ f_{dead}) & (f_{dead} > 0) \end{cases}$ where $f_{dead}=LAI_{Dead}/LAI_{Total}$	C_w
LDW, LAI_{Total}	$C_m=10^{-4}*LDW/LAI_{Total}$	C_m
CNC, LAI_{Total}	$C_{ab}=26*LNC$ where $LNC=CNC/LAI_{Total}$	C_{ab}
CNC, LAI_{Total}	$C_{ar}=0.216*C_{ab}$	C_{ar}
/	Fixed C_{ant} to 0	C_{ant}
/	Fixed C_{brown} to 0	C_{brown}
/	Fixed p_{soil} to 1	p_{soil}
LAI_{Total}	$LAI = LAI_{Total}$	*LAI*
/	Fixed ALA to 50˚	*ALA*
LAI_{Total}	$hspot=a/LAI_{Total}$ (a is an empirical parameter and is set as 0.5)	*hspot*
L, *DOY*	$\cos(SZA)=\sin(L)\sin(\delta)+\cos(L)\cos(\delta)\cos(h)$ given $\delta=23.45\sin(360365(284+DOY))$ $h=15(AST-12)$	*SZA*
/	Fixed VZA to 0	*VZA*
L, *DOY*	RAA = SAA – VAA given VAA=0 (when VZA=0) $\sin(SAA)=\cos(\delta)\sin(h)\cos(90°-SZA)$	*RAA*

Notes:

"/" denotes no output for equivalent input;

SLA ($cm^2\ mg^{-1}$): leaf area per unit leaf dry weight;

APSIM output variable	Transformation formula	PROSAIL input variable

LAI_{Total} (m^2 m^{-2}): total leaf area index;

LDW (g m^{-2}): leaf dry weight per unit planting area;

Z_s: decimal zadok score for the growth stage;

f_{dead}: the fraction of dead leaves;

LAI_{Dead} (m^2 m^{-2}): leaf area index of senesced or dead leaves;

LNC (g m^{-2}): leaf nitrogen content per unit leaf area;

CNC (g m^{-2}): canopy nitrogen content per unit planting area;

L (°): local latitude;

δ (°): solar declination angle;

DOY: day of year;

h (°): hour angle;

AST (h): apparent solar time for phenotyping crops and here is set at three levels: 10:00, 12:00 and 14:00.

2.3 Synthetic dataset generation

At four sites used to represent diverse conditions across the Australia wheatbelt (Table 3), simulations were run with historical weather records from 2000-2019, the typical soil condition with best initial soil water and local management practices (i.e. fertilization, Table 3, (Chenu et al., 2013)). For each site at each year, 9 cultivars (varying in habit and/or development speed) and 9 sowing dates (from 1-May to 30-June in 1-week interval) were selected to characterize different wheat growth patterns (Table 3). In total, 6 480 simulation seasons were performed using APSIM Next Generation. Major crop traits (Table 2) were output in daily step from emergence to harvest, resulting in 1 080 680 daily records. Based on three defined observation times, these APSIM output records were converted into 3 243 040 PROSAIL input records using variable transformation formulas presented in Table 2, which were then used to determine variation range of each PROSAIL input variable (Table 4).

204 *Table 3 Information of genotype, environment and management used for simulation to represent Australia wheatbelt*

Cropping Area	Site	Latitude	Longitude	Soil Classification	PAWC at sowing (mm)	Nitrogen (kg ha^{-1})	Sowing Date	Genotype (habit / development speed)
West	Merredin	-31.48	118.28	Shallow loamy duplex	86	20-20-30	1-May; 8-May; 15-May; 22-May; 29-May; 5-Jun; 12-Jun; 19-Jun; 26-Jun	Young (Spring / Very-fast); Gauntlet (Spring /Fast); Ellison (Spring / Mid-fast); Wills (Spring / Mid); Lancer (Spring / Slow); Forrest (Spring / Very-slow); Longsword (Winter / Fast); Kittyhawk (Winter / Mid); Manning (Winter / Slow)
South-east	Yanco	-34.61	146.42	Brown sodosol	191	40-40a-40b		
East	Narrabri	-30.34	149.76	Grey vertosol	218	130-0-0		
East	Gatton	-27.54	152.33	Black vertosol	225	20-30-0		

Notes: The first three sites (Merredin, Yanco and Narabri) are Managed Environment Facilities of Australia and Gatton is an experimental station in the University of Queensland. Plant available water content (PAWC) at sowing is indicated for each soil at its best level of initial soil water (referring to (Chenu et al., 2013)). Applied nitrogen is represented by 'x-y-z': x, nitrogen (urea for the East and nitrate for the rest areas) applied at sowing; x, y, nitrogen (nitrate) applied at stage of 'beginning of stem elongation' and 'flag leaf just visible', respectively. [a] applied only if more than 100 mm rainfall from sowing to stage of 'beginning of stem elongation'; [b] applied only if soil PAWC > 60% of maximum.

205

206 *Table 4 Setting of PROSAIL input parameters used for sensitivity analysis and synthetic dataset generation*

Parameters	Local sensitivity analysis (OAT)		Global sensitivity analysis (EFAST) / PROSAIL dataset		
	Baseline	Levels	Lower Bound	Upper Bound	Distribution
N	2.25	1, 1.6, 2.3, 2.9, 3.5	1	3.5	Uniform
C_w	0.017	0.003, 0.01, 0.017, 0.023, 0.03	0.003	0.03	Uniform
C_m	0.006	0.002, 0.004, 0.006, 0.008, 0.01	0.002	0.01	Uniform

Parameters	Local sensitivity analysis (OAT)		Global sensitivity analysis (EFAST) / PROSAIL dataset		
	Baseline	Levels	Lower Bound	Upper Bound	Distribution
C_{ab}	48	5, 26, 48, 79, 90	5	90	Uniform
C_{ar}	10	1, 5, 10, 15, 20	1	20	Uniform
C_{ant}	0	0	0	0	Fixed
C_{brown}	0	0	0	0	Fixed
p_{soil}	1	1	1	1	Fixed
LAI	3.5	0.5, 1.5, 2.5, 3.5, 4.5, 5.5, 6.5	0.01	7	Uniform
ALA	50	50	50	50	Fixed
hspot	1.7	0, 0.8, 1.7, 2.6, 3.5	0	3.5	uniform
SZA	35	0, 18, 35, 52, 70	0	70	Uniform
VZA	0	0	0	0	Fixed
RAA	45	0, 22, 45, 68, 90	0	90	Uniform

The synthetic dataset was a set of data pairs constructed from PROSAIL input variables and corresponding outputs. It is an ideal means to use synthetic dataset to demonstrate the application of PROSAIL inversion for estimation of crop traits in theoretical dimensionality, since such a synthetic dataset represents the whole range of possible situations varying in crop types and growth status as well as observation conditions. Based on combination mode of PROSAIL input parameters, the synthetic dataset can be classified into two types: PROSAIL dataset and APSIM-PROSAIL dataset. For PROSAIL dataset, the entire parameter space of input variables consists of arbitrary combination of each parameter changing in their ranges defined in Table 4. The input variables of this type of datasets were generated by sampling a subset from this parameter space based on EFAST's resampling scheme with given minimum sample size (n). EFAST sensitivity analysis was undertaken multiple times to determine the appropriate n size. Our results show that sensitivity indices can converge to a relative robust value where $n \geq 5000$ (Figure S3 in supplementary materials), indicating a subset

with sample number of 45000 (=5000*9) is sufficient to reflect relationships between model's input and output. Compared with input variables of PROSAIL dataset, input variables of APSIM-PROSAIL dataset are not combined arbitrarily but are constrained to wheat growth pattern. The input variables of APSIM-PROSAIL dataset were those directly converted from ASPIM output variables, which contained 2 149 226 unique records after omitting 1 092 814 duplicated ones from the total 3 242 040 records. To our objectives, several synthetic datasets were generated in the way mentioned above (Table 5) and the density distribution and spatial co-distribution of PROSAIL input variables of these datasets in one- or two-dimensional space are presented in Figure 2.

Table 5 Information of synthetic dataset's name, type, number and description

Dataset Name	Dataset Type	Sample Number	Description
p_initialTrain	PROSAIL	45 000	Sampling from parameter space based on EFAST's resampling scheme with n=5000; used to determine FFNN model's hyperparameters in initial training
p_test	PROSAIL	10 000	Selecting from p_initialTrain; used as test set to evaluate performance of trained FFNN models (trained by p_train1) in experiments SE1-SE7 (see Table 6)
p_train1	PROSAIL	90 000	Sampling from parameter space based on EFAST's resampling scheme with n=10000; used as train set in in experiments SE1-SE7 (see Table 6)
ap_test	APSIM-PROSAIL	10 000	Selecting from total 2 149 226 unique samples; used as test set to evaluate performance of trained FFNN models (trained by ap_train1 or ap_train2) in experiment SE8-SE9 (see Table 6)
ap_train1	APSIM-PROSAIL	90 000	Selected from the remaining 2 139 226 samples; used as train set in experiment SE8 (see Table 6)
ap_train2	APSIM-PROSAIL	2 139 226	The remaining 2 139 226 samples; used as train set in experiment SE9 (see Table 6)

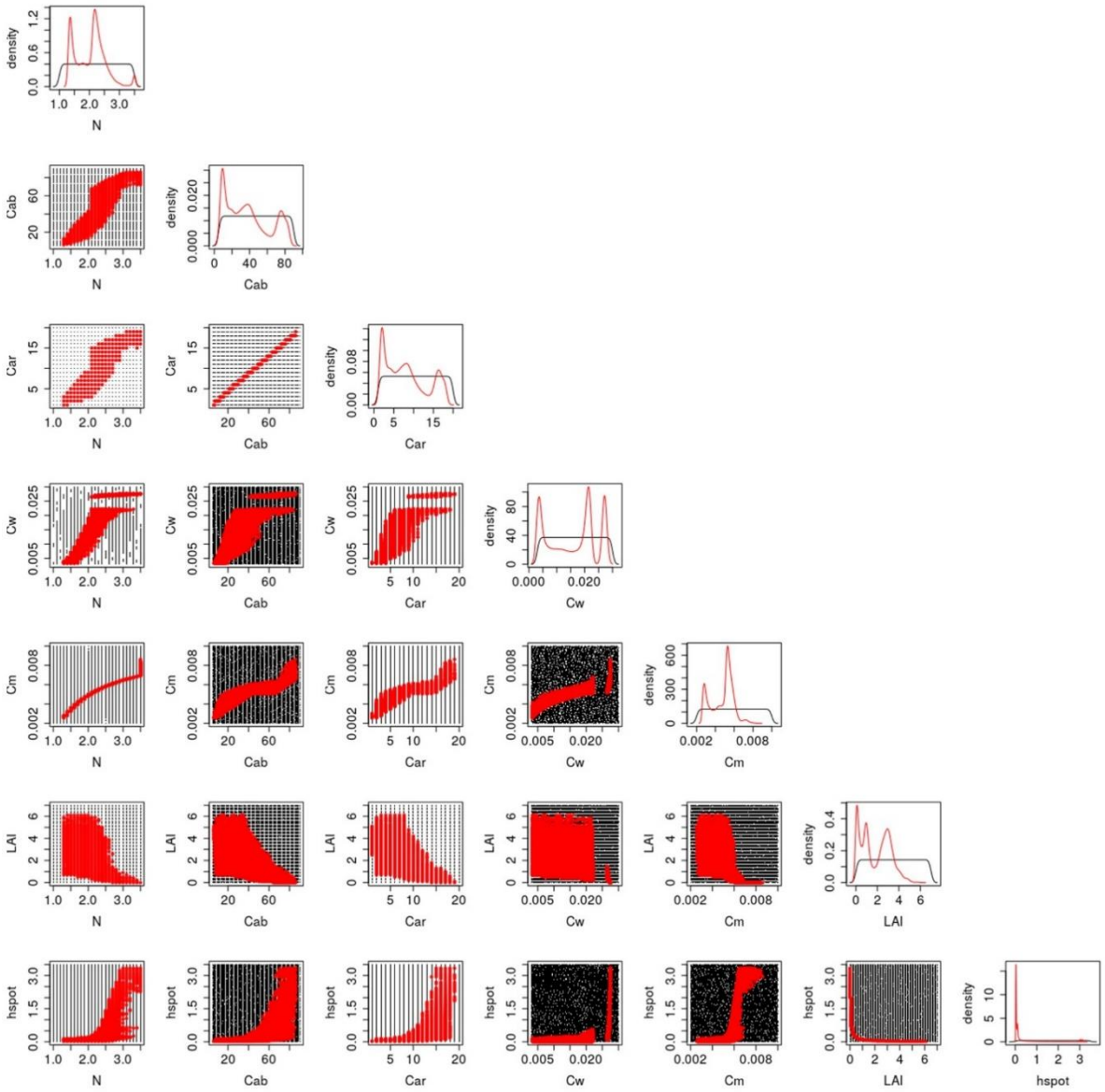


Figure 2 Density distribution and spatial co-distribution of PROSAIL input variables of two types of dataset. Black symbol represents PROSAIL dataset (p_train1) while red symbol represents APSIM-PROSAIL dataset (ap_train1).

2.4 Sensitivity analysis

For the PROSAIL model, not all canopy reflectance at wavelength from 400 to 2500 nm are sensitive to variation of input parameters. A simplified model with fewer insensitive outputs is superior to the full model for variable retrieval through inversion against a spectral image. For example, inverting variable from a parameter space with fewer possible solutions can improve inversion efficiency and mitigate the ill-posed problem (Verrelst and Rivera, 2017). In order to

generate a more representative synthetic dataset used for variable retrieval, a sensitivity analysis was conducted to evaluate the relative importance of each input variable in PROSAIL model and subsequently identify the most influential variables and most sensitive wavelength range.

Sensitivity analysis includes local and global sensitivity analysis. Local sensitivity analysis is commonly referred to "one-at-a-time (OAT)", which changes one input variable at a time while holding others at their central values for measurement of variation in the model outputs compared with outputs at central point (Verrelst and Rivera, 2017). In contrast, global sensitivity analysis explores the entire variable space and simultaneously changes all variables. The "extended-FAST" (EFAST) method (Saltelli et al., 1999) is a variance-based method and is frequently used in global sensitivity analysis. This method allows the estimation of the first-order and total effect indices for all the input parameters at a total cost of n × p simulations (p is the number of parameters, n is the sample size). First-order effect indices (S_{1i}) represent the isolated contribution of i^{th} parameter to the variance of the model output (i.e. canopy reflectance in this study). Total effect indices (S_{ti}) represent the total contribution of i^{th} parameter: the isolated contribution of a parameter plus its interactions with other input parameters. The normalized total effect indices (RC_S_{ti}) are appropriate to represent the relative contribution of i^{th} input parameter to variation of model output and the normalized first-order effect indices (RC_S_{1i}) represent the relative isolated contribution.

$$RC\_S_{ti} = \frac{S_{ti}}{\sum_{i=1}^{p} S_{ti}} \times 100\% \quad (1)$$

$$RC\_S_{1i} = \frac{S_{1i}}{\sum_{i=1}^{p} S_{1i}} \times 100\% \quad (2)$$

2.5 In this research, OAT analysis was undertaken to present how canopy reflectance responds to variation of each input variable in the wavelength range from 400 nm to 2500 nm. In

addition, EFAST was chosen as the global sensitivity analysis method to quantify relative contribution of each input variable on canopy reflectance in 400-2500 nm range. Values and ranges of PROSAIL input parameters used for sensitivity analysis were set as shown in Table 4. **Feedforward neural network (FFNN)**

Feedforward neural networks (FFNNs), also called deep feedforward networks, or multilayer perceptrons, are the quintessential deep learning models (Goodfellow et al., 2016). A FFNN defines a mapping $y=f(x;\theta)$ and learns the value of the parameters θ that result in the best function approximation for a prediction: either a classification (discrete) or a regression (continuous). Within this network structure, data passes from the input "x" corresponds to the raw data, which goes through intermediate computations in the function "f" with the parameters "θ", to the output "y" in a single pass without any feedbacks or cycles.

As presented in Figure 3, FFNN has a multilayer structure consisting of an input layer (the first layer of a network), an output layer (the final layer of a network), and one or more hidden layers (the remaining layers of a network). The total number of layers is called the depth of a network, and each hidden layer of the network consists of many neurons (or units). The dimensionality of these hidden layers determines the width of the network. According to the universal approximation theorem (Cybenkot, 1989; Hornik et al., 1989), a feedforward network can approximate an arbitrary function even with only one hidden layer that is sufficiently wide. However, simply increasing the number of neurons can easily lead to over-parameterization, hence increasing depth seems to be an alternative as experiences from previous studies showed greater depth typically resulted in better generalization (Lin et al., 2014; Zhang et al., 2017). In addition to depth and width, a FFNN has other necessary components: activation, optimizer and loss function. The activation function is a function used to transform data, which allows the layer to learn not only the linear transformation but also the non-linear transformations of the input data and to increase the capacity for better learning of the complex mapping from the

input to the output. The optimizer specifies how the training (learning) proceeds through updating model parameters θ (weights) towards a better prediction based on feedback signal from loss function. The magnitude of the move of weights update in training is controlled by learning rate. The loss function quantifies the accuracy of a model on the training data and is used to navigate the training process so as to minimize the training loss. Since a network is trained by iteratively going through the training data, the loss score decreases as training proceeds and finally, it yields a trained network that can accurately estimate the output y with $f(x;\theta)$ when consistent minimal loss is observed.

The FFNN model was implemented using Keras in TensorFlow 2.3.0 (https://www.tensorflow.org/). Based on research objectives, the PROSAIL and APSIM-PROSAIL datasets generated in previous steps were reconstructed and several simulation experiments were designed to facilitate other steps in the method (Table 1). In the following sections, we demonstrated how to build, train and evaluate FFNN.

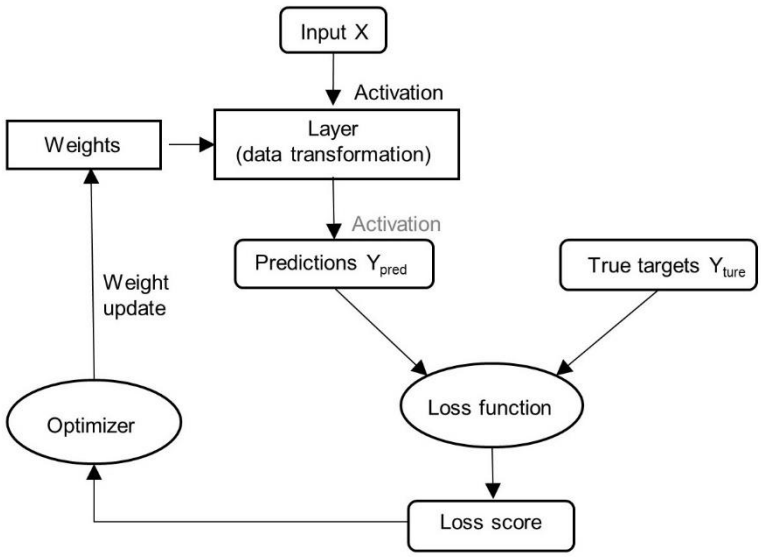


Figure 3 Feedforward neural network working roadmap (adapted from (Chollet, 2017)). 'Activation' coloured in black is necessary while that one coloured in grey is unnecessary.

2.5.1 Hyperparameter tuning

Hyperparameters denote variables that govern the training process and the architecture of an FFNN model. Hyperparameters include model hyperparameters (influencing model basic architecture such as the number and width of hidden layers) and algorithm hyperparameters (influencing the speed and quality of training process such as learning rate, activation and optimizer). The process of determining the optimal hyperparameters is called hyperparameter optimization or hyperparameter tuning. For a regression task presented here, the most common loss function is mean squared error (MSE).

In the initial training experiment 'hypertuning', dataset 'p_initialTrain' (Table 6) was used for hyperparameter tuning, and the best combination of hyperparameters was determined via a two-step optimization process. At the first step, the best combination of learning rate, activation and optimizer was selected by changing these three hyperparameters and holding the number and width of hidden layers to default values (3 hidden layers with 256 units for each layer). At the second step, the best combination of depth and width was chosen based on the algorithm hyperparameters selected at the first step. During this hyperparameter tuning, 3 common values of learning rate, 10 activation functions, 9 optimizers, 5 levels of hidden layer number and 8 levels of unit number were evaluated. In the end, the model structure with 3 hidden layers and 512 units for each hidden layer and using 0.001 as learning rate, 'softplus' as activation, 'Adamax' as optimizer, was selected as the optimal FFNN structure (Figure S4 and Figure S5) used in later training experiments.

2.5.2 Training and evaluation

The best model architecture determined in experiment 'hypertuning' was used in the following experiments SE1 to SE9. Experiment SE1 was designed to check how well PROSAIL inversion based on FFNN can retrieve target crop traits (i.e. *LAI*, *Cab*, *Cm* and *Cw*) from canopy reflectance. Experiments SE2 to SE6 were used to check whether it is possible to reduce

hyperspectral bands used in FFNN's input on the base of ensuring model's prediction precision by comparing performance of models using different wavelength range in input. Experiments SE5 and SE7 were used to check whether observation geometry information is necessary to achieve good prediction of crop traits by comparing performance of models including or excluding this information from the input. In order to verify the hypothesis that APSIM-PROSAIL dataset outperforms PROSAIL dataset when being used for traits retrieval, experiments SE7 to SE9 were designed to use different types of datasets for training model.

Loss, the total mean squared error of all output variables after normalization, was used to evaluate FFNN model's overall performance for joint estimation of all target variables: a smaller loss indicates a higher precision. Absolute error (AE, including its variation range (AE range), standard deviation (std AE) and mean (MAE)) and root mean squared error (RMSE) were used for measurement of each target variable after de-normalization. In particular, AE range and std AE were used to evaluate model's stability for estimating each target variable while MAE and RMSE were used to evaluate model's average precision for this variable: a narrower AE range, a smaller std AE, MSE and RMSE indicate a better performance.

Table 6 Details of simulation experiment's name, original and reconstructed synthetic dataset, and evaluation metric

Experiment name	Original synthetic dataset	Reconstructed synthetic dataset		Evaluation metric
	Original Dataset	Reconstructed input	Reconstructed output	
Hypertuning	p_initialTrain	2101 reflectance bands (400-2500 nm), *SZA*, *RAA*	*LAI*, *Cab*, *Cm*, *Cw*	Loss
SE1	p_train1, p_test	2101 reflectance bands (400-2500 nm), *SZA*, *RAA*		Loss, AE, RMSE
SE2	p_train1, p_test	351 reflectance bands (400-750 nm), *SZA*, *RAA*		
SE3	p_train1, p_test	351 reflectance bands (750-1100 nm), *SZA*, *RAA*		

Experiment name	Original synthetic dataset	Reconstructed synthetic dataset		Evaluation metric
	Original Dataset	Reconstructed input	Reconstructed output	
SE4	p_train1, p_test	1401 reflectance bands (1100-2500 nm), *SZA*, *RAA*		
SE5	p_train1, p_test	701 reflectance bands (400-1100 nm), *SZA*, *RAA*		
SE6	p_train1, p_test	1751 reflectance bands (750-2500 nm), *SZA*, *RAA*		
SE7	p_train1, p_test	701 reflectance bands (400-1100 nm)		
SE8	ap_train1, ap_test	701 reflectance bands (400-1100 nm)		
SE9	ap_train2, ap_test	701 reflectance bands (400-1100 nm)		

Notes: The description of original synthetic datasets refers to Table 5 and the meaning of *SZA*, *RAA*, *LAI*, *Cab*, *Cm* and *Cw* refers to Table 1.

3. Results and Discussion

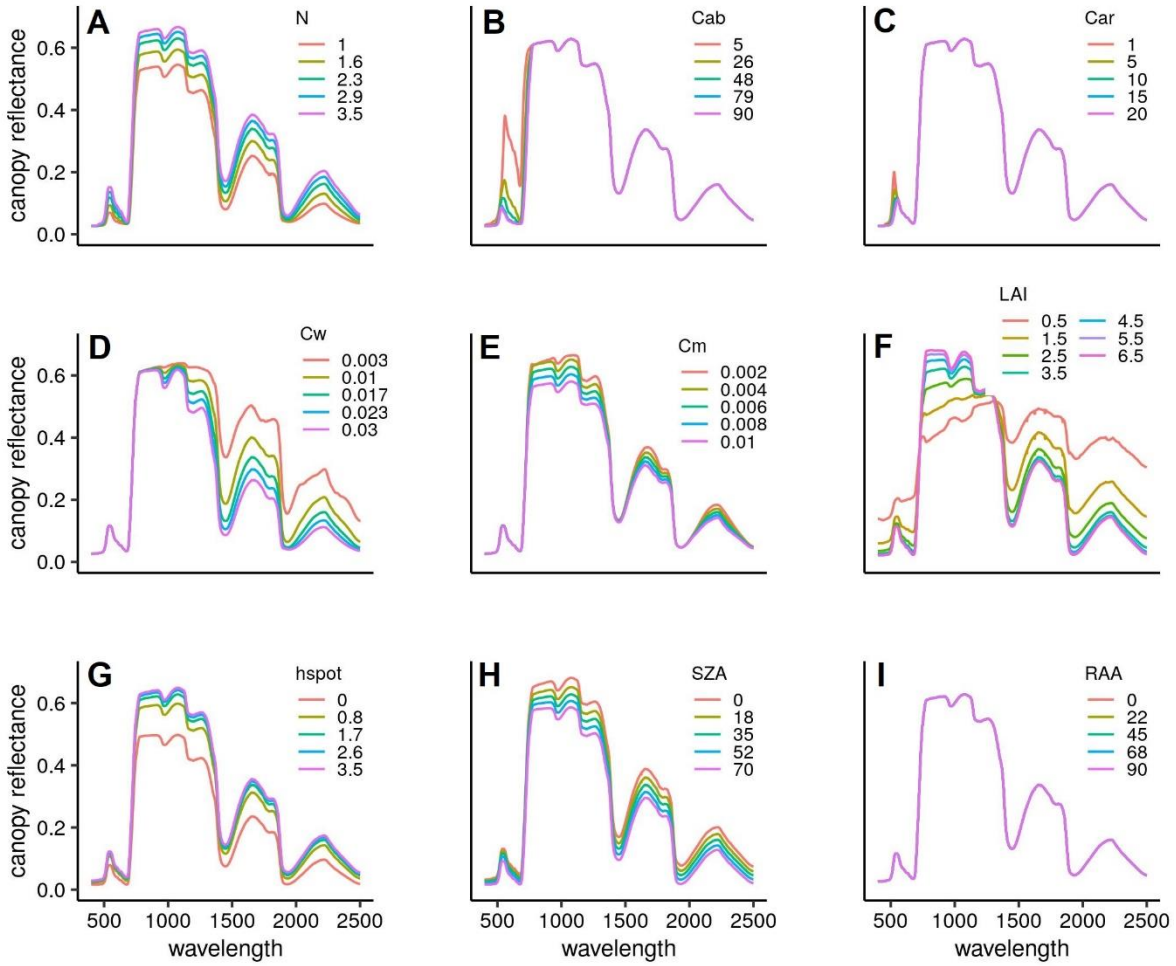
3.1 Sensitivity analysis of PROSAIL

By applying a local sensitivity analysis (OAT), we characterised a baseline situation for mid-season wheat crop and typical measurement scenario in order to demonstrate the impacts of independently varying input variables on canopy reflectance in 400-2500 nm range (Figure 4). An EFAST analysis qualified relative importance of each variable to total variability of reflectance across 400-2500 nm range in the entire input variable space (Figure 5).

In the OAT, the variables *N*, *hspot*, *Cm* and *SZA* had a consistent either negative or positive effect on reflectance across the whole range (although with small effect of *Cm* in visible region) (Figure 4). Increasing *N* and *hspot* resulted in increases in canopy reflectance (Figure 4A, G) while increasing *Cm* and *SZA* resulted in decreases in canopy reflectance (Figure 4E, H). Increases of two pigment variables *Cab* and *Car* only decreased reflectance in the visible region (Figure 4B, C) where Cab accounted for more than 60% of total variability around 560 nm and

358 710 nm while influence of *Car* was much less and only in 400-550 nm range (Figure 5).
359 Increasing *Cw* only decreased reflectance in 750-2500 nm range (Figure 4D) and its influence
360 appeared larger at wavelength > 1200 nm (Figure 5), making it become another important driver
361 except for *LAI* in this range. Compared with variables mentioned above, influence of *LAI* was
362 more important and complicated: contributing more than 60% of the total variability especially
363 in range of 400-500 nm, 750-1200 nm as well as 1400-2500 nm (Figure 5) and its increases
364 only increased reflectance in the range around 750-1350 nm but decreased reflectance in the
365 remaining range (Figure 4F). In particular, the contribution of *RAA* was negligible and it only
366 influenced reflectance through interaction with other variables (Figure 5).

367 Overall, our sensitivity analysis indicated that *LAI* dominated variability of canopy reflectance
368 across 400-2500 nm spectral range, while *Cab* and *Cw* played a key role only at visible range
369 (400-750 nm) and shortwave infrared range (1100-2500 nm), respectively. Similar findings
370 could also be found in other publications (e.g., Danner et al., 2019; Verrelst and Rivera, 2017),
371 although the magnitude of relative contribution of these variables were slightly different due to
372 various method or variables used for sensitivity analysis. The dominance of *LAI* across the
373 whole range is realised via its leaf elements which results in *LAI*. *LAI* indirectly controls the
374 soil reflectance propagating to canopy in low ground cover and controls light absorption in
375 different ranges via leaf optical properties (pigment and water content). Nevertheless, the
376 influence from soil background can be neglected for a canopy with LAI > 3 (Atzberger et al.,
377 2003) in which situation canopy generally reaches a high ground cover (Ramirez-Garcia et al.,
378 2012). The fact that increasing *LAI* decreased total reflectance in 400-750 nm and 1100-2500
379 nm indicates that photosynthetic pigments (chlorophyll and carotenoid) strongly absorb visible
380 light at wavelengths around 450 nm and 680 nm with leaf water having an influence on light
381 absorption coefficient at wavelength > 1200 nm (Feret et al., 2008).

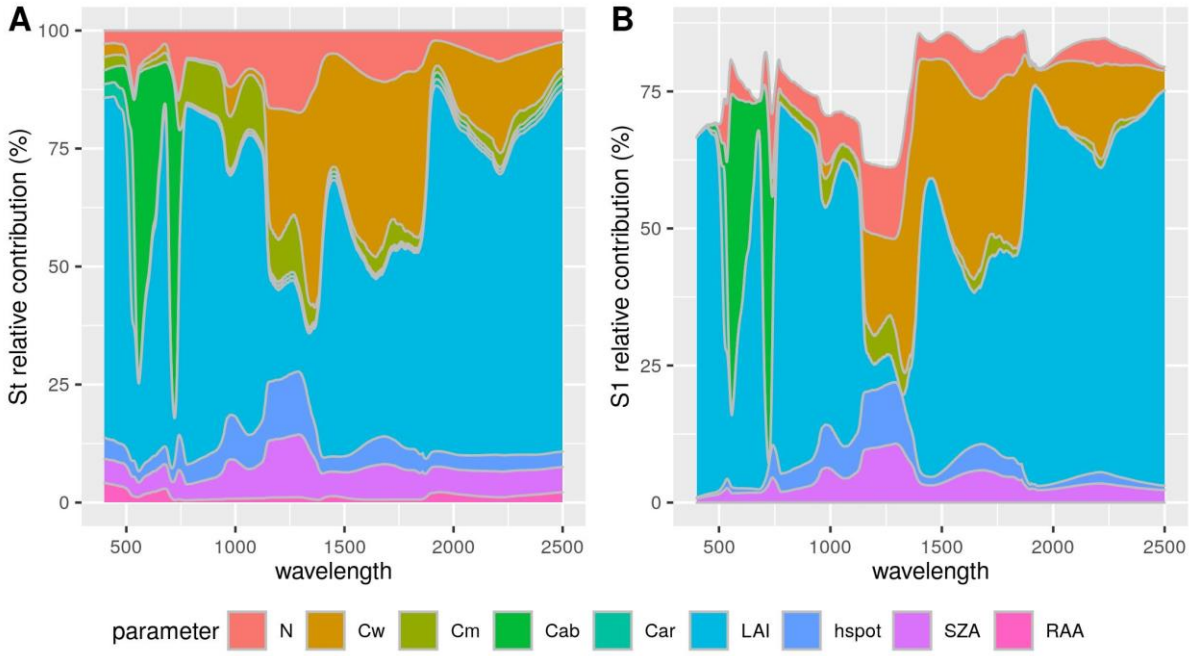


382

383 *Figure 4 Response of canopy reflectance to change of each parameter. Except for the study parameter changing at different*

384 *levels, the others are fixed to the baseline according to setting of OAT in Table 3 (Baseline: N=2.25, Cab=48, Car=10, Cant=0,*

385 *Cbrown=0, psoil=1,Cw=0.017, Cm=0.006, LAI=3.5, ALA=50, hspot=1.7, SZA=35, VZA=0, RAA=45).*



386

387 *Figure 5 Relative contribution of total effect (A) and first-order effect (B) for each input parameter on canopy reflectance*
388 *computed by PROSAIL using EFAST global sensitivity analysis*

389 **3.2 PROSAIL inversion based on FFNN can predict crop traits from spectral reflectance**

390 To investigate the precision of PROSAIL inversion for variable retrieval, we applied FFNN as
391 model inversion method to retrieve target variables from PROSAIL dataset (p_train1) which
392 was generated by PROSAIL with samples from full parameter space for arbitrary combination
393 of input variables (Figure 1). Performance of the trained FFNN model for target variable
394 retrieval from canopy reflectance in 400-2500 nm range and observation geometry information
395 (here is *SZA* and *RAA*) is reported in Table 7 (results of SE1), Figure 6 and Figure 7. The small
396 values of MAE and RMSE show the trained FFNN reached high precision for joint estimation
397 of *LAI*, *Cab*, *Cm* and *Cab* when the input variables were allowed to explore all combinations
398 across valid physiological ranges. However, the uncertainty range of estimations were
399 occasionally large, for instance, the absolute error of 95% samples for *Cab* estimation was
400 within 2.19 µg cm^{-2} while the maximum error was more than 10 times bigger up to 26.92 µg
401 cm^{-2} (see SE1 in Table 7). The same issue occurred to estimation for the other three variables.

Additionally, the trained FFNN continued to make good estimations at different levels of the true value, although larger true values tended to have a higher probability of larger absolute errors (Figure 6).

The trained model might not perform as well as presented here when it was applied to retrieve variables from real observation data due to measurement and model uncertainties. Thus, it would be unfair to compare our results with those from real observation data and should be more reasonable to compare with those also from simulation data. As presented in Table 8, our results were favourable, with approximately 10 times smaller RMSE for estimation of *LAI*, *Cab*, *Cm* and *Cw*, compared with simulated results from other model inversion studies. Reasons for such a good estimation exhibited in our research are likely due to the use of better architecture and algorithm used in neural network as well as more complete information included in massive hyperspectral bands as other studies inverted variables from only a few broad/narrow bands (Atzberger, 2004; Baret et al., 2007; Upreti et al., 2019) or derived VIs (le Maire et al., 2008; M. Xu et al., 2019). This result highlights the advantages of new deep learning techniques applied to high spectral resolution hyperspectral data and, demonstrate that it is viable to use deep learning approach to invert hyperspectral data to retrieve variables.

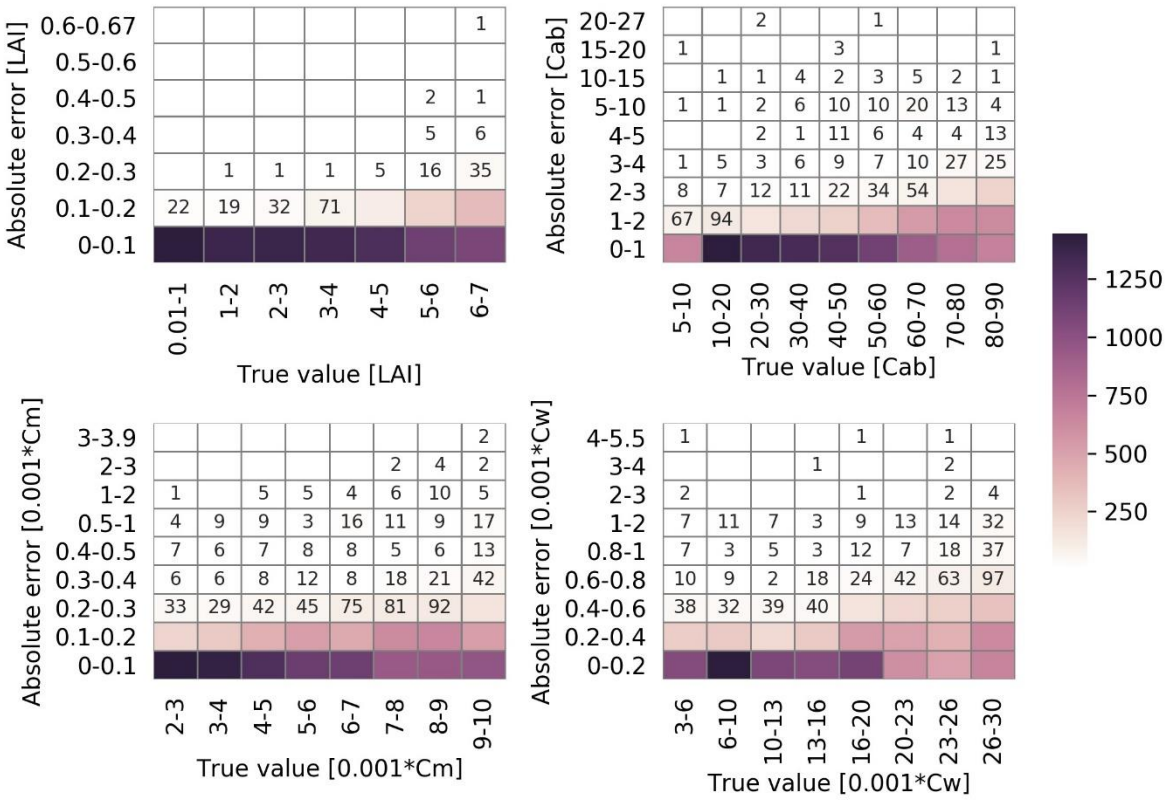
418 *Table 7 Statistical information of model's prediction error for each variable in each experiment*

Experiment Name	Crop Trait	min AE	5% AE	25% AE	50% AE	75% AE	95% AE	max AE	std AE	MAE	RMSE
SE1	*LAI*	0.000	0.003	0.014	0.031	0.060	0.123	0.666	0.042	0.044	0.061
	Cab	0.00	0.07	0.32	0.63	1.10	2.19	26.92	1.12	0.88	1.42
	Cm	0.000000	0.000007	0.000034	0.000072	0.000122	0.000242	0.003911	0.000146	0.000099	0.000176
	Cw	0.000000	0.000014	0.000071	0.000153	0.000291	0.000583	0.005438	0.000235	0.000216	0.000319
SE2	*LAI*	0.000	0.006	0.027	0.057	0.106	0.246	1.810	0.095	0.083	0.126
	Cab	0.00	0.09	0.45	0.96	1.67	3.20	41.42	1.39	1.25	1.87
	Cm	0.000000	0.000012	0.000067	0.000143	0.000265	0.000561	0.004661	0.000251	0.000206	0.000324
	Cw	0.000000	0.000100	0.000515	0.001215	0.002872	0.011521	0.028487	0.003829	0.002675	0.004671
SE3	*LAI*	0.000	0.006	0.031	0.067	0.127	0.311	1.382	0.112	0.101	0.151
	Cab	0.00	0.10	0.54	1.23	2.38	5.87	55.44	2.70	1.94	3.33
	Cm	0.000000	0.000012	0.000063	0.000142	0.000274	0.000620	0.003094	0.000224	0.000208	0.000306
	Cw	0.000000	0.000036	0.000186	0.000412	0.000811	0.002139	0.013600	0.000876	0.000671	0.001103
SE4	*LAI*	0.000	0.005	0.027	0.060	0.107	0.214	1.105	0.080	0.080	0.113
	Cab	0.00	0.48	2.33	5.55	12.20	33.66	87.15	11.02	9.53	14.57
	Cm	0.000000	0.000009	0.000042	0.000095	0.000175	0.000388	0.003169	0.000159	0.000136	0.000209
	Cw	0.000000	0.000023	0.000124	0.000263	0.000483	0.001086	0.007668	0.000416	0.000377	0.000561
SE5	*LAI*	0.000	0.005	0.024	0.051	0.104	0.235	0.828	0.077	0.077	0.109
	Cab	0.00	0.04	0.22	0.49	0.92	1.97	32.71	0.89	0.70	1.13
	Cm	0.000000	0.000007	0.000037	0.000076	0.000129	0.000268	0.003600	0.000126	0.000101	0.000161
	Cw	0.000000	0.000027	0.000139	0.000300	0.000519	0.001162	0.009210	0.000557	0.000426	0.000701
SE6	*LAI*	0.000	0.003	0.016	0.034	0.062	0.129	0.558	0.045	0.046	0.064
	Cab	0.00	0.07	0.34	0.76	1.53	5.30	50.26	3.14	1.57	3.51
	Cm	0.000000	0.000007	0.000034	0.000074	0.000136	0.000304	0.003580	0.000165	0.000111	0.000199
	Cw	0.000000	0.000016	0.000080	0.000170	0.000319	0.000728	0.009617	0.000329	0.000255	0.000416
SE7	*LAI*	0.000	0.004	0.020	0.044	0.080	0.180	1.093	0.063	0.061	0.087

Experiment Name	Crop Trait	min AE	5% AE	25% AE	50% AE	75% AE	95% AE	max AE	std AE	MAE	RMSE
	Cab	0.00	0.05	0.24	0.51	0.93	2.28	42.17	1.72	0.85	1.92
	Cm	0.000000	0.000007	0.000037	0.000077	0.000141	0.000383	0.005052	0.000267	0.000135	0.000299
	Cw	0.000000	0.000021	0.000108	0.000231	0.000459	0.001689	0.020608	0.001164	0.000511	0.001271
SE8	*LAI*	0.000	0.000	0.002	0.004	0.006	0.012	0.041	0.004	0.005	0.006
	Cab	0.00	0.01	0.03	0.06	0.13	0.78	6.89	0.41	0.18	0.45
	Cm	0.000000	0.000001	0.000006	0.000012	0.000019	0.000065	0.000549	0.000034	0.000020	0.000039
	Cw	0.000000	0.000003	0.000016	0.000034	0.000062	0.000141	0.001721	0.000054	0.000048	0.000072
SE9	*LAI*	0.000	0.000	0.001	0.003	0.004	0.008	0.031	0.003	0.003	0.004
	Cab	0.00	0.00	0.02	0.05	0.11	0.28	3.14	0.19	0.10	0.22
	Cm	0.000000	0.000000	0.000002	0.000005	0.000008	0.000025	0.000181	0.000012	0.000008	0.000014
	Cw	0.000000	0.000002	0.000013	0.000026	0.000044	0.000083	0.000317	0.000026	0.000032	0.000041

Notes: For experiments SE1 to SE7, models were evaluated with dataset p_test, while models in experiments SE8 and SE9 were evaluated with dataset ap_test. Both p_test and ap_test contains 10 000 examples. Values with '% 'symbol (i.e. 5%, 25%, 50%, 75%, 95%) represent absolute error (AE) at corresponding quantile level. 'min AE', 'max AE' and 'std AE' represents the minimum, maximum, standard deviation of AE, respectively. MAE and RMSE represents mean absolute error and root mean squared error, respectively. Unit of AE (MAE, RMSE) for LAI, Cab, Cm, Cw is m^2 m^{-2}, µg cm^{-2}, g cm^{-2}, and g cm^{-2}, respectively.

419



420

421 *Figure 6 Heatmap of model performance at different levels (SE1). Numbers in cells of the figure show the number of samples*
422 *whose prediction absolute errors of true values in a range were within a specific range. Numbers less than 100 are shown to*
423 *differ small values from zero. The unit of absolute error (true value) of LAI, Cab, Cm and Cw is* $m^2\ m^{-2}$, $\mu g\ cm^{-2}$, $g\ cm^{-2}$ *and g*
424 cm^{-2}, *respectively.*

425 *Table 8 Root mean squared error (RMSE) of target variable estimation from this and other inversion studies*

LAI ($m^2\ m^{-2}$)	*Cab* ($\mu g\ cm^{-2}$)	*Cm* ($g\ cm^{-2}$)	*Cw* ($g\ cm^{-2}$)	Inversion method	Reference
0.81 (0.59)	9.94 (7.84)	/	0.0037 (0.0028)	Neural network	Atzberger, 2004
1.10	/	/	/	Neural network	Baret et al., 2007
1.31	9.84	0.001414	/	VI empirical regression	le Maire et al., 2008
/	10.5 (7.9)	/	/	VI-based look-up table	M. Xu et al., 2019
0.99	11.4	/	/	Bagging trees	
1.04	11.02	/	/	Neural network in ARTMO	
1	11.55	/	/	Random forest tree bagger	Upreti et al., 2019
1.18	11.16	/	/	Partial least square regression	

LAI (m^2 m^{-2})	*Cab* (μg cm^{-2})	*Cm* (g cm^{-2})	*Cw* (g cm^{-2})	Inversion method	Reference
1.18	11.15	/	/	Least square linear regression	
1.14	11.61	/	/	Boosting trees	
1.3	14.99	/	/	Regression trees	
1.4	16.03	/	/	Random forest fit ensemble	
0.94	10.55	/	/	Gaussian process regression	
0.061	1.42	0.000176	0.000319	Feedforward neural network	SE1 in Table 7
0.087	1.92	0.000299	0.001271		SE7 in Table 7

Notes: Figures in brackets are the RMSEs after improvement by using a more novel estimation approach. In Atzberger's study, an object-based method was introduced to add spatial constraints on variables. In the study of M. Xu et al, a matrix-based method was used to constrain co-distribution of *LAI* and *Cab*.

426

427 **3.3 FFNN could still achieved compatible prediction after reducing wavelength bands and**
428 **excluding observation geometry in model input**

429 We have demonstrated trained FFNN can well predict crop traits from input features consisting
430 of spectral reflectance in 400-2500 nm range and observation configuration (solar and viewing
431 angles) and in this section, we explored the possibility of hyperspectral bands reduction and
432 effect of observation configuration in order to simplify FFNN model (Figure 1).

433 There are two main reasons to investigate effects of the use of reflectance bands in different
434 ranges for variable retrieval. Firstly, most sensors used currently only cover part of the whole
435 400 to 2500 nm wavelength range. Secondly, a large number of reflectance bands used as input
436 to FFNN model results in a complicated model structure, which might make it less suitable for
437 practical application. Considering both the wavelength range classification and ranges covered
438 in current hyperspectral sensors, the model was retrained using five sets of wavelength ranges:
439 400-750 nm, 750-1100 nm, 1100-2500 nm, 400-1100 nm, and 750-2500 nm. Performance of

trained FFNN using reflectance bands in these ranges alone are shown in Table 6 (results of SE 2 to S6) and Figure 7. It is clear that reflectance bands in 400-750 nm and 1100-2500 nm were not suitable for joint estimation of target variables (Figure 7). That is because the use of reflectance bands in these two ranges produced poor estimation for *Cab* and *Cw*, respectively, in which ranges canopy reflectance was insensitive to variation of *Cab* or *Cw* as demonstrated in Section 3.1. For joint retrieval of four target variables, the range of 400-1100 nm could be used as an alternative to replace the whole range for the aim of model simplification: the use of reflectance bands in this narrower range provided nearly same good estimation for *Cab* and *Cm* but less precise estimation for *LAI* and *Cw* by comparing SE1 and SE5.

The reason to attempt to exclude observation configuration (solar and viewing angles) from input of FFNN model is based on two considerations. One consideration is about information necessity and redundancy: the influence of observation configuration on canopy reflectance is likely to be implicitly reflected in hundreds of reflectance bands, so the observation configuration is unnecessary if a large number of reflectance bands are used in FFNN inputs. The other consideration is about information acquisition and availability: observation condition (solar and viewing angles) varies across pixels for an image, so observation configuration of each pixel must be calculated alone from this pixel's latitude, DOY and day time for phenotyping this pixel if precise retrieval using this method is expected at pixel-level. However, the gaining of observation configuration for every pixel is normally inaccessible especially for an image with large cover consisting of massive number of pixels. By comparing results of SE5 (400-1100 nm with solar/viewing angles) and SE7 (400-1100 nm without angles), it shows that exclusion of observation configuration in input of FFNN model just slightly decreased estimation precision for retrieval of *Cab*, *Cm* and *Cw* and even slightly increased overall estimation for *LAI* (RMSE was 0.109 μg cm^{-2} for SE5 and 0.087 μg cm^{-2} for SE7). Regardless, the results show that the trained FFNN model with input only including reflectance bands in

400-1100 nm range without angles still produced much smaller RMSE than those in previous model inversion studies for variable retrieval (see Table 8).

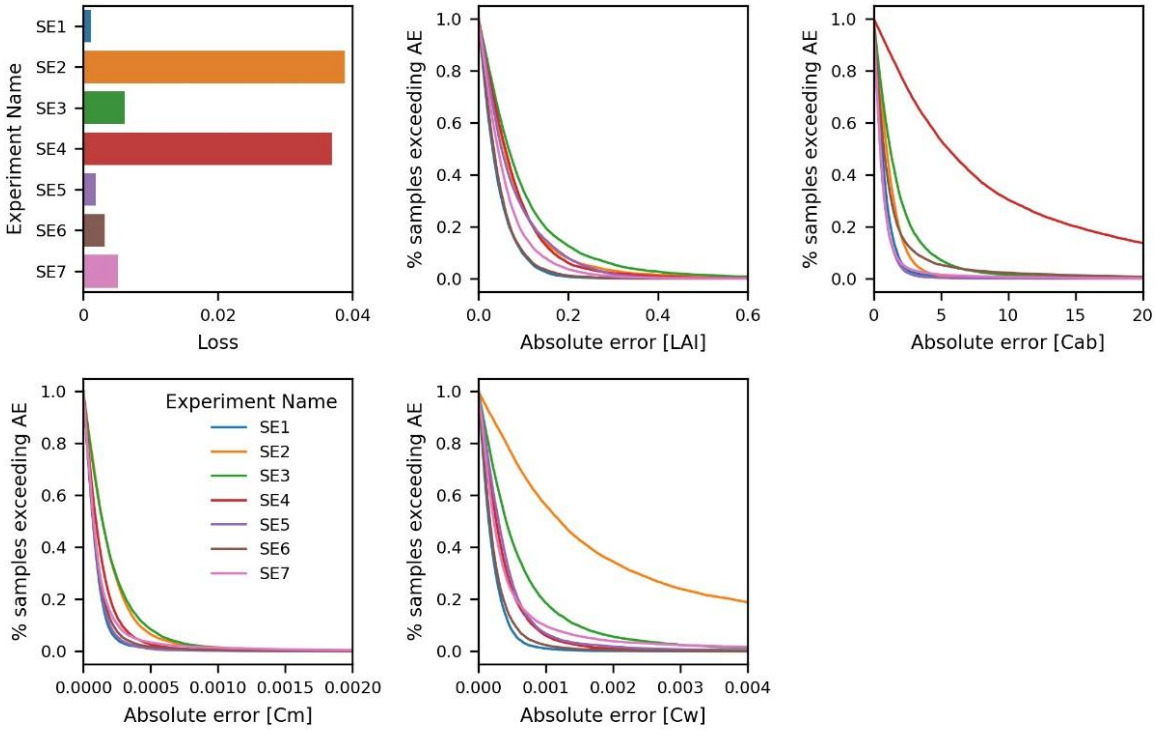


Figure 7 Total performance and its empirical distribution for models with different input features. Loss is a unitless indicator and represents the total mean squared error of joint estimation of four model output variables after normalization. Absolute error (AE) is the difference between the true value of each variable and its prediction after de-normalization. The unit of AE of LAI, Cab, Cm and Cw is m^2 m^{-2}, μg cm^{-2}, g cm^{-2} and g cm^{-2}, respectively.

3.4 Coupling APSIM and PROSAIL can further improve FFNN's prediction precision

While the FFNN model was able to perform well when trained across sets of variables that explore the entire ranges of realistic values, the key part of this paper was to investigate the impact of constraining the input variables to be limited to 'physiologically realistic' combinations (Figure 1). Figure 8 indicates that the use of APSIM-PROSAIL dataset (ap_train1 and ap_train2) generated by coupling APSIM and PROSAIL significantly improved FFNN's performance for estimation of all target variables using reflectance bands in 400-1100 nm without observation configuration (solar/viewing angles). This improvement of model

performance presented in both average precision (smaller MAE and RMSE) and stability (narrower uncertainty range, smaller standard deviation of AE) according to statistical results in Table 7. For instance, compared with results using PROSAIL dataset (p_train1, SE7), the use of APSIM-PROSAIL dataset (ap_train1, SE8) for *LAI* estimation narrowed the uncertainty range of AE from 0~1.093 m^2 m^{-2} to 0~0.041 m^2 m^{-2} and also reduced standard deviation of AE (from 0.063 to 0.004 m^2 m^{-2}), MAE (from 0.061 to 0.005 m^2 m^{-2}) and RMSE (from 0.087 to 0.006 m^2 m^{-2}). The magnitude of improvement resulting from model integration is much larger than that resulting from adding spatial constraints on related variables as demonstrated in the study of Atzberger (2004) where RMSE reduced from 0.81 to 0.59 m^2 m^{-2} for LAI estimation as exhibited in Table 8. It deserves to emphasize that this improvement is irrelevant to the number of samples used for training (sample size of p_train1 and ap_train1 is the same) and increasing sample size of training set tended to improve model performance in further (sample size of ap_train2 (2 139 226) is larger than ap_train1 (90 000)).

The success of this integration in building improved inverse models seems to lie in providing a higher quality training dataset as FFNN totally learns from data provided. CGMs are powerful to predict crop growth and development characterised by a series of physiological processes, thus using CGM to produce sets of input variables of RTM can create a distribution and co-distribution of related variables closer to real situations defining canopy architecture and potentially remove cases that can be simulated but may not be actually observed (sampled) in a real-world from training dataset. Figure 2 clearly shows difference of variable distribution in two types of training dataset whose input parameters vary in the same range. It should be noticed that the difference of distributions projected in low-dimensional space of two datasets can indicate that their distributions projected in high-dimensional space must be different, but that a similarity between distributions in low-dimensional space does not guarantee that their distributions in high-dimensional space are also the same.

The continuous improvement in model performance as sample size increased, it is likely associated to a richer diversity of samples rather than growing number of same samples. Adding more existing samples in training dataset does not enhance learning of relationships between variables within a network. It is reported that a uniform distribution of the variables may obtain a more even distribution of the uncertainties in spite of a poor RMSE (Baret and Buis, 2008). However, our results show a non-uniform distribution closer to true distribution (variables in APSIM-PROSAIL dataset) can also obtain a relative even uncertainty distribution (see Figure S6 and Figure S7) and a better RMSE (see SE8 and SE9 in Table 8) compared with results from PROSAIL dataset with uniform distribution variables (see Figure 6 and SE7 in Table 8). A small consequence is that there is a higher probability of larger error when estimating larger true value. Overall, these findings offered evidence to support that integrating CGM and RTM is a promising way for variable retrieval from hyperspectral data in theoretical dimensionality.

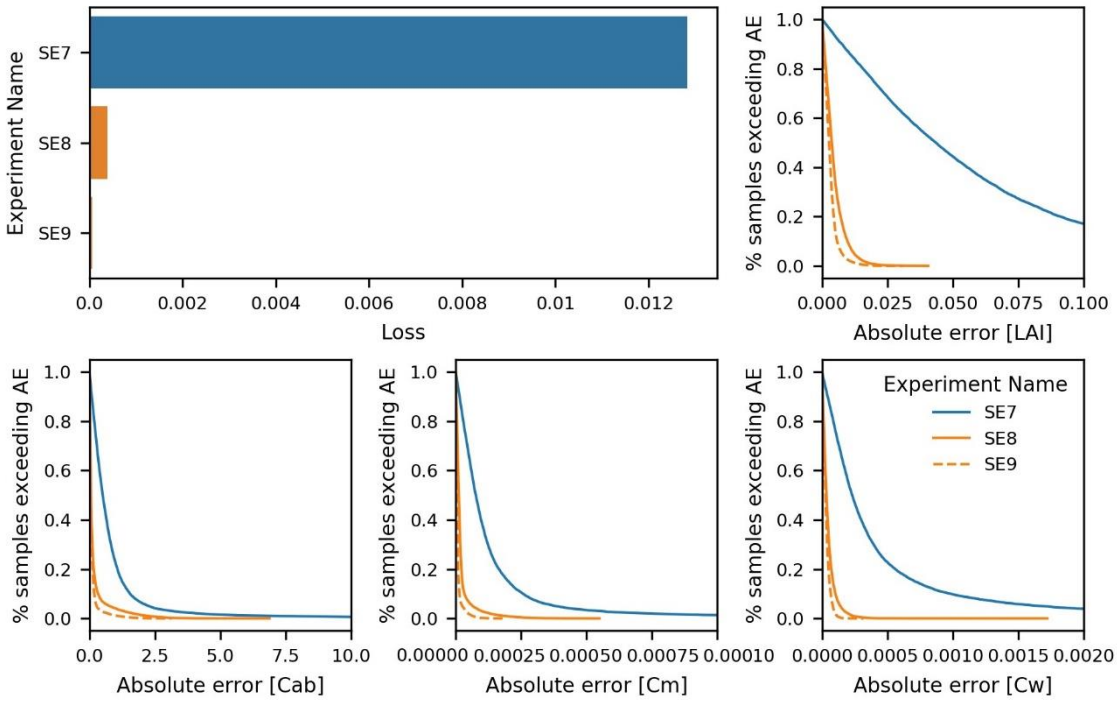


Figure 8 Total performance and its empirical distribution for models trained using different datasets. Loss is a unitless indicator and represents the total mean squared error of joint estimation of four model output variables after normalization.

Absolute error (AE) is the difference between the true value of each variable and its prediction after de-normalization. The unit of AE of LAI, Cab, Cm and Cw is m^2 m^{-2}, μg cm^{-2}, g cm^{-2} and g cm^{-2}, respectively.

4. Conclusion

In general, a completed workflow of integrating APSIM and PROSAIL was demonstrated in this research, resulting in an applicable coupling APSIM-PROSAIL model. Additionally, this research also illustrated the process of choosing appropriate hyperparameter for FFNN model and presented the advantages of using FFNN for crop traits retrieval. It was also demonstrated that the model could be applied to subsets of different wavelength ranges in order to suit different types of instrumentation and applications. However, the major contribution of this research is to demonstrate a practical way to generate higher quality training data which can better characterise the real canopy realization and, prove its practicability in theory. Although this trained FFNN model might not perform as well as presented here when it is applied to retrieve variables from real observation data due to measurement and model uncertainties, it is expected to be able to make relative good performance according to difference of estimation precision on simulated and observed data from other model inversion studies. In future, we are going to investigate the performance of trained FFNN model using simulation data in different simulated situations as well as practical production environments.

Author's responsibility

QC implemented this research and wrote the draft of the manuscript. SC helped formulate research ideas and BZ provided guidance of use of APSIM Next Generation model and outlined the framework of this paper with QC. TC provided useful technical advices in the application of the deep learning models. All authors engaged in reviewing and editing this manuscript.

Acknowledgement

QC is receiving a scholarship co-funded by the China Scholarship Council and the University of Queensland and has been supported by CSIRO as a visiting student.

References

Atzberger, C., 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. Remote Sens. Environ. 93, 53–67. https://doi.org/10.1016/j.rse.2004.06.016

Atzberger, C., Jarmer, T., Schlerf, M., et al., 2003. Retrieval of wheat bio-physical attributes from hyperspectral data and SAILH+ PROSPECT radiative transfer model. 3rd EARSeL Work. Imaging Spectrosc. 473–482.

Bacour, C., Baret, F., Béal, D., et al., 2006. Neural network estimation of LAI, fAPAR, fCover and LAI×Cab, from top of canopy MERIS reflectance data: Principles and validation. Remote Sens. Environ. 105, 313–325. https://doi.org/10.1016/j.rse.2006.07.014

Bacour, C., Jacquemoud, S., Leroy, M., et al., 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

Baret, F., Buis, S., 2008. Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems. Adv. L. Remote Sens. Syst. Model. Invers. Appl. 173–201. https://doi.org/10.1007/978-1-4020-6450-0_7

Baret, F., Hagolle, O., Geiger, B., et al., Weiss, M., Samain, O., Roujean, J.L., Leroy, M., 2007. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm. Remote Sens. Environ. 110, 275–286. https://doi.org/10.1016/j.rse.2007.02.018

Berger, K., Atzberger, C., Danner, M., et al., 2018. Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: A review study. Remote Sens. 10. https://doi.org/10.3390/rs10010085

568 Campos-Taberner, M., Moreno-Martínez, Á., García-Haro, F.J., et al., 2018. Global estimation
569 of biophysical variables from Google Earth Engine platform. Remote Sens. 10, 1–17.
570 https://doi.org/10.3390/rs10081167

571 Chapman, S.C., Zheng, B., Andries, B., et al., 2018. Visible, Near Infrared, and Thermal
572 Spectral Radiance on-Board UAVs for High-Throughput Phenotyping of Plant Breeding
573 Trials, in: Biophysical and Biochemical Characterization and Plant Species Studies. pp.
574 275–299.

575 Chenu, K., Deihimfard, R., Chapman, S.C., 2013. Large-scale characterization of drought
576 pattern: A continent-wide modelling approach applied to the Australian wheatbelt - spatial
577 and temporal trends. New Phytol. 198, 801–820. https://doi.org/10.1111/nph.12192

578 Chollet, F., 2017. Deep Learning with Python, Deep Learning with Python.
579 https://doi.org/10.1007/978-1-4842-2766-4

580 Combal, B., Baret, F., Weiss, M., et al., 2003. Retrieval of canopy biophysical variables from
581 bidirectional reflectance using prior information to solve the ill-posed inverse problem.
582 Remote Sens. Environ. 84, 1–15. https://doi.org/10.1016/S0034-4257(02)00035-4

583 Cybenkot, G., 1989. Mathematics of Control, Signals, and Systems Approximation by
584 Superpositions of a Sigmoidal Function. Math. Control Signals Syst. 2, 303–314.

585 Danner, M., Berger, K., Wocher, M., et al., 2019. Fitted PROSAIL parameterization of leaf
586 inclinations,water content and brown pigment content for winter wheat and maize
587 canopies. Remote Sens. 11. https://doi.org/10.3390/rs11101150

588 Dhakar, R., Sehgal, V.K., Chakraborty, D., et al., 2019. Field scale wheat LAI retrieval from
589 multispectral Sentinel 2A-MSI and LandSat 8-OLI imagery: effect of atmospheric
590 correction, image resolutions and inversion techniques. Geocarto Int. 0, 1–21.

591 https://doi.org/10.1080/10106049.2019.1687591

592 Dorigo, W.A., Zurita-Milla, R., de Wit, A.J.W., et al., 2007. A review on reflective remote
593 sensing and data assimilation techniques for enhanced agroecosystem modeling. Int. J.
594 Appl. Earth Obs. Geoinf. 9, 165–193. https://doi.org/10.1016/j.jag.2006.05.003

595 Eon, R.S., Goldsmith, S., Bachmann, C.M., et al., 2019. Retrieval of salt marsh above-ground
596 biomass from high-spatial resolution hyperspectral imagery using PROSAIL. Remote
597 Sens. 11, 1–18. https://doi.org/10.3390/rs11111385

598 Feret, J.B., François, C., Asner, G.P., et al., 2008. PROSPECT-4 and 5: Advances in the leaf
599 optical properties model separating photosynthetic pigments. Remote Sens. Environ. 112,
600 3030–3043. https://doi.org/10.1016/j.rse.2008.02.012

601 Féret, J.B., Gitelson, A.A., Noble, S.D., et al., 2017. PROSPECT-D: Towards modeling leaf
602 optical properties through a complete lifecycle. Remote Sens. Environ. 193, 204–215.
603 https://doi.org/10.1016/j.rse.2017.03.004

604 García-Haro, F.J., Campos-Taberner, M., Muñoz-Marí, J., et al., 2018. Derivation of global
605 vegetation biophysical parameters from EUMETSAT Polar System. ISPRS J.
606 Photogramm. Remote Sens. 139, 57–74. https://doi.org/10.1016/j.isprsjprs.2018.03.005

607 Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Feedforward Networks, Deep Learning.
608 The MIT Press, Cambridge, MA, USA.

609 Guo, C., Tang, Y., Lu, J., et al., 2019. Predicting wheat productivity: Integrating time series of
610 vegetation indices into crop modeling via sequential assimilation. Agric. For. Meteorol.
611 272–273, 69–80. https://doi.org/10.1016/j.agrformet.2019.01.023

612 Holzworth, D., Huth, N.I., Fainges, J., et al., 2018. APSIM Next Generation: Overcoming
613 challenges in modernising a farming systems model. Environ. Model. Softw. 103, 43–51.

614 https://doi.org/10.1016/j.envsoft.2018.02.002

615 Holzworth, D.P., Huth, N.I., deVoil, P.G., et al., 2014. APSIM - Evolution towards a new
616 generation of agricultural systems simulation. Environ. Model. Softw. 62, 327–350.
617 https://doi.org/10.1016/j.envsoft.2014.07.009

618 Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal
619 approximators. Neural Networks 2, 359–366. https://doi.org/10.1016/0893-
620 6080(89)90020-8

621 Huang, J., Ma, H., Sedano, F., et al., 2019. Evaluation of regional estimates of winter wheat
622 yield by assimilating three remotely sensed reflectance datasets into the coupled
623 WOFOST–PROSAIL model. Eur. J. Agron. 102, 1–13.
624 https://doi.org/10.1016/j.eja.2018.10.008

625 Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra.
626 Remote Sens. Environ. 34, 75–91. https://doi.org/10.1016/0034-4257(90)90100-Z

627 Jacquemoud, S., Verhoef, W., Baret, F., et al., 2009. PROSPECT + SAIL models: A review of
628 use for vegetation characterization. Remote Sens. Environ. 113, S56–S66.
629 https://doi.org/10.1016/j.rse.2008.01.026

630 Jay, S., Baret, F., Dutartre, D., et al., Weiss, M., Maupas, F., 2019. Exploiting the centimeter
631 resolution of UAV multispectral imagery to improve remote-sensing estimates of canopy
632 structure and biochemistry in sugar beet crops. Remote Sens. Environ. 231, 0–1.
633 https://doi.org/10.1016/j.rse.2018.09.011

634 le Maire, G., François, C., Soudani, K., et al., 2008. Calibration and validation of hyperspectral
635 indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area,
636 leaf area index and leaf canopy biomass. Remote Sens. Environ. 112, 3846–3864.

637 https://doi.org/10.1016/j.rse.2008.06.005

638 Lin, M., Chen, Q., Yan, S., 2014. Network in network. 2nd Int. Conf. Learn. Represent. ICLR
639 2014 - Conf. Track Proc. 1–10.

640 Lunagaria, M.M., Patel, H.R., 2019. Evaluation of PROSAIL inversion for retrieval of
641 chlorophyll, leaf dry matter, leaf angle, and leaf area index of wheat using
642 spectrodirectional measurements. Int. J. Remote Sens. 40, 8125–8145.
643 https://doi.org/10.1080/01431161.2018.1524608

644 Ramirez-Garcia, J., Almendros, P., Quemada, M., 2012. Ground cover and leaf area index
645 relationship in a grass, legume and crucifer crop. Plant, Soil Environ. 58, 385–390.
646 https://doi.org/10.17221/195/2012-pse

647 Roosjen, P.P.J., Brede, B., Suomalainen, J.M., et al., 2018. Improved estimation of leaf area
648 index and leaf chlorophyll content of a potato crop using multi-angle spectral data –
649 potential of unmanned aerial vehicle imagery. Int. J. Appl. Earth Obs. Geoinf. 66, 14–26.
650 https://doi.org/10.1016/j.jag.2017.10.012

651 Saltelli, A., Tarantola, S., Chan, K.P.S., 1999. A quantitative model-independent method for
652 global sensitivity analysis of model output. Technometrics 41, 39–56.
653 https://doi.org/10.1080/00401706.1999.10485594

654 Shibayama, M., Watanabe, Y., 2007. Estimating the mean leaf inclination angle of wheat
655 canopies using reflected polarized light. Plant Prod. Sci. 10, 329–342.
656 https://doi.org/10.1626/pps.10.329

657 Thenkabail, P.S., Lyon, J.G., Huete, A., 2019. Fifty Years of Advances in Hyperspectral
658 Remote Sensing of Agriculture and Vegetation—Summary, Insights, and Highlights of
659 Volume II. Biophys. Biochem. Charact. Plant Species Stud. 303–341.

https://doi.org/10.1201/9781315159331-10

Thorp, K.R., Wang, G., West, A.L., et al., 2012. Estimating crop biophysical properties from remote sensing data by inverting linked radiative transfer and ecophysiological models. Remote Sens. Environ. 124, 224–233. https://doi.org/10.1016/j.rse.2012.05.013

Upreti, D., Huang, W., Kong, W., et al., 2019. A comparison of hybrid machine learning algorithms for the retrieval of wheat biophysical variables from sentinel-2. Remote Sens. 11. https://doi.org/10.3390/rs11050481

Verhoef, W., 1998. Theory of radiative transfer models applied in optical remote sensing of vegetation canopies. Wageningen Agric. Univ. https://doi.org/ISBNL 90-5485-804-4

Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. Remote Sens. Environ. 16, 125–141. https://doi.org/10.1016/0034-4257(84)90057-9

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., et al., 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

Verrelst, J., Rivera, J.P., 2017. A Global Sensitivity Analysis Toolbox to Quantify Drivers of Vegetation Radiative Transfer Models, Sensitivity Analysis in Earth Observation Modelling. Elsevier Inc. https://doi.org/10.1016/B978-0-12-803011-0.00016-1

Weiss, M., Baret, F., Myneni, R.B., et al., 2000. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. Agronomie 20, 3–22. https://doi.org/10.1051/agro:2000105

Wu, C., Liu, M., Liu, X., et al., 2019. Developing a new spectral index for detecting cadmium-induced stress in rice on a regional scale. Int. J. Environ. Res. Public Health 16.

683 https://doi.org/10.3390/ijerph16234811

684 Xia, Z., Peng, Y., Liu, S., et al., 2019. The optimal image date selection for evaluating cultivated
685 land quality based on gaofen-1 images. Sensors (Switzerland) 19.
686 https://doi.org/10.3390/s19224937

687 Xu, J., Meng, J., Quackenbush, L.J., 2019. Use of remote sensing to predict the optimal harvest
688 date of corn. F. Crop. Res. 236, 1–13. https://doi.org/10.1016/j.fcr.2019.03.003

689 Xu, M., Liu, R., Chen, J.M., et al., 2019. Retrieving leaf chlorophyll content using a matrix-
690 based vegetation index combination approach. Remote Sens. Environ. 224, 60–73.
691 https://doi.org/10.1016/j.rse.2019.01.039

692 Xu, X.Q., Lu, J.S., Zhang, N., et al., 2019. Inversion of rice canopy chlorophyll content and
693 leaf area index based on coupling of radiative transfer and Bayesian network models.
694 ISPRS J. Photogramm. Remote Sens. 150, 185–196.
695 https://doi.org/10.1016/j.isprsjprs.2019.02.013

696 Yu, F.H., Xu, T.Y., Du, W., et al., 2017. Radiative transfer models (RTMs) for field
697 phenotyping inversion of rice based on UAV hyperspectral remote sensing. Int. J. Agric.
698 Biol. Eng. 10, 150–157. https://doi.org/10.25165/j.ijabe.20171004.3076

699 Zhang, Y., Chan, W., Jaitly, N., 2017. Very deep convolutional networks for end-to-end speech
700 recognition. ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. 4845–4849.
701 https://doi.org/10.1109/ICASSP.2017.7953077

702 Zhu, W., Sun, Z., Huang, Y., et al., 2019. Improving field-scale wheat LAI retrieval based on
703 UAV remote-sensing observations and optimized VI-LUTs. Remote Sens. 11.
704 https://doi.org/10.3390/rs11202456

705