26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Integrating crop modelling and machine learning for non-destructive estimation of crop traits

Qiaomin Chen [a], **Pengcheng Hu** [b], **Bangyou Zheng** [c], **and Scott C. Chapman** [a]

[a] *School of Agriculture and Food Sustainability, The University of Queensland, Brisbane, 4067, Australia*

[b] *CSIRO Agriculture & Food, Canberra, 2601, Australia*

[c] *CSIRO Agriculture & Food, Brisbane, 4067, Australia*

Email: qiaomin.chen@uq.edu.au

Abstract: Accurate estimation of crop traits is critical for advancing plant phenotyping, improving crop management, and accelerating breeding efforts. However, current approaches often face challenges such as limited high-quality training data, poor transferability across environments, and a lack of biological realism in model predictions. In this study, we present a biologically informed modelling framework that integrates crop growth models, radiative transfer models, and machine learning to enhance crop trait estimation. This approach implicitly and/or explicitly incorporates prior knowledge of crop physiology and biophysical relationships to refine trait prediction, thereby improving model accuracy, robustness, and interpretability.

We first developed a conceptual framework (i.e., APSIM-PROSAIL coupling model) that synthesizes crop growth simulations with radiative transfer models to generate biologically constrained synthetic datasets for wheat crops. This allowed the inclusion of realistic trait covariations and physiological constraints in training data, which significantly improved the retrieval accuracy of key wheat traits such as leaf area index, chlorophyll content, dry matter, and water content. Deep learning models trained on this constrained dataset outperformed those trained on unconstrained data, particularly in transferring to real multispectral UAV imagery collected from wheat plots with diverse canopy structures.

Building on this, we implemented the APSIM-PROSAIL model to estimate aboveground biomass (AGB) dynamics at small plot scales using remote sensing data. By deriving radiative transfer parameters from APSIM-Wheat outputs, the model was able to simulate canopy reflectance and constrain the estimation of unknown environmental variables, such as initial soil water and nitrogen content, across tens of wheat variety trials in Australia. This biologically informed inversion framework enabled accurate reproduction of Sentinel-2 NDVI dynamics and retrieval of plot-level AGB, demonstrating the capacity to assess biomass variability and genotype–environment–management interactions at fine spatial scales using moderate-resolution satellite data.

To support rapid, scalable, and generalizable biomass estimation, we further developed a biophysics-informed neural network that integrates multispectral and daily weather data with embedded physiological knowledge. Unlike traditional models that require frequent recalibration, the neural network model provides consistent dynamic biomass predictions from sowing to harvest across years, locations, and genotypes. Validation using satellite and drone imagery confirmed the framework's effectiveness at both trial and plot scales, highlighting its practicality for multiscale phenotyping and crop monitoring.

Together, our integrated efforts establish a novel class of biologically informed modelling frameworks that bridge mechanistic understanding and data-driven learning. By embedding physiological constraints into machine learning models, we offer new tools to enhance the reliability and scalability of crop trait estimation from remote sensing, paving the way for smarter agricultural decision-making and more precise breeding strategies.

Keywords: *Above-ground biomass, crop growth model, leaf area index, machine learning, radiative transfer model, wheat*

