HI from the Sky : Estimating harvest index from UAVs combined with machine learning

Alexandra J. Burgess[1]

[1]Agriculture and Environmental Sciences, School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough, UK

alexandra.burgess@nottingham.ac.uk

Crop yields must increase in order to meet food security goals. Estimating crop yield- related traits accurately and rapidly is therefore of utmost importance, with applications in breeding, precision farming and yield monitoring. Hight throughput phenotyping (HTPP) refers to the use of non-invasive techniques to measure physiological and agronomical traits of crop plants. Optical remote sensing techniques are widely used in HTPP. They encompass data captured from areas of the visible (400–700 nm, captured using red green blue- RGB- sensors), near infrared (NIR, 700–1350 nm), red edge (680–730 nm), and shortwave infrared (SWIR; 1350–2500 nm) regions of the spectrum (Gamon et al., 2019). This spectral information has numerous applications within crop monitoring direct prediction of a trait of interest such as pigment content or biomass area, of through using spectral indices, otherwise known as vegetative indices, which correlate spectral reflectance with physiological traits such as green area, portrayed by the normalised difference vegetation index (NDVI) (Fu et al., 2020; Tattaris et al., 2016).

One of the most important components determining crop yield is the Harvest Index (HI); the ratio between harvestable yield and above ground biomass. However, relative to other yield-related traits, HI is poorly understood. This is partly a result of the fact that HI is traditionally measured via destructive field sampling, which is both time-consuming and labour intensive. Within many of the staple crops, such as wheat and rice, HI is approaching the theoretical maximum value. However, rapid screening of HI is valuable within breeding programmes and for the temporal evaluation of growth status.

As the name suggest, remote sensing via optical, or other, techniques, permits data to be collected from a distance, without the need for physical interaction with a plant. These sensors may be fixed or in motion. For example, remote sensing from satellite data has many uses within agricultural research, however in some instances, application is limited by low resolution and fixed measurement times. Therefore, an alternative approach for field data collection is through the use of unmanned aerial vehicles (UAVs). UAVs are able to provide high resolution spatial- and temporal- data for a relatively low cost. So far, UAV-based data has been used to estimate crop traits including seedling emergence, plant height, leaf area index, above ground biomass and yield (Tsouros et al., 2019).

Following collection, data analysis can be supported by a variety of techniques including machine learning approaches. Ensemble Learning (EL), refers to a subset of machine learning, whereby multiple learning algorithms are combined within a single framework to obtain higher predictive performance. EL methods are able to overcome problems associated with small training sets, such as overfitting, because outputs of independent base models are integrated through secondary learning methods (Zhang et al., 2019). This integration can be facilitated by Bayesian Model Averaging (BMA), in which the posterior probability of each of the basic models are taken as weights for the secondary learning step. As such, BMA is able to overcome uncertainty in the modelling process, leading to a higher estimation accuracy, and so has been widely applied for many fields (Shu et al., 2022).

Within this issue of *Plant Physiology,* Ji et al. (2023) used multi-spectral data captured from a UAV combined with machine learning to calculate harvest index of pea (*Pisum sativum* L.) and faba bean (*Vicia faba* L.) at four different growth stages. Ji et al. (2023) combined data collection from RGB, multispectral (MS) and thermal infrared (TIR) sensors, as well as four different machine learning models and an integrated EL model based on MBA (Figure 1). Combining multi-modal data provided a better overall estimate of HI compared to any single sensor with improvements in model fitting (i.e. R^2) of up to 18% and 31% for bean and pea, respectively. Similarly to spectral information, the combined EL model provided the best and most stable estimation of HI, and combining information from all four growth stages provided a better estimate compared to any single growth stage.

Together, the results of Ji et al. (2023) indicate the power of combining multi-sensor and multi-model information for the prediction of plant physiological traits. The improved HI prediction power reflects the results of previous studies which demonstrate how multi-sensor information can improve crop trait estimation (Feng et al., 2020; Li et al., 2022). However, species-specific differences were seen in the estimation accuracy of HI, with pea achieving more accurate predictions relative to bean. This improved HI estimation in pea was predicted to be a result of reduced canopy cover, and thus reduced saturation of the optical sensors. In some instances, combining information from two sensors yielded greater predictive power than three indicating possible data redundancy. Furthermore, in some cases, the base machine learning models performed better than the EL approach. Thus the optimal combination of sensors and models is likely to be both species-, growth stage- and trait- specific.

Despite the recent rise in machine learning approaches for plant science disciplines, its application for crop yield prediction is not yet viable for wide-scale use. Nevertheless, whilst improvements can still be made for the estimation of HI, the study of Ji et al. (2023) presents a further step in the generation of HTPP for monitoring crop performance. Combining low-cost sensor information and machine learning permits robust and accurate measurements of physiological traits, ultimately providing an acceleration in the breeding process. Future improvements will require large datasets encompassing multiple genotypes across a variety of sites, and assessment of the redundancy between multi-sensor data.

Figure Legends

***Figure 1*: *Overview of the remote sensing approach designed by Ji et al. (2023) for the estimation of Harvest Index (HI) of field grown faba bean (Vicia faba L.) and pea (Pisum sativum L.). Optical data was collected using a UAV from red green blue (RGB), multispectral (MS) and thermal infrared (TIR) sensors. Data was analysed using a variety of machine learning approaches singularly and in combination using Ensemble Learning (EL) and Bayesian Model Averaging (BMA). The most accurate estimation of HI was obtained using multi-modal data combined with the EL model.*

Acknowledgements

AB wishes to thank Plant Physiology for the opportunity to act as an Assistant Features Editor.

References

Feng, A., Zhou, J., Vories, E., Sudduth, K., Zhang, M., 2020. Yield estimation in cotton using UAV-based multi-sensor imagery. Biosyst Eng 193, 101–114.

Fu, P., Meacham-Hensold, K., Guan, K., Wu, J., Bernacchi, C., 2020. Estimating photosynthetic traits from reflectance spectra: a synthesis of spectral indices, numerical inversion, and partial least square regression. Plant Cell Environ 43, 1241–1258.

Gamon, J., Somers, B., Malenovský, Z., Middleton, E., Rascher, U., Schaepman, M., 2019. Assessing vegetation function with imaging spectroscopy. Surv Geophys 40, 489–513.

Ji, Y., Liu, Z., Cui, Y., Liu, R., Chen, Z., Zong, X., Yang, T., 2023. Faba bean and pea harvest index estimations using aerial-based multimodal data and machine learning algorithms. Plant Physiol kiad.

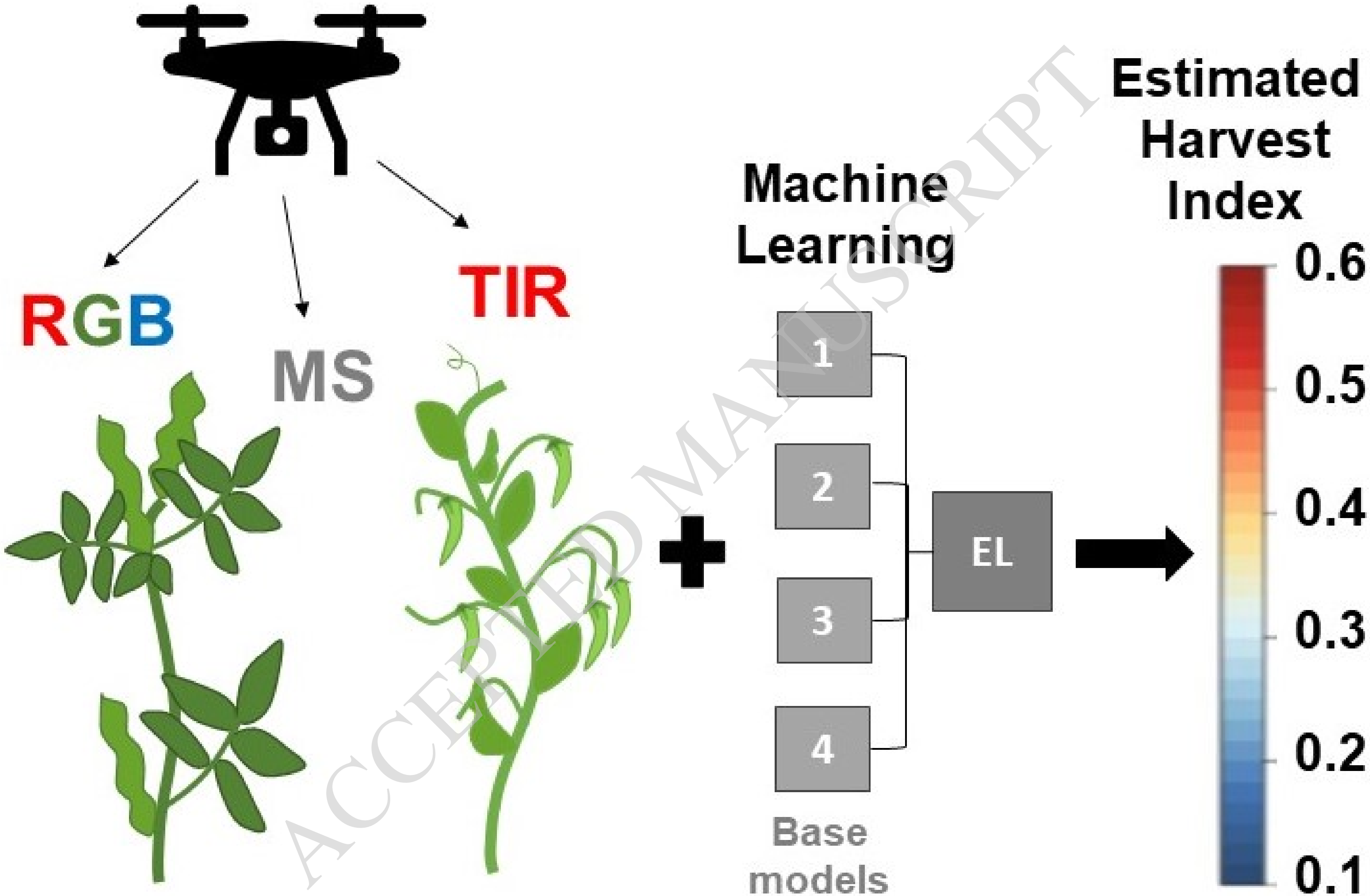
Li, J., Schachtman, D., Creech, C., Wang, L., Ge, Y., Shi, Y., 2022. Evaluation of UAV-derived multimodal remote sensing data for biomass prediction and drought tolerance assessment in bioenergy sorghum. Crop J 10, 1363–1375.

Shu, M., Fei, S., Zhang, B., Yang, X., Guo, Y., Li, B., Ma, Y., 2022. Application of UAV multisensor data and ensemble approach for high-throughput estimation of maize phenotyping traits. Plant Phenomics.

Tattaris, M., Reynolds, M., Chapman, S., 2016. A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. Front Plant Sci 7, 1131.

Tsouros, D., Bibi, S., Sarigiannidis, P., 2019. A review on UAV-based applications for precision agriculture. Information 10, 349.

Zhang, Z., Pasolli, E., Crawford, M., 2019. An adaptive multiview active learning approach for spectral–spatial classification of hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing 58, 2557–2570.


RGB
MS
TIR
Machine Learning
1
2
3
4
EL
Base models
Estimated Harvest Index
0.6
0.5
0.4
0.3
0.2
0.1

Parsed Citations

Feng, A., Zhou, J., Vories, E., Sudduth, K., Zhang, M., 2020. Yield estimation in cotton using UAV-based multi-sensor imagery. Biosyst Eng 193, 101–114.
Google Scholar: Author Only Title Only Author and Title

Fu, P., Meacham-Hensold, K., Guan, K., Wu, J., Bernacchi, C., 2020. Estimating photosynthetic traits from reflectance spectra: a synthesis of spectral indices, numerical inversion, and partial least square regression. Plant Cell Environ 43, 1241–1258.
Google Scholar: Author Only Title Only Author and Title

Gamon, J., Somers, B., Malenovský, Z., Middleton, E., Rascher, U., Schaepman, M., 2019. Assessing vegetation function with imaging spectroscopy. Surv Geophys 40, 489–513.
Google Scholar: Author Only Title Only Author and Title

Ji, Y., Liu, Z., Cui, Y., Liu, R., Chen, Z., Zong, X., Yang, T., 2023. Faba bean and pea harvest index estimations using aerial-based multimodal data and machine learning algorithms. Plant Physiol kiad.
Google Scholar: Author Only Title Only Author and Title

Li, J., Schachtman, D., Creech, C., Wang, L., Ge, Y., Shi, Y., 2022. Evaluation of UAV-derived multimodal remote sensing data for biomass prediction and drought tolerance assessment in bioenergy sorghum. Crop J 10, 1363–1375.
Google Scholar: Author Only Title Only Author and Title

Shu, M., Fei, S., Zhang, B., Yang, X., Guo, Y., Li, B., Ma, Y., 2022. Application of UAV multisensor data and ensemble approach for high-throughput estimation of maize phenotyping traits. Plant Phenomics.
Google Scholar: Author Only Title Only Author and Title

Tattaris, M., Reynolds, M., Chapman, S., 2016. A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. Front Plant Sci 7, 1131.
Google Scholar: Author Only Title Only Author and Title

Tsouros, D., Bibi, S., Sarigiannidis, P., 2019. A review on UAV-based applications for precision agriculture. Information 10, 349.
Google Scholar: Author Only Title Only Author and Title

Zhang, Z., Pasolli, E., Crawford, M., 2019. An adaptive multiview active learning approach for spectral–spatial classification of hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing 58, 2557–2570.
Google Scholar: Author Only Title Only Author and Title