25th International Congress on Modelling and Simulation, Darwin, NT, Australia, 9 to 14 July 2023
mssanz.org.au/modsim2023

Building trust in continental-scale modelling in agriculture

Jonathan Richetti [a], Bangyou Zheng [b], Javier Navarro Garcia [b] and Roger Lawes [a]

[a] CSIRO Agriculture and Food, Perth, Australia
[b] CSIRO Agriculture and Food, Brisbane, Australia
Email: Jonathan.Richetti@csiro.au

Abstract: All models are wrong, but some are useful. This famous sentence summarises modelling in all areas; however, determining model usefulness is challenging, particularly in complex systems such as agriculture. Building trust in models means providing evidence that their uncertainties and errors are acceptable for the decision-making process. Here we explore two ways to understand model usefulness in agriculture; a data-driven one and a mechanistic one. Lastly, we provide an example of combing data-driven and mechanist models to provide further insights into the agricultural landscape.

Data-driven methods such as modern machine learning algorithms relying on techniques such as cross-validation to reliably determine the model's usefulness (i.e. accuracy assessment). Cereal yields were forecasted with Deep Learning (DL) to show the implications of different architectures and provide some cross-validation recommendations for the agricultural context (Richetti et al. 2023). It is shown that DL performance can vary with architecture and that the optimal choice is task-dependent. Practical aspects for applying DL models for agricultural datasets are emphasised, such as dataset size (26 representative samples in each field sufficed) and cross-validation (indispensable on small datasets). However, mechanistic models such as crop models (i.e. APSIM) are usually not assessed similarly. Thus, a different approach is often used to build trust and determine their usefulness. To determine how a crop model could be assessed on a continental scale, the APSIM Next Gen was combined with gridded soil and weather data and run across 14,796 national variety trials from 2005 to 2022 across the Australian continent with minimal management information (sowing date and cultivar). This would allow previously calibrated cultivars to be assessed and an understanding of how the model is expected to perform if extrapolated further. Finally, we combined APSIM Next Gen outputs and machine learning to provide insights into crop rotations across every paddock across 20 million hectares in Western Australia, where we show that the benefits of break-crops and pastures to farmers are less than the 400 to 600 kg/ha benefit commonly reported from field experiments (Lawes et al. 2022).

Understanding the uncertainties of models is crucial for successfully deploying them. When modelling agriculture at a continental scale, the uncertainties are accentuated, and error sources are broadened. Nonetheless, by carefully crafting the accuracy assessments that reflect the reality in which the models are applied, the models can be useful.

REFERENCES

Lawes, R., Mata, G., Richetti, J., Fletcher, A., Herrmann, C., 2022. Using remote sensing, process-based crop models, and machine learning to evaluate crop rotations across 20 million hectares in Western Australia. Agron. Sustain. Dev. 42, 120. https://doi.org/10.1007/s13593-022-00851-y

Richetti, J., Diakogianis, F.I., Bender, A., Colaço, A.F., Lawes, R.A., 2023. A methods guideline for deep learning for tabular data in agriculture with a case study to forecast cereal yield. Computers and Electronics in Agriculture 205, 107642. https://doi.org/10.1016/j.compag.2023.107642

Keywords: *APSIM, deep learning, cross-validation, crop rotation*