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

More Profits in Good Years, Less Risk in Bad: Physics-Informed RL to Inform Crop Selection

Wenze Wang [a], **Matthew J. Knowling** [b]

[a] School of Computer and Mathematical Sciences, The University of Adelaide, Adelaide, South Australia, Australia [b] School of Agriculture, Food and Wine, The University of Adelaide, Adelaide, South Australia, Australia
Email: wenze.wang@student.adelaide.edu.au

Crop selection decision-making is influenced by a range of uncertain factors. In practice, growers make decisions with incomplete knowledge of how much soil water is available at the start of the season, how much rain will fall during the season, and what future commodity prices and input costs will be. We present a reinforcement learning (RL) framework that supports annual crop selection decisions (canola, lentil, or wheat) by maximising an RL reward built from clipped gross margin, with physics-informed shaping terms that encode plant available water (PAW) signals and simple rotation frequency and diversity terms. RL operates at the annual decision scale, using a per-year observation vector derived from APSIM Classic v7.10 (Holzworth et al., 2014) outputs and ancillary records (pre-sowing PAW mean, seasonal weather summaries, and available commodity prices and input costs). We apply the framework to a case-study farm at Roseworthy (Lower North, South Australia). The raw daily weather series drives APSIM only and is not part of the RL observation.

We aim to recover simple, threshold-like heuristics ("tipping points") that inform decision-making across seasons. We frame crop choice as a partially observable problem (Kaelbling, Littman & Cassandra, 1998): true plant-available water at sowing and future crop prices are uncertain. At each decision time, we construct a per-year observation vector from mean PAW, summary statistics of the weather forcing, and the available price and variable-cost records; the raw daily weather series is used only to drive APSIM (Holzworth et al., 2014). A Proximal Policy Optimization (PPO) agent is trained on these observations with a reward that equals annual gross margin capped within bounds to limit outliers, plus a bonus that increases with mean water via a logistic curve, minus a small penalty when the same crop is repeated too often in recent years, plus a small bonus when the last three years include different crops (Schulman et al., 2017).

We compare two learned policies in a like-for-like, controlled comparison over 1990–2022: one trained with a PAW bonus in the reward and one otherwise identical baseline with the PAW term removed. The observation vector is identical in both cases. Including the PAW bonus increases the mean evaluated return (R) by 10.7%, and the policy allocates crops sensibly by season. For example, in wet starts it shifts toward canola (51%); in normal years lentil is most common (69%); and in dry starts wheat is favoured (58%) with lentil around 30%. Here, "wet", "normal", and "dry" denote the upper, middle, and lower terciles of pre-sowing PAW (1990–2022). A simple rule that summarises the learned policy is: if pre-sowing PAW is in the upper third, choose canola; if in the middle third, prefer lentil; if in the lower third, favour wheat. This is supported by industry reports (e.g., GRDC 2022), emphasising soil moisture triggers and seasonal break timing as key determinants of canola adoption, with higher uptake in wet starts. When we replay each learned policy year-by-year on the historical sequence without exploration, the cumulative gross-margin trajectory is generally above and ends higher for the PAW-aware policy, indicating more upside in good seasons and less downside in bad seasons. The gains are modest but persistent from 1990 to 2022, suggesting value for risk-aware decision support.

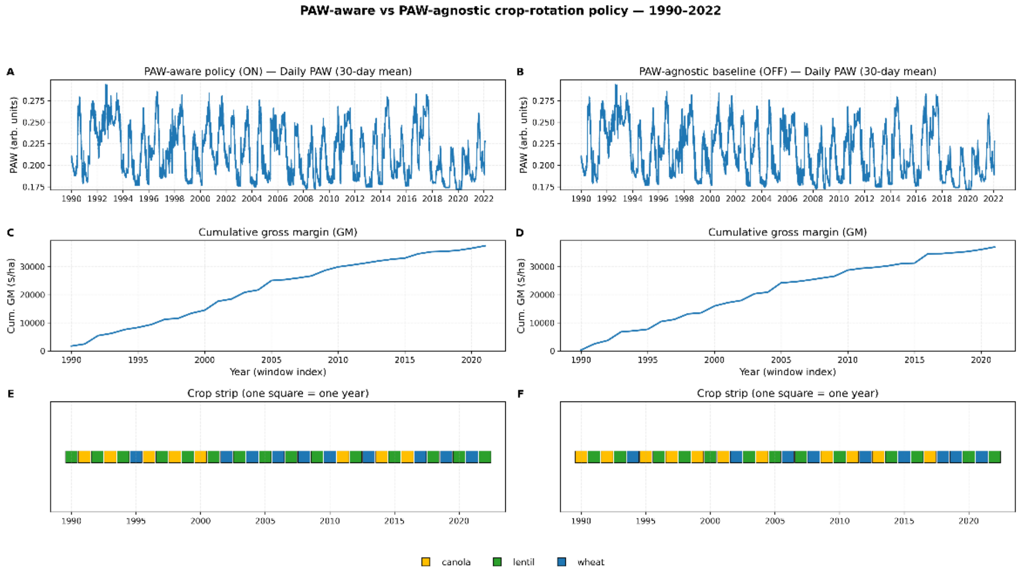
Keywords: *crop rotation; plant-available water; gross margin; reinforcement learning; decision support*

Table 1 Key metrics comparing PAW-aware (ON) vs PAW-agnostic baseline (OFF) at Eval 30k

Policy	R	Mix C/L/W	H^e(even.)	Gini (GM)	Seasonal mix D | N | W (C/L/W)
ON	240.5	30.5/49.6/19.8	0.938	0.311	12/30/58 | 29/69/2 | 51/49/0
OFF	217.3	38.5/33.8/27.7	0.992	0.329	14/19/67 | 49/33/18 | 51/49/0

Note: ON = PAW bonus in reward; OFF = same reward without PAW (observations identical). R = mean evaluated return; Mix C/L/W = % canola/lentil/wheat; H^e = Shannon evenness; Gini (GM) = Gini of annual gross margin (A$/ha); Seasonal mix D|N|W (C/L/W) = % by Dry|Normal|Wet terciles; "%" omitted.

Figure 1. PAW-aware (ON) vs PAW-agnostic (OFF) crop-rotation policy at Roseworthy, 1990–2022.



Note: **A–B:** daily PAW (30-day mean water storage, used as a PAW). The same PAW sequence is only shown for comparison. **C–D:** show cumulative gross margin from deterministic replay (A$ per ha). **E–F:** crop-strip timelines (one square = one year; canola/lentil/wheat).

REFERENCES

Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., et al. 2014, ‘APSIM—evolution towards a new generation of agricultural systems simulation’, Environmental Modelling & Software, vol. 62, pp. 327–350.

Kaelbling, L.P., Littman, M.L. & Cassandra, A.R. 1998, ‘Planning and acting in partially observable stochastic domains’, Artificial Intelligence, vol. 101, no. 1–2, pp. 99–134.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O. (2017). Proximal Policy Optimization Algorithms. arXiv preprint arXiv:1707.06347.

Grains Research & Development Corporation (GRDC) 2022, Genetic improvement of canola establishment, GRDC Update Paper, GRDC, Barton, ACT.

