An evidence-bounded crop-risk tool for pre-season planning

Adrian Dunkley
Climate Studies Group Mona, Department of Physics
The University of the West Indies, Mona

13 November 2026

Abstract

This software protocol describes an AI-facing decision tool that exposes the assumptions behind a crop plan rather than presenting an unsupported recommendation. Its numerical core uses actual NASA POWER daily rainfall estimates for a Pedro Plains-area point in Jamaica, while crop response, soil conditions and economic values remain explicit scenario inputs. A reproducible Monte Carlo engine compares planting and irrigation choices with a fixed baseline. Field-level yield, irrigation and sales data are not yet available, so the tool's agronomic benefit and interface effectiveness are unvalidated. The proposed study tests whether users can identify the decisive evidence, recognise adverse scenarios and reject a plan when essential field constraints are missing. The implemented numerical engine and proposed AI interface are reported separately.

1 Intended decision

A grower deciding on 1 March needs enough time to obtain planting material, prepare land and arrange water before April or May establishment. The tool should return a comparison of feasible options, the information used and what could reverse the ranking. It must not turn a conditional simulation into a guaranteed yield or income forecast.

The companion rainfall framework contains the executed water-balance analysis and complete equations, including the general crop response factors from the Food and Agriculture Organization [3]. This protocol reuses that engine and input archive; it is not a second independent crop experiment. Its separate research question is whether an evidence-linked interface helps people identify the data and constraints needed before acting. No user study or language-model trial has been run.

2 Agronomic evidence and software responsibilities

A crop interface must distinguish the result of a model from evidence that its inputs and outputs fit a particular farm. AquaCrop and the Decision Support System for Agrotechnology Transfer (DSSAT) already provide crop-management modelling frameworks [4, 5, 7]. Soil, crop development and evaporative demand matter to their use; the constant-demand bucket is a restricted substitute for that information, not an equivalent agronomic model [6]. Rankine et al.'s Jamaican cassava evaluation used local field trials [8]; those observations do not establish performance for this tool's tomato or watermelon scenarios.

Physics-informed neural methods introduce another distinction between a training objective and a verified output [9]. Soft penalties differ from architectural guarantees [10], and optimisation can fail to recover the intended physics [11]. The earlier Jamaican rainfall-allocation benchmark by Dunkley also separates conservation and predictive skill [12]. Here those distinctions define the interface's responsibilities: identify the engine, display the comparison's

assumptions and reject unsupported decision claims. The neural-surrogate experiment belongs to the companion numerical study. This paper's contribution is the proposed evidence contract and evaluation design, not another crop-model result.

3 Data contract and workflow

The core archive contains 8,766 NASA POWER daily precipitation values for 2001–2024. The sampling point is 17.93° N, 77.77° W, not a verified farm boundary. The provider identifies MERRA2 reanalysis; rainfall is therefore a gridded scientific estimate, not a field gauge [1, 2]. Only 2001–2018 informs candidate selection, with 2019–2024 reserved for retrospective checks. The reanalysis vintage was retrieved in 2026, so operational availability at historical decision dates remains untested.

Before an operational comparison, the interface would require the crop and cultivar, field area, planting and harvest constraints, water availability and cost, input budget, intended buyer or price range, and an agreed baseline. Missing values must stay visible. A farmer's missing soil measurement cannot silently become the simulator's default. All numerical calculations should come from a versioned engine; an AI explanation may summarise them but must not invent results or change a constraint.

The prototype provides nine options per crop from three establishment dates and three irrigation rates. The future interface would first eliminate infeasible options, then show their estimated margin differences and adverse cases. The current numerical demonstration assumes all options feasible. It does not establish that two millimetres of daily irrigation can be supplied at the example location.

4 What the executed example shows

The same 5,000 draws of historical weather year, soil and economics are paired across options. Only eighteen distinct historical weather years support the sampling. Parameter ranges are assumed independent uniform distributions, not elicited beliefs or calibrated probabilities.

Under those assumptions, the mean-margin rule selects 1 May with two millimetres per day irrigation for both tomato and watermelon. Held-out rainfall-year mean differences from the fixed 1 April rainfed baseline are USD 718 and USD 970 per hectare respectively. These are simulated differences, not observed earnings. The tomato difference is negative in two of the six held-out years; the watermelon difference is negative in one. A mean alone would conceal the decision risk.

A suitable explanation would state the selected option, its comparator, why the model favours it, the assumptions carrying that advantage and circumstances in which it loses. It would also say that unknown water access and field performance prevent an operational recommendation. The historical-rainfall animation is an input explainer, not an animation of a validated harvest forecast.

5 Neural-surrogate rejection evidence

The companion study trained a 291-parameter physics-informed neural surrogate of a single water-balance step, with actual rainfall forcing and synthetic bucket targets. Across three matched seeds, the physical penalty reduced balance root mean squared error from 0.930–0.977 mm to 0.560–0.931 mm but increased storage error from 1.790–2.398 mm to 2.429–2.679 mm. Negative outputs persisted. This experiment provides a concrete reason to retain the exact simulator rather than label a neural output trustworthy because it uses physics.

The network is not used to produce the scenario margins. A proposed interface should expose the model identity and reject physically invalid outputs before displaying a crop comparison.

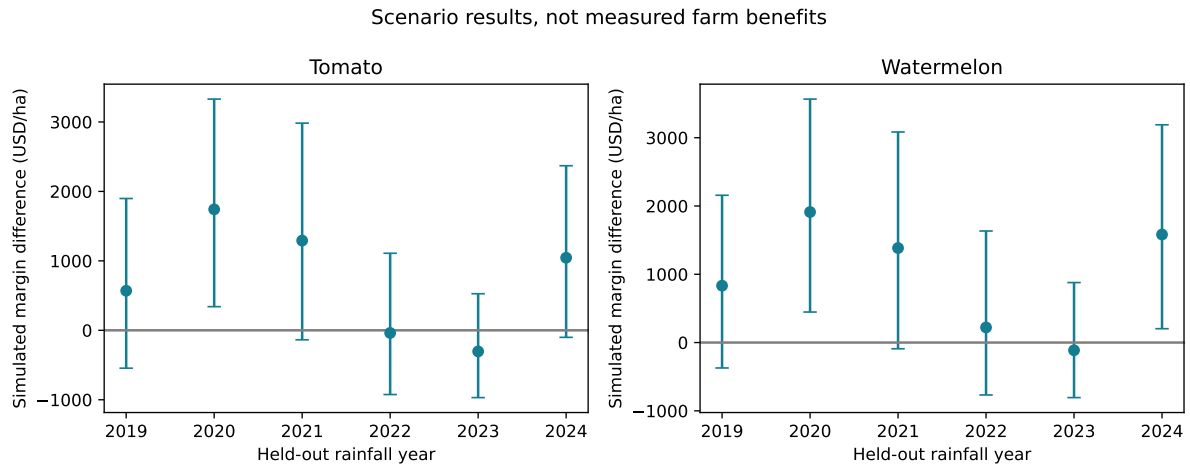


Figure 1: Same numerical experiment as the companion rainfall framework. Points and ranges show simulated differences under held-out rainfall. The reuse is explicit; no independent field replication is claimed.

The neural results are shared with the companion paper, not a new experiment reported twice. Their purpose here is to motivate an explicit dispatch and rejection contract.

6 Related interface evidence

Rose et al. [13] identify usability, relevance and performance among factors associated with agricultural tool uptake in a UK study; transfer to Jamaican growers remains a research question. Buçinca et al. [14] found a trade-off between reduced overreliance and users' subjective ratings in a non-agricultural experiment. A cautious interface is therefore not assumed to be a usable one. Amershi et al. [15] support communicating capabilities, failure limits and correction routes. The National Institute of Standards and Technology's voluntary framework [16] motivates contextual evaluation but does not certify a crop tool or establish legal compliance.

7 Proposed risk controls

The first control is an information cutoff: no data with a future valid date may influence candidate selection. The second is a feasibility check: a plan requiring unavailable irrigation is rejected. The third is a provenance check that links every displayed result to the engine output and parameter set. The fourth is an abstention rule when field inputs needed for the intended decision are missing.

A generated explanation must preserve units, baseline, uncertainty and the distinction between acquired rainfall estimate and simulated output. An interface test should deliberately introduce a wrong unit, a missing price, a cheaper but unavailable water supply and a request for guaranteed profit. Passing requires identifying the relevant evidence and explaining the consequence, not merely choosing a cautious-sounding response. No safety or alignment success is claimed before those tests run.

8 Proposed evaluation

A small usability pilot should first check whether growers and extension advisers understand the output and can supply the required inputs. A subsequent preregistered comparison could randomise an interface explanation versus the same numerical results in a plain table. Both

groups would receive identical evidence and feasible options. Outcomes would include correct identification of the decisive constraint, unsupported-recommendation rate, unit errors and the completeness of a written decision rationale.

Sample size requires a pilot estimate of variability and a stated minimum useful effect. Time to decision is secondary and should not reward a fast wrong answer. Real farming outcomes require a separate prospective agronomic study; interface comprehension does not establish increased yield or income. Appropriate consent and institutional requirements must be established before recruitment.

9 Future work

Field-level soil, irrigation, yield and sales observations are not available. Access to those records would allow crop calibration and a prospective test of practical value. The separate interface study needs a working implementation, participant recruitment with appropriate consent and a fixed evaluation plan. Its central comparison is constraint identification and unsupported-recommendation rates under identical evidence, not whether participants prefer a polished explanation.

The technical study should compare exact constraints with soft penalties and examine whether errors accumulate over repeated steps. No claim of implemented interface effectiveness, field forecast skill or improved income follows from the current numerical tests.

10 Conclusion and availability

The proposed tool separates rainfall evidence, scenario assumptions and decision permission. Water-cost sensitivity and neural constraint violations show why these distinctions must remain visible. A grower should receive a conditional comparison only after feasible options are established, with the missing evidence stated before any action is considered.

The numerical engine, rainfall archive, trained surrogate weights, stress results and verification records are supplied with the companion framework. No human-subject study was conducted. This paper provides an interface-study protocol and an evidence contract, not a deployed service or an independently validated agricultural intervention.

References

[1] NASA POWER. Data questions. Accessed 13 November 2026. https://power.larc.nasa.gov/docs/faqs/data/.

[2] NASA POWER. Daily API documentation. https://power.larc.nasa.gov/docs/services/api/temporal/daily/.

[3] Food and Agriculture Organization. Annex II: Irrigation scheduling. https://www.fao.org/4/X5647E/x5647e0e.htm.

[4] Food and Agriculture Organization. About AquaCrop: crop water productivity model. https://www.fao.org/aquacrop/overview/en/. Accessed 13 November 2026.

[5] J. W. Jones et al. (2003). The DSSAT cropping system model. European Journal of Agronomy 18, 235–265. https://doi.org/10.1016/S1161-0301(02)00107-7.

[6] R. G. Allen, L. S. Pereira, D. Raes and M. Smith (1998). Crop evapotranspiration: guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. https://www.fao.org/4/X0490E/x0490e00.htm.

[7] P. Steduto, T. C. Hsiao, E. Fereres and D. Raes (2012). Crop yield response to water. FAO Irrigation and Drainage Paper 66. https://www.fao.org/4/i2800e/i2800e00.htm.

[8] D. Rankine et al. (2021). Evaluation of DSSAT-MANIHOT-Cassava model to determine potential irrigation benefits for cassava in Jamaica. Agronomy Journal 113, 5317–5334. https://doi.org/10.1002/agj2.20876.

[9] M. Raissi, P. Perdikaris and G. E. Karniadakis (2019). Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045.

[10] T. Beucler et al. (2021). Enforcing analytic constraints in neural networks emulating physical systems. Physical Review Letters 126, 098302. https://doi.org/10.1103/PhysRevLett.126.098302.

[11] A. Krishnapriyan, A. Gholami, S. Zhe, R. Kirby and M. W. Mahoney (2021). Characterizing possible failure modes in physics-informed neural networks. Advances in Neural Information Processing Systems 34. https://papers.nips.cc/paper/2021/hash/df438e5206f31600e6ae4af72f2725f1-Abstract.html.

[12] A. Dunkley (2026). Exact volume conservation improves a neural rainfall downscaler over Jamaica and does not beat seasonal allocation: a conditional CHIRPS benchmark with a driver-error floor. Unpublished research manuscript, 10 November.

[13] D. C. Rose et al. (2016). Decision support tools for agriculture: towards effective design and delivery. Agricultural Systems 149, 165–174. https://doi.org/10.1016/j.agsy.2016.09.009.

[14] Z. Buçinca, M. B. Malaya and K. Z. Gajos (2021). To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. Proceedings of the ACM on Human-Computer Interaction 5(CSCW1). https://doi.org/10.1145/3449287.

[15] S. Amershi et al. (2019). Guidelines for human-AI interaction. CHI. https://doi.org/10.1145/3290605.3300233.

[16] National Institute of Standards and Technology (2023). Artificial Intelligence Risk Management Framework. NIST AI 100-1. https://doi.org/10.6028/NIST.AI.100-1.