Bayesian Methods for Crop Trait Optimization in Precision Agriculture

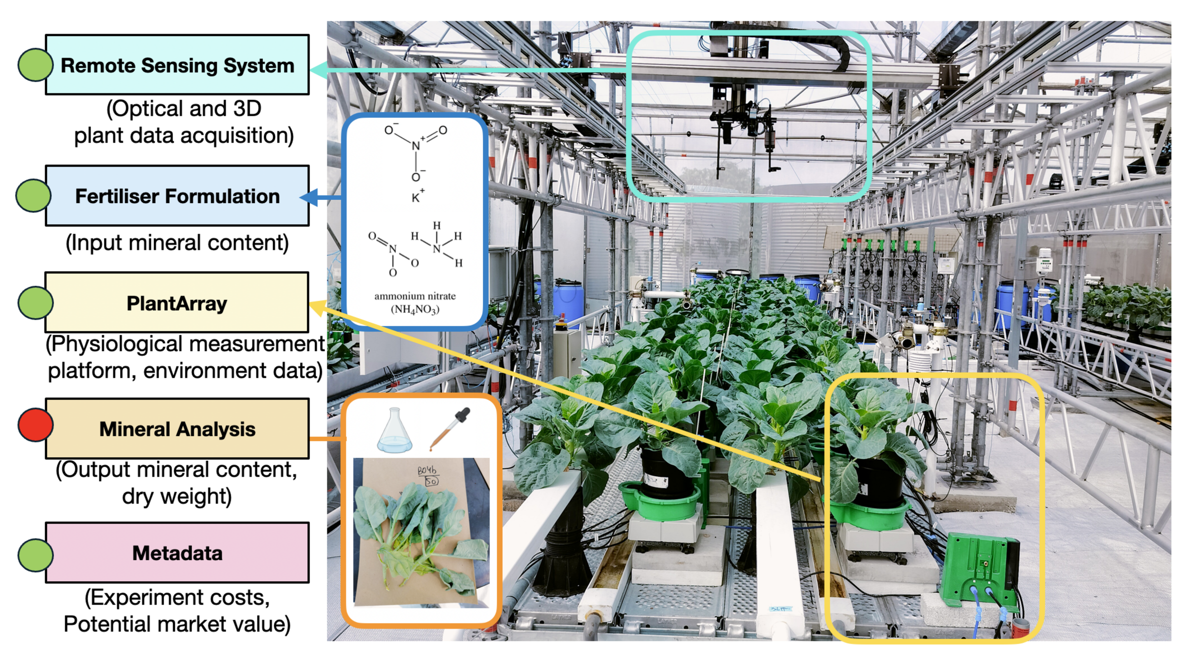
Miguel R. Pebes-Trujillo[1,4], Itamar Shenhar[1,2,4], Aravind Harikumar[1,4], Ittai Herrmann[2,4], Menachem Moshelion[2,4], Kee Woei Ng[1,4], Matan Gavish[3,4]

[1]School of Materials Science & Engineering, Nanyang Technological University; [2] The Robert H. Smith Faculty of Agriculture, Food and Environment, Hebrew University of Jerusalem (HUJ); [3]School of Computer Science and Engineering, HUJ; [4]Singapore-HUJ Alliance for Research and Enterprise (SHARE)

Abstract

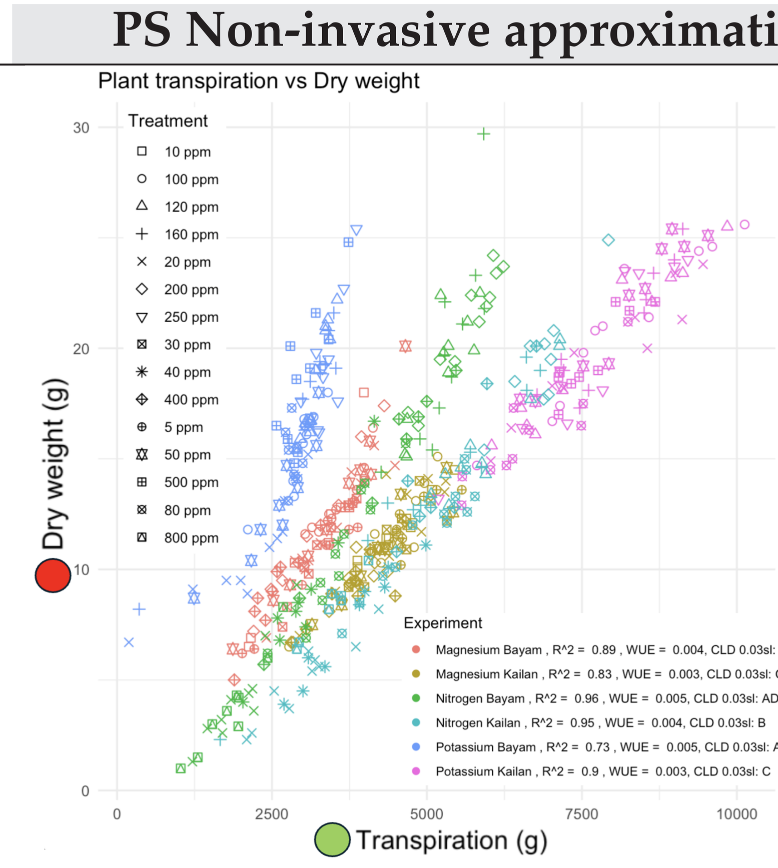
We develop Bayesian methods for urban precision agriculture to boost crop trait and resource use optimization through principle models and high-throughput data sources. E.g., a common objective is to maximize yield, nutritional value, and taste while minimizing costs and environmental impact.

Phenotyping facility: Oasis Living Lab, Singapore



Availability during the experiment: Green=Yes, Red=No
Relevant trait data is collected destructively after harvest.
We use in-situ (PS) and Remote Sensing (RS) capabilities.

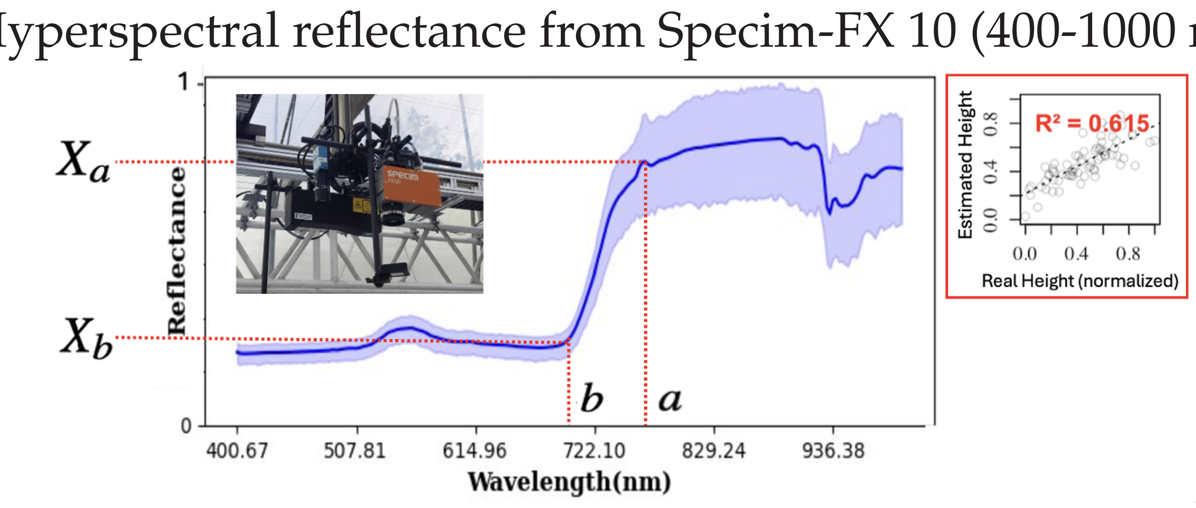
PS Non-invasive approximation



Dry weight (obtained after harvest) can be estimated from *transpiration* (collected near-continuously by PlantArray).

RS Non-invasive approximation

Hyperspectral reflectance from Specim-FX 10 (400-1000 nm)



We learn the set of wavelegnth indices (a, b, c, d) whose reflectance values (X_a, X_b, X_c, X_d) predicts the trait $\mathbf{y}$.
We fit: $\mathbf{y} = \theta_0 \cdot \mathbf{1} + \theta_1 \cdot \mathbf{z}_1 + \theta_2 \cdot e^{\circ(\theta_3 \cdot \mathbf{z}_2)} + \boldsymbol{\varepsilon}$
with weights θs, $e^{\circ(\theta_3 \cdot \mathbf{z}_2)} = \left[e^{\theta_3 z_{21}}, e^{\theta_3 z_{22}}, \ldots, e^{\theta_3 z_{2n}}\right]^\top$, and $\mathbf{z}_1$ and $\mathbf{z}_2$ are *Normalized Difference Indices* defined as

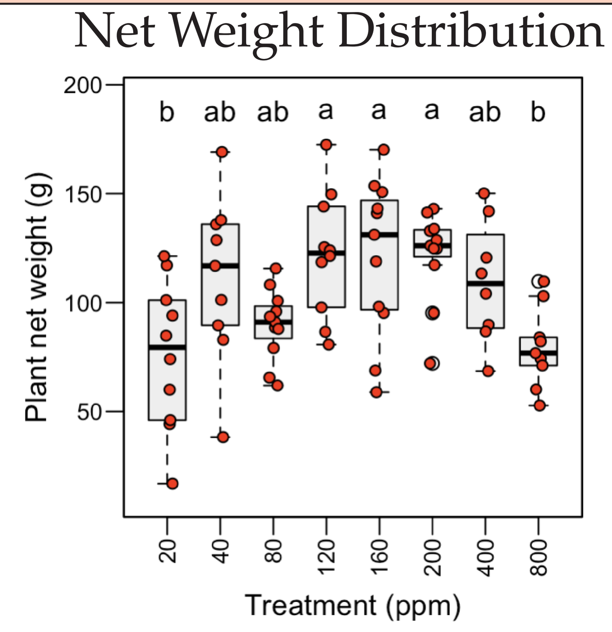
$$\mathbf{z}_1 = \left(\frac{X_{ai} - X_{bi}}{X_{ai} + X_{bi}}\right)_{i=1}^n \qquad \mathbf{z}_2 = \left(\frac{X_{ci} - X_{di}}{X_{ci} + X_{di}}\right)_{i=1}^n.$$

Methods for Experiment Evaluation and Feedback Intervention

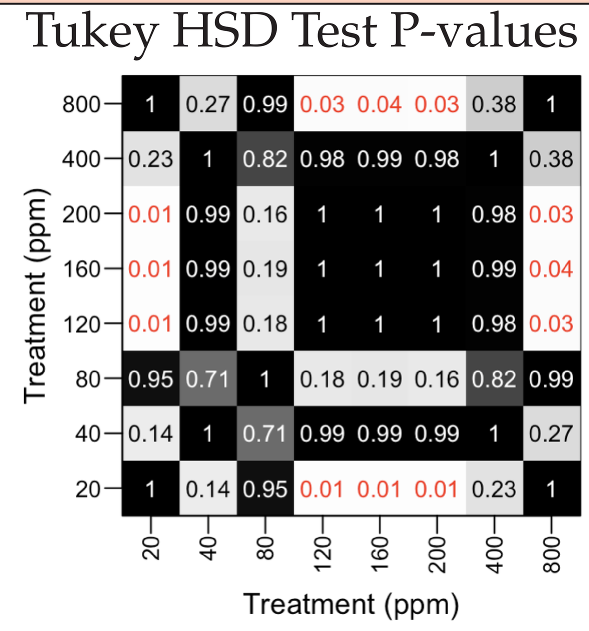
Example 1: Evaluation. We ran an experiment on the PlantArray system where treatments varied in the supplied nitrogen concentration: 20, 40, 80 120, 160, 200, 400, or 800 mg L^{-1} or parts per million (ppm). The objective is to identify the treatment that maximizes *plant fresh weight* at harvest.

Typical analysis (inputs data)

Net Weight Distribution



Tukey HSD Test P-values

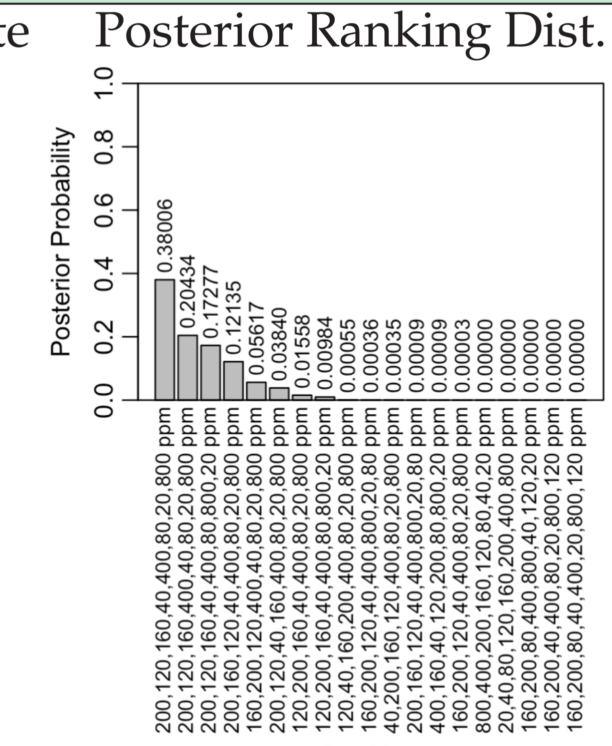


Our method (inputs data + optimization criteria)

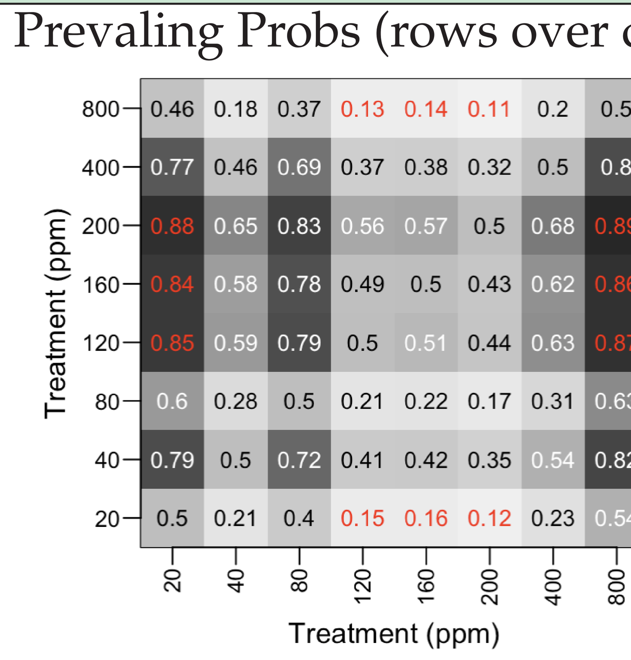
Ranking Point Estimate

1. **200 ppm**
2. 120 ppm
3. 160 ppm
4. 40 ppm
5. 400 ppm
6. 80 ppm
7. 20 ppm
8. 800 ppm

Posterior Ranking Dist.

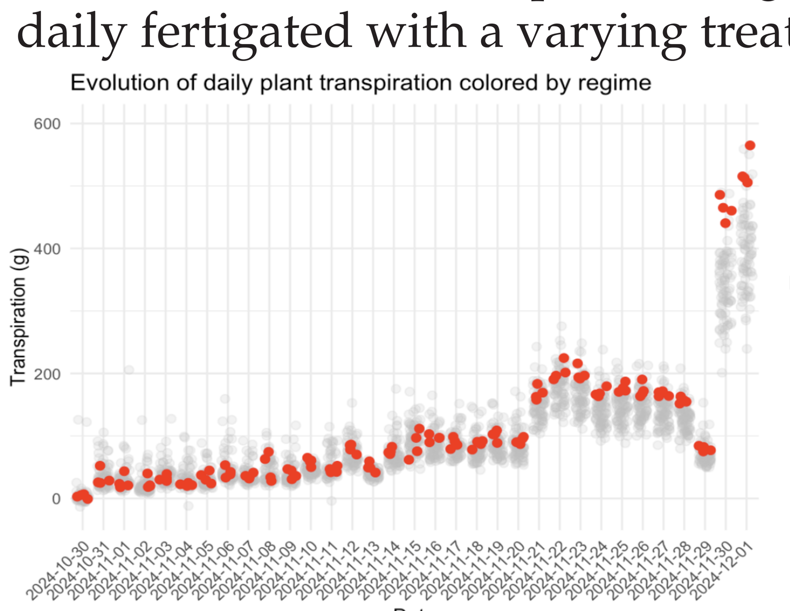


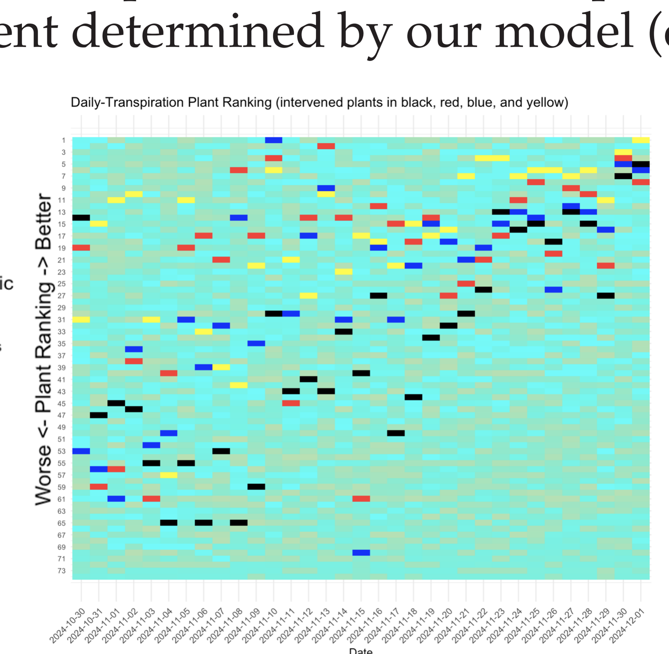
Prevaling Probs (rows over cols)

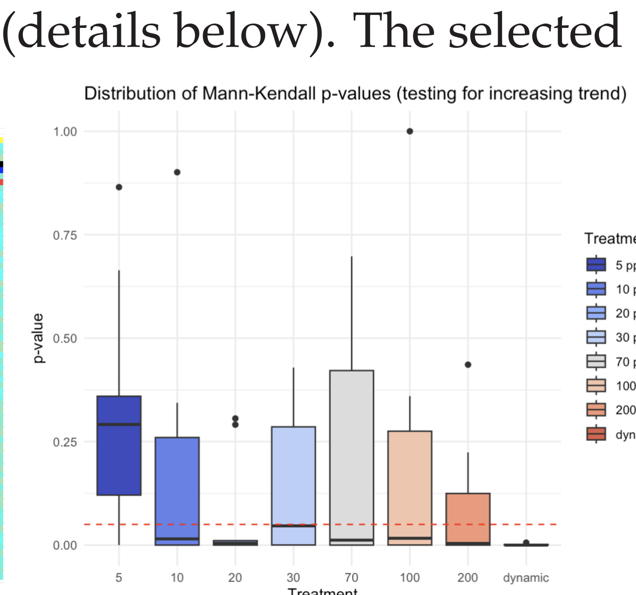


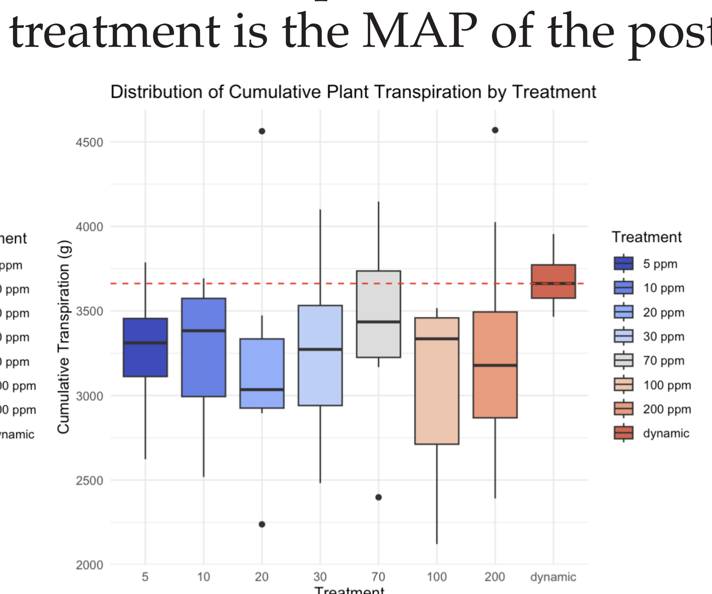
In practice, we want to achieve multiple (possibly conflicting) traits. For example, maximizing yield and nutritional content while minimazing environmental impact and cost.

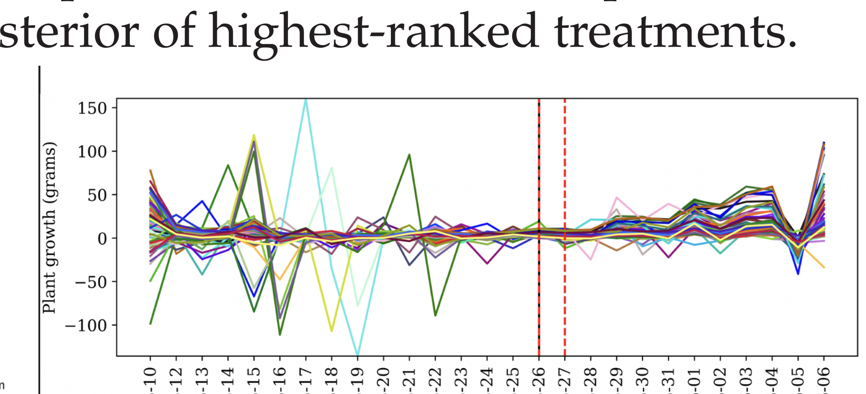
Example 2: Intervention. We address **yield maximization** using as informing measurement *plant transpiration* which is cumulatively linear on dry weight (yield). Thus, better treatments correspond to higher transpiration levels. We experimented with 70 Kailan (Chinese broccoli) plants under 7 Phosphorus treatments. Four plants were daily fertigated with a varying treatment determined by our model (details below). The selected treatment is the MAP of the posterior of highest-ranked treatments.











$$p(\mathbf{y} \mid \boldsymbol{\tau}, \mu_1, \nu_1, \beta_0, \beta_1, \sigma^2) = \prod_{i=1}^{n} \left(\prod_{t=1}^{\tau_i} \text{Normal}(y_t \mid \mu_1, \nu_1) \times \prod_{t=\tau_{i+1}}^{T} \text{Normal}\left(\epsilon_t \mid y_t - \left(\beta_0 e^{\beta_1(t-\tau_{i+1}+1)}\right), \sigma^2\right)\right)$$

Note: The denoising problem

Key Ideas on Characterization and Modeling

Problem: Which treatment $1, \ldots, K$ is the optimal? (for M ideal features $\boldsymbol{X}^* = \{X_1^*, \ldots, X_M^*\} \in \mathbb{R}^M$).
Data: $\boldsymbol{X}_k = [X_{kij}] \in \mathbb{R}^{n_k \times M}; k = 1, \ldots, K; i = 1, \ldots, n_k; j = 1, \ldots, M$.
Deviation RVs: $D_{kij} = |X_j^* - X_{kij}|$ (deviation symmetry assumption)
Comparison RVs: For two treatments r and s, $Y_{jrs} = \sum_{a=1}^{n_r} \sum_{b=1}^{n_s} I(D_{raj} < D_{sbj})$. Then $\boldsymbol{Y} = [Y_{jrs}] \in (\{0\} \cup \mathbb{Z}^+)^{M \times K \times K}$ is the collection of Y_{jrs} for all features j and all treatment pairs (r, s); and $\boldsymbol{y}$ = observed realizations.
Treatment Performance: via *dominance indexes* $\boldsymbol{d} = d_1, \ldots, d_K \in \mathbb{R}^K$.
Multiple features: Each treatment has M sub-indexes $d_{k1}, \ldots, d_{kM} \in [0, u]^M$, s.t. $d_k = d_{k1} + \ldots + d_{kM}$, $\sum_{j=1}^M w_{kj} = 1$, and $d_{kj} = d_k w_{kj}$.
Denote weights $\boldsymbol{w}_k = [w_{k1}, \ldots, w_{kM}]$, and $\boldsymbol{w} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K]^\top$.
The Bradley-Terry Model for Experiments (BTME) Consider the BT model with $\phi : \mathbb{R} \to (0, 1)$, e.g. sigmoid $\phi(x) = (1 + e^{-x})^{-1}$, for $\boldsymbol{y}$:

$$\begin{aligned} p(\boldsymbol{y} \mid \boldsymbol{n}, \boldsymbol{d}, \boldsymbol{w}) &= \prod_{j=1}^{M} \prod_{r=1}^{K} \prod_{s=r+1}^{K} \binom{n_r \times n_s}{y_{jrs}} \times \left(\phi\left(d_r w_{rj} - d_s w_{sj}\right)\right)^{y_{jrs}} \\ &\times \left(1 - \phi\left(d_r w_{rj} - d_s w_{sj}\right)\right)^{n_r \times n_s - y_{jrs}}. \end{aligned}$$

Key Steps on Inference

We use vanilla M-H to estimate $(\boldsymbol{d}, \boldsymbol{w})$, which, jointly, answer our question and enable uncertainty quantification. Sample parameterization:

$$\begin{aligned} d_1, \ldots, d_K &\overset{iid}{\sim} \text{Uniform}[0, u \times M] \quad (k = 1, \ldots, K) \\ w_{k1}, \ldots, w_{kM} &\sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_K) \\ \pi_{rs} &= \text{Sigmoid}(d_r w_{rj} - d_s w_{sj}) \quad (j = 1, \ldots, M; r < s) \\ y_{jrs} \mid n_r, n_s, \pi_{rs} &\overset{ind}{\sim} \text{Binomial}(n_r \times n_s, \pi_{rs}) \quad (j = 1, \ldots, M; r < s); \end{aligned}$$

Proposals: Trucated-Normal$(0, u \times M)$ for ds, and we randonly walk on the $(M-1)$-Simplex for $\boldsymbol{w}$s. Finally, we estimate vector $\boldsymbol{d}$ and matrix $\boldsymbol{w}$.

[1] Miguel R. Pebes-Trujillo, Itamar Shenhar, Aravind Harikumar, Ittai Herrmann, Menachem Moshelion, Kee Woei Ng, and Matan Gavish. Ranking of multi-response experiment treatments. *arXiv*. 2024.

Acknowledgement

Research supported by the NRF Singapore, under its Campus CREATE and program SHARE REQ0246254 (NRF2020-THE003-0008).