# Algorithmically-Guided Postharvest by Experimental Combinatorial Optimization

**Ofer Shir**, Boris Yazmir, Assaf Israeli, and Dan Gamrasni

Tel-Hai College

MIGAL Galilee Research Institute

July 2022, GECCO'22, Boston MA, USA

# Postharvest

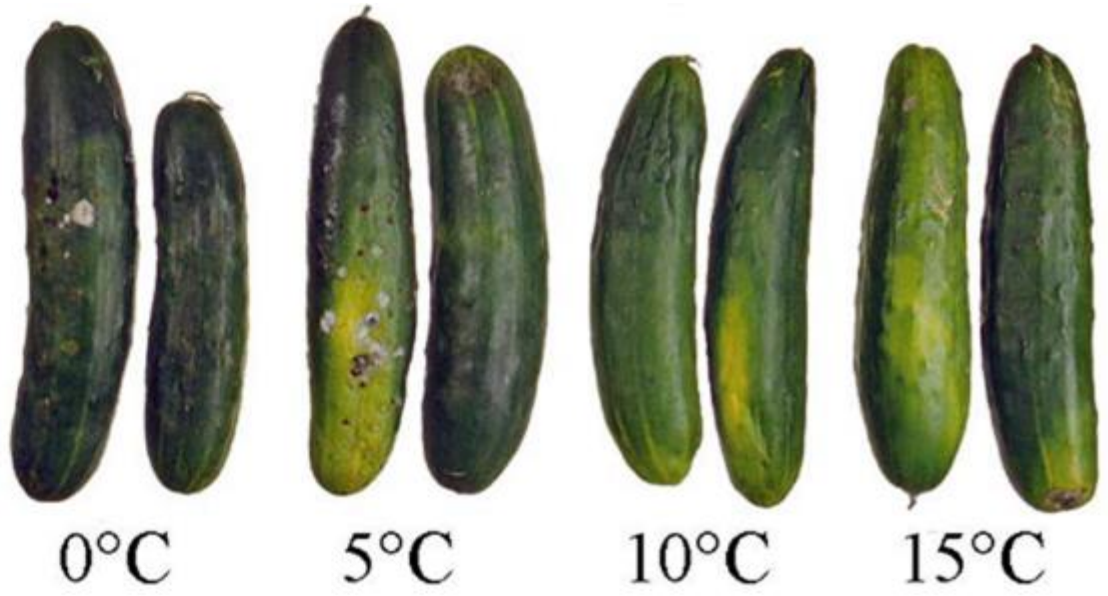
- The collection of practices for handling crops immediately following their harvest, with the explicit goal of maintaining their quality, while boosting their shelf-life.
- Postharvest technologies are a cornerstone of modern sustainability, and influence food security directly, with a potentially vast economic impact on the global food supply-chain.
- Yet, they impose significant scientific challenges concerning treatment protocols for fresh fruit and vegetables.

# Cucumbers as a model for the postharvest challenge

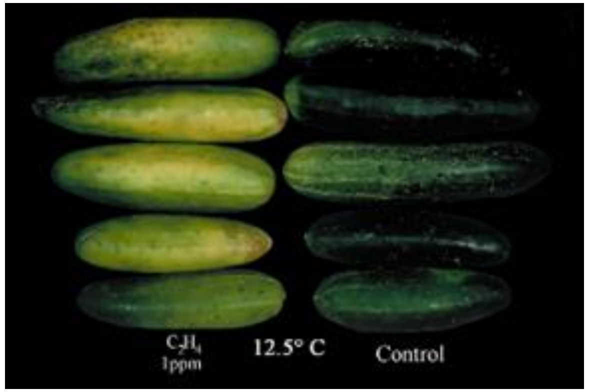
**Cucumbers are extremely sensitive with high postharvest losses.**

**Postharvest blemishes:**

- Rots
- Chilling injury
- Color loss (yellowing)
- Weight loss/shriveling
- Softening




Cucumbers are available all year round
➔ sequential experimentation




Tel-Hai College

# Project’s Goal

## Hypothesis:

AN *OPTIMIZED COMBINATION* OF TWO TREATMENTS AND A PACKAGE WILL SIGNIFICANTLY IMPROVE CUCUMBERS’ QUALITY DURING STORAGE AND MARKETING

Goal: minimize cucumbers’ postharvest quality loss
(*i.e.*, deterioration after harvest ➔ min)



Tel-Hai College

MIGAL Galilee Research Institute LTD.

# Experimental design

- **Treatments**: selected according to existing literature and preliminary research: Plant growth regulators, U.V. radiation, volatiles, antioxidants, edible coating, wax…
  - Each treatment has multiplicity of activation levels (concentration, time, etc.)
- **Package Type**: LDPE (Control), MAP- ROP/ ZOEPAK.
- The algorithm suggests the combinations:

  Treatment A→ Treatment B → Package(x/y/z)
- **Storage temperatures:**

  20ºC- immediate marketing

  10ºC- cold storage
- **Quality assessment:**

  The fruit was measured following 4 weeks in storage

# Search-Space Definition

Given $n_t$ postharvest treatments and a set of packages, a candidate treatment is denoted as:

$$\vec{\tau} \in \pi \circ \mathcal{T}, \quad \pi \in P_{\pi}^{(n_t)}, \quad \mathcal{T} = \mathcal{T}_1 \times \mathcal{T}_2 \times \cdots \times \mathcal{T}_{n_t} \times \mathcal{P}$$

( $\mathcal{T}_j$ lists the levels/categories of each treatment).

The obtained search-space cardinality:

$$|\pi \circ \mathcal{T}| = n_t! \cdot \left[ \prod_{j=1}^{n_t} \left( \mathcal{T}_j^{\max} - \mathcal{T}_j^{\min} \right) \right] \cdot |\mathcal{P}|$$

Tel-Hai College

MIGAL Galilee Research Institute LTD.

# Objective Function Definition

Given a combinatorial search-space of possible postharvest treatments, obtain a protocol that minimizes the deterioration.

Color deviation (normalized): $\Delta c\left(\vec{\tau}\right)$
Stiffness deviation (normalized): $\Delta s\left(\vec{\tau}\right)$
Mass reduction (normalized): $\Delta m\left(\vec{\tau}\right)$
Expert's score (normalized): $\text{score}_{\text{exp}}\left(\vec{\tau}\right)$

$$\mathcal{L}_{i\rightarrow f}\left(\vec{\tau}\right) := \Delta c\left(\vec{\tau}\right) + \Delta s\left(\vec{\tau}\right) + \Delta m\left(\vec{\tau}\right) + \text{score}_{\text{exp}}\left(\vec{\tau}\right) \mapsto \min$$

In practice, the current project targets a reduced form of 2 treatments followed by a packaging (~$10^6$ combinations altogether).

Tel-Hai College

MIGAL Galilee Research Institute LTD.

# Pragmatic Settings and Compact Representation

Pragmatic industrial settings: applying only 2 treatments:

$$\begin{aligned} &\text{minimize}_{\vec{\tau} \in \pi \circ \mathcal{T}} && \mathcal{L}_{i \to f}(\vec{\tau}) \\ &\text{subject to:} && \#\left\{ j : \tau_j \neq 0,\ j = 1, \ldots, n_t \right\} == 2, \\ & && \{\text{decay} < \epsilon\} \end{aligned}$$

Getting the following pleasantly-compact representation:

$$\vec{\varphi} := \left( \overbrace{1^{st}\ \text{treatment},\ 2^{nd}\ \text{treatment}, \text{package},}^{\vec{d}:\ \text{categorical}} \overbrace{1^{st}\ \text{level}, 2^{nd}\ \text{level}}^{\vec{z}:\ \text{integers}} \right)^T .$$

# Taken Approach and Setup

The compact representation reduces the search-space cardinality from $10^{17}$ to $10^{6}$.

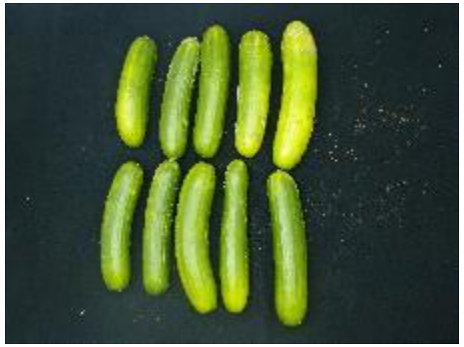
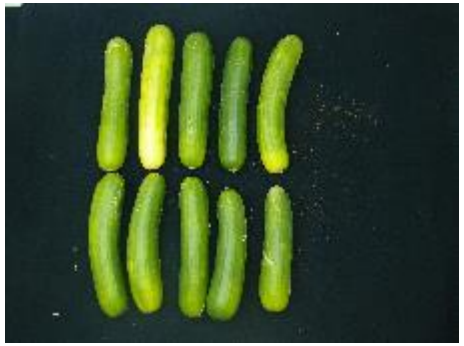
Yet, the laboratory program approved 7 iterations, each having 26 applications (with 10 biological repetitions):

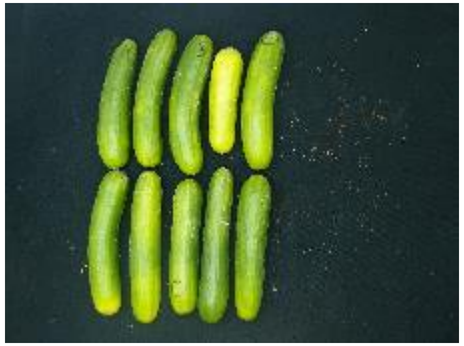
**7 iterations X (11 candidates + 2 references) X 2 Systems {10C,20C}**

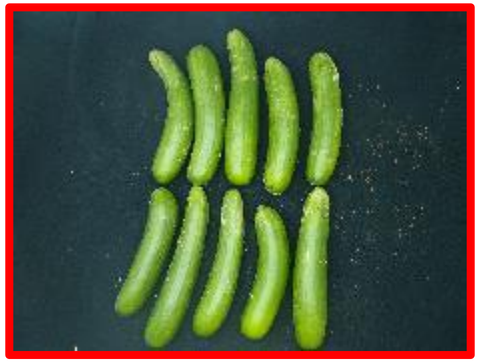
The chosen strategy: a Categorical Evolution Strategy (details omitted – see *Reehuis and Bäck GECCO'10*).

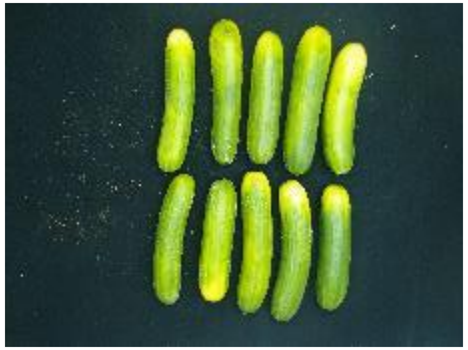
# 20°C system: 28 days postharvest outcome



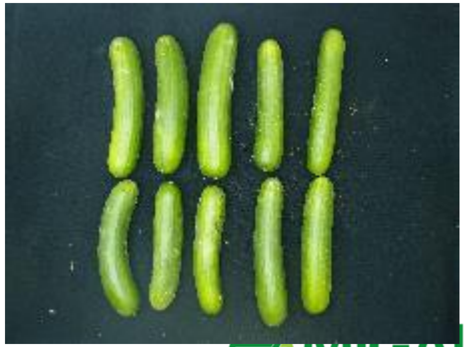





**Control**
(Untreated)

**'In house' treatment**

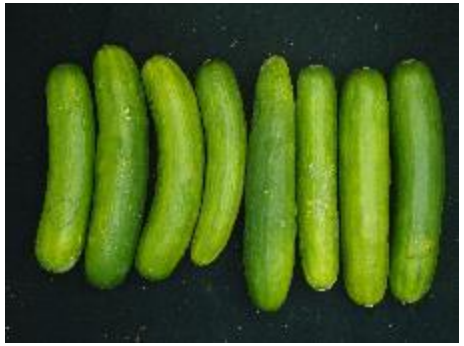


College

MIGAL Galilee Research Institute LTD.

# 10°C system: 28 days postharvest outcome

Control (Untreated)



'In house' treatment

# Prolonged Cucumbers' storage up to 63(!) days at 10°C



# Normalization ?

Using no-treatment (“control”), denoted ‘$\tau_0$’ , and “inhouse” (best human practice), denoted ‘$\tau_{ih}$’, we considered two forms:

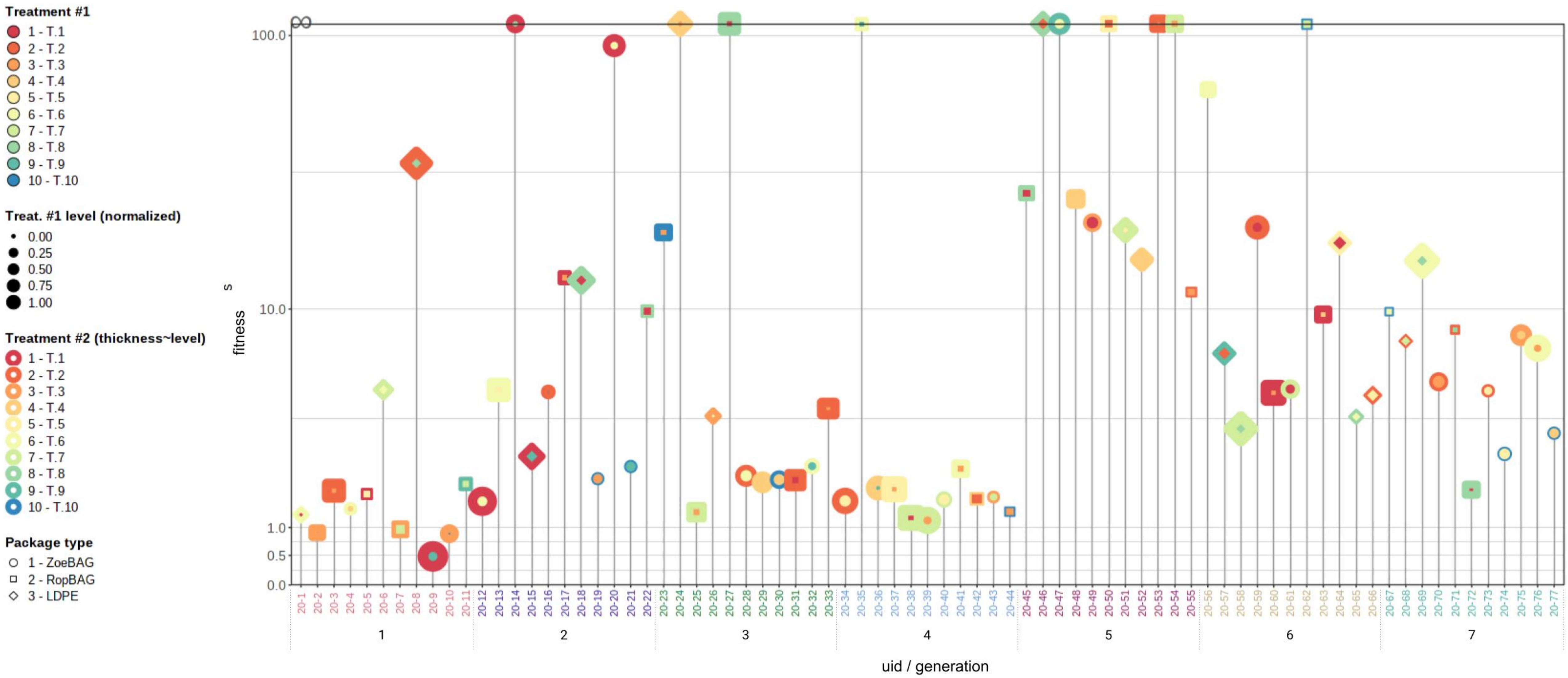
(*)

$$f\left(\vec{\tau}^{(g)}\right) := \frac{\mathcal{L}_{i\to f}\left(\vec{\tau}^{(g)}\right) - \mathcal{L}_{i\to f}\left(\vec{\tau}^{(g)}_{\mathrm{ih}}\right)}{\left(\mathcal{L}_{i\to f}\left(\vec{\tau}^{(g)}_{0}\right) - \mathcal{L}_{i\to f}\left(\vec{\tau}^{(g)}_{\mathrm{ih}}\right)\right)},$$

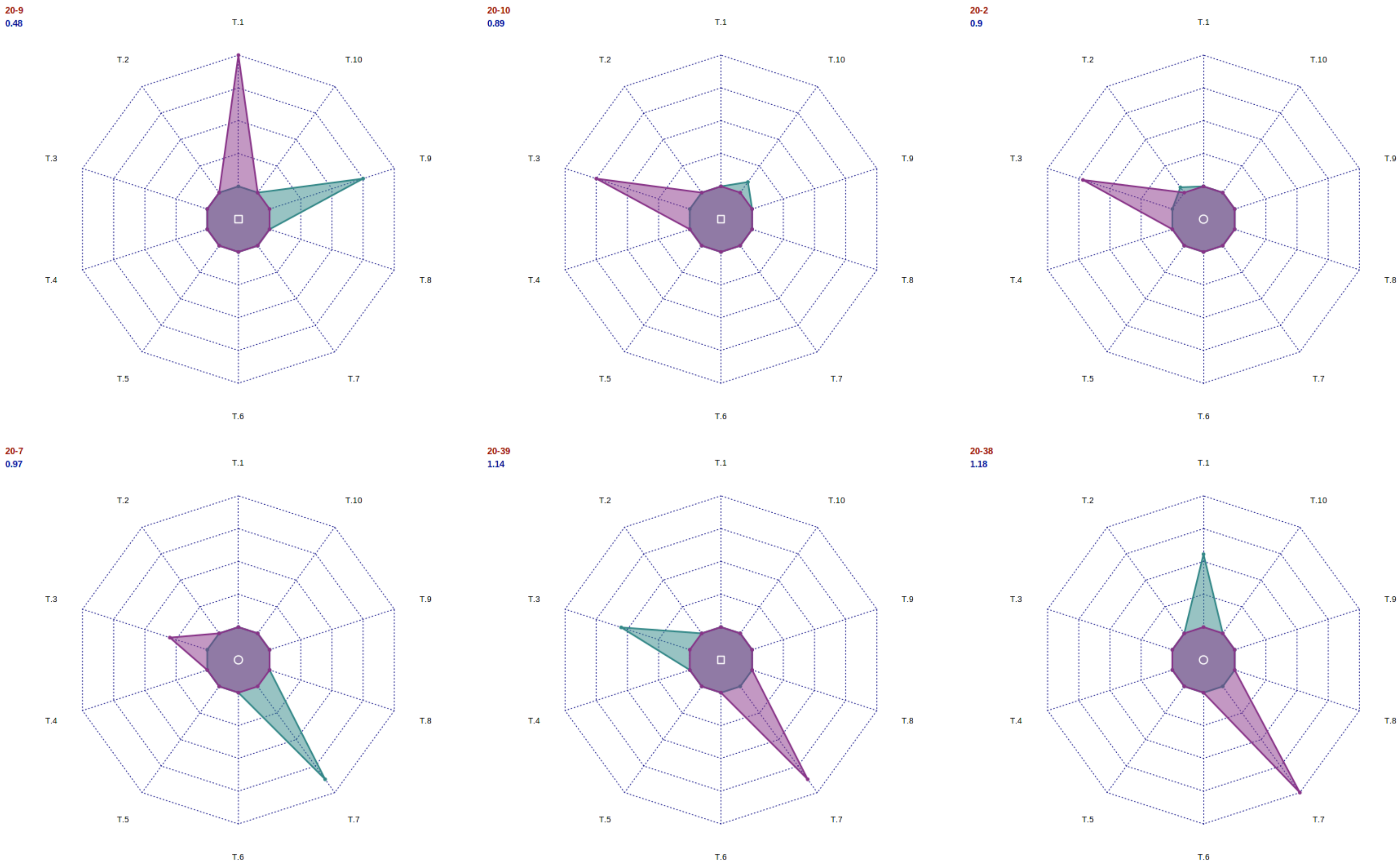
(**)

$$\tilde{f}\left(\vec{\tau}^{(g)}\right) := \frac{\mathcal{L}_{i\to f}\left(\vec{\tau}^{(g)}\right) - \mathcal{L}_{i\to f}\left(\vec{\tau}^{(g)}_{\mathrm{ih}}\right)}{\mathcal{L}_{i\to f}\left(\vec{\tau}^{(g)}_{\mathrm{ih}}\right)}$$

# Lollipop visualization of the entire campaign [20°C system]



# The top 6 protocols obtained for the 20°C system



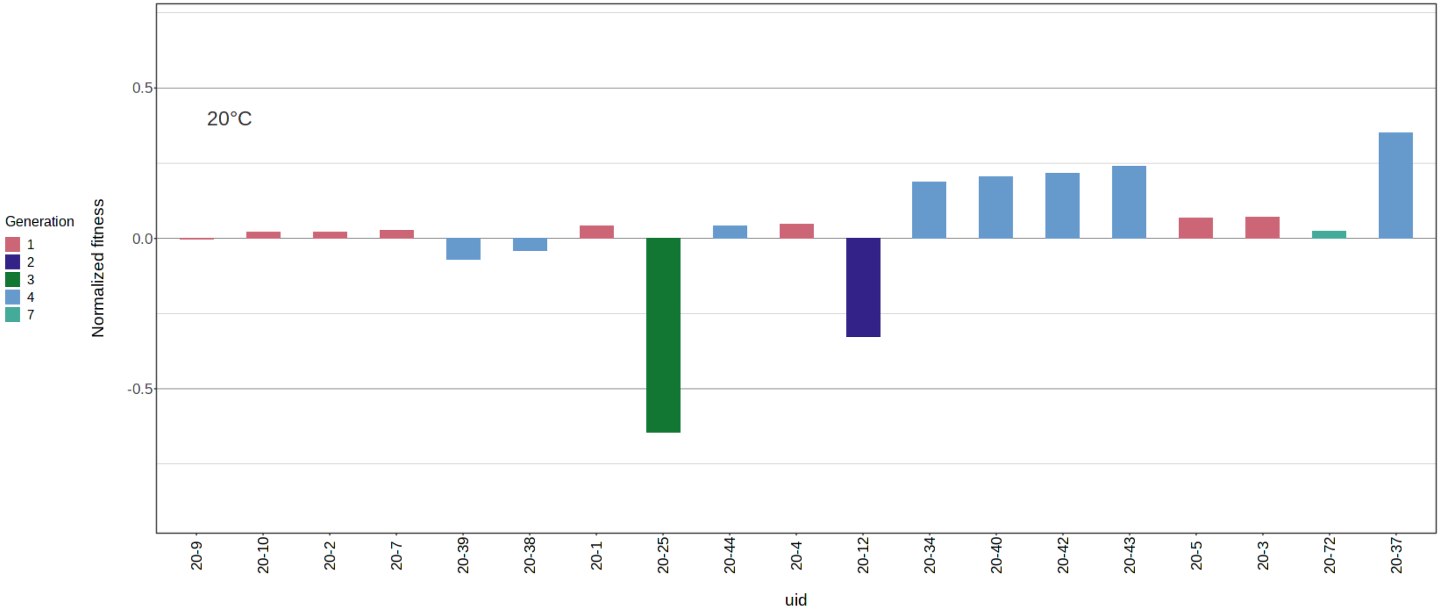

➔ They all improve, sometimes dramatically, their iteration’s “in-house” protocol.

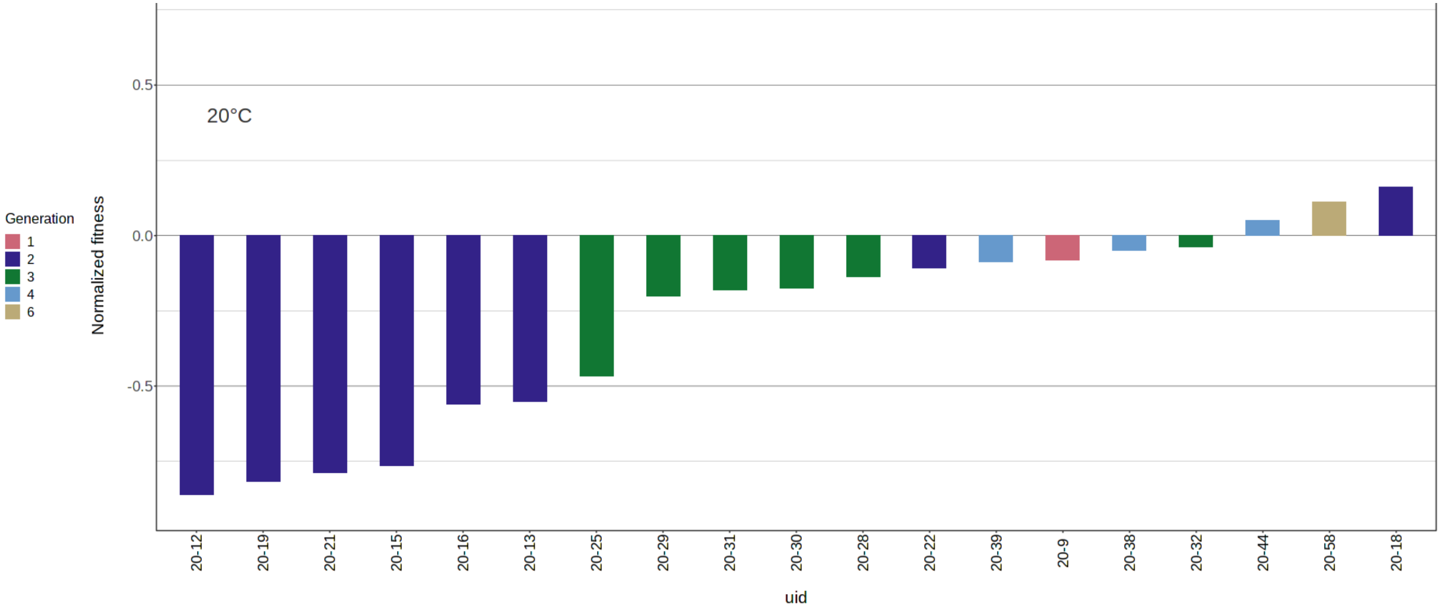
➔ Retrospectively, the combination of treatments may be explained biologically.

# But normalization does not seem to work well

- Using (*) at TOP – ordered by raw values.
- Using (**) at BOTTOM; ordered by normalized values.

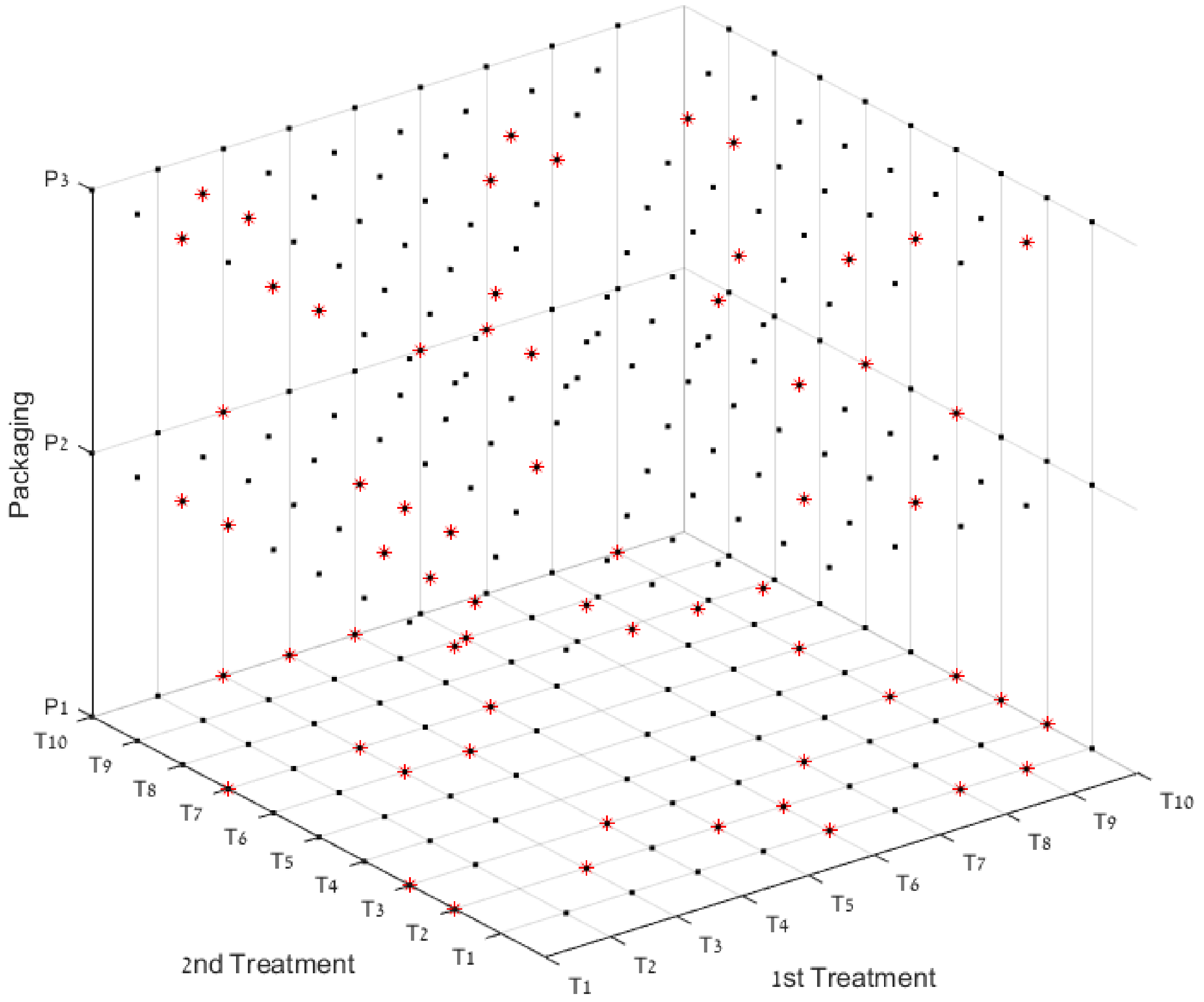
Statistical tests to quantify the correlation between the raw values to each of the two normalization forms:
Pearson’s $\boldsymbol{r}$-values read $\boldsymbol{r}$*=0.48 and $\boldsymbol{r}$**=0.55 - reflecting low to moderate correlation.
➔ Applying normalization requires further investigation.



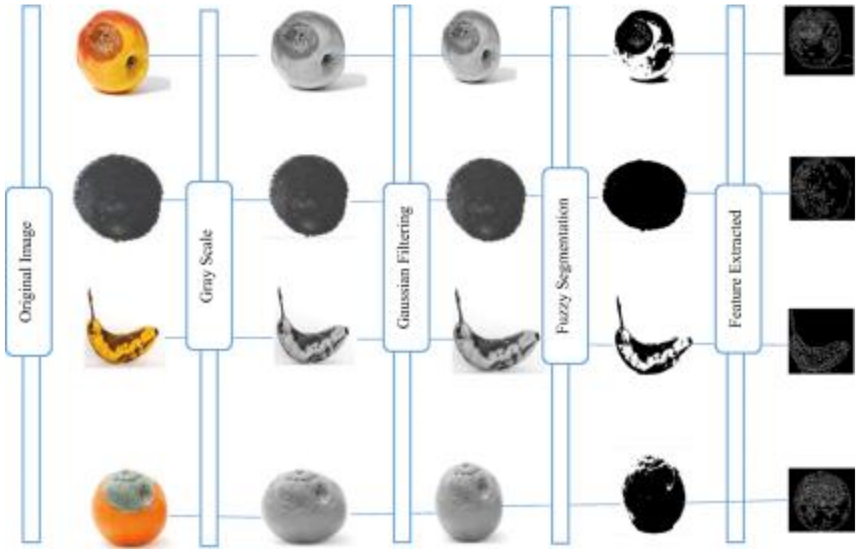
# Coverage of the categorical sub-space by the algorithm

All 270 feasible search-points (black points) within the 3-dimensional categorical sub-space (lacking activation levels!), and those visited *de facto* by the algorithm per the 20°C system (64 red stars).



# Next steps

- Multiobjective Pareto consideration.
- Learning the response surface(s) – might depend on normalization
- ‘Optima transfer’ of protocols from one crop to another (e.g. cucumbers’ protocol to zucchini).
- ‘One shot optimization’ algorithm for seasonal crops (berries!)
- Objective function evaluation via *image analysis*:



```
root@PLEXUS:~$ thank you
root@PLEXUS:~$ any questions?
```