Self-supervised learning predicts plant growth trajectories from multi-modal industrial greenhouse data

Adam J Riesselman [1] **Evan M Cofer** [1] **Therese LaRue** [1] **Wim Meeussen** [1]

Abstract

Quantifying organism-level phenotypes, such as growth dynamics and biomass accumulation, is fundamental to understanding agronomic traits and optimizing crop production. However, quality growing data of plants at scale is difficult to generate. Here we use a mobile robotic platform to capture high-resolution environmental sensing and phenotyping measurements of a large-scale hydroponic leafy greens system. We describe a self-supervised modeling approach to build a map from observed growing data to the entire plant growth trajectory. We demonstrate our approach by forecasting future plant height and harvest mass of crops in this system. This approach represents a significant advance in combining robotic automation and machine learning, as well as providing actionable insights for agronomic research and operational efficiency.

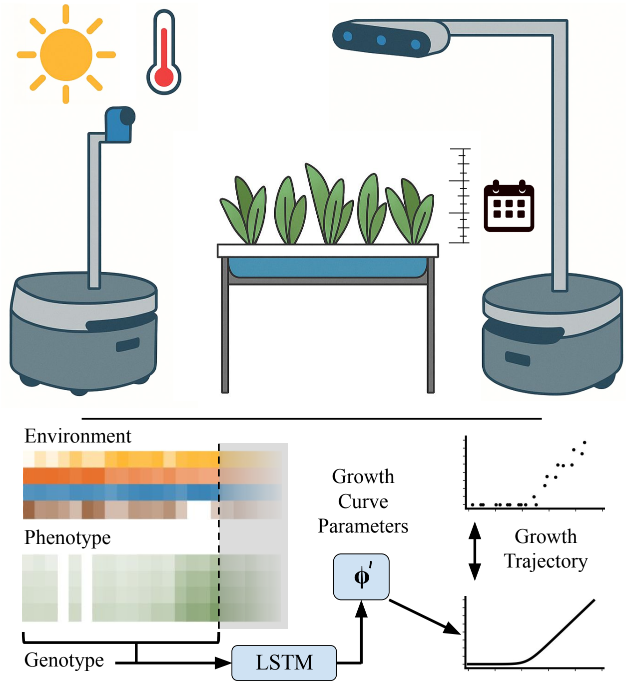

Figure 1. Top: Robots and sensors collect data of leafy greens grown hydroponically: left logs environment (temperature, light, humidity), and right captures phenotype (height over time). Bottom: The entire trajectory of growth is generated using currently observed environment, phenotype, and genotype data.

1. Introduction

Understanding how plants grow is crucial for predicting biomass production and maturity; these important agronomic traits are directly linked to crop yield. However, reliably estimating these traits is challenging because plant growth is sensitive to the environment, and varies across species and cultivars. Robustly capturing the relationships between genotype, environment, and organismal phenotype demands extensive data.

Traditionally, efforts to build such predictive models have been hindered by the cost and labor involved in collecting detailed phenotypic and environmental information. In this study, we deploy an automated robotic system managing a large-scale hydroponic production facility growing lettuce and spinach varieties. These robots systematically gather high-dimensional environmental and phenotypic data across thousands of growing trays daily.

Leveraging this comprehensive dataset, we develop HINTS (Harvest Intelligence from Neural Time Series), a generative model that links environmental and phenotypic inputs to key outcomes like growth rates and harvest mass. Our approach uniquely benefits from self-supervision, where the natural physical growth process provides the supervisory signal. Models trained on this data can generate growth parameters, which can be used to predict plant height and weight at harvest. This approach can fully leverage automated robotic phenotyping and environmental sensing to understand plant biology.

---

[1] Hippo Harvest. Correspondence to: Adam J Riesselman <adam@hippoharvest.com>, Wim Meeussen <wim@hippoharvest.com>.

Proceedings of the Workshop on Generative AI for Biology at the 42^{nd} *International Conference on Machine Learning*, Vancouver, Canada. PMLR 267, 2025. Copyright 2025 by the author(s).

2. Related Work

A standard approach to estimating plant maturity is growing degree days: This heuristic models plant growth as the accumulation of heat each day throughout the life cycle, with species or variety specific temperature thresholds (de Réaumur, 1735; Miller et al., 2001). These approaches have been further refined by incorporating additional constraints, like plant stress (Hatfield et al., 2011), and environmental data (Brechner et al., 1996). In addition, statistical models of plant growth (Hunt, 1982) and machine learning methods (Sharmin et al., 2025) can improve estimates agricultural productivity.

The development of accurate models require substantial amounts of quality data. High-throughput phenotyping facilities automate this data collection process (Abebe et al., 2023; Atefi et al., 2021) using a combination of computer vision and sensor data to estimate plant growth and maturity (Li et al., 2021; Xu et al., 2022). However, these facilities are expensive to build and maintain, and are not integrated into large-scale commercial operations.

Extracting useful features to represent biology from vast datasets is also a challenge (Wang et al., 2023). Self-supervised learning has been frequently employed to build biological insights without extensive labeled data (Liu et al., 2021; Lopez et al., 2018; Notin et al., 2024; Lu et al., 2019; Brixi et al., 2025), and has been incorporated into industrial phenotyping efforts in biomedical research (Fay et al., 2023; Sivanandan et al., 2023). These computational approaches, when paired with robotic data acquisition, make them scalable biology learners.

3. Methods

3.1. Data collection

Lettuce and spinach are cultivated hydroponically in an approximately 1 m^2 growing tray across an 3000 m^2 sized facility (Figure 1). Hydroponic water volume is managed by custom-outfitted freight robots (Meeussen et al., 2021): one set of robots determines water volume in each tray, while another set delivers precise amounts of water and nutrients to sustain growth.

Additional robots capture environmental and phenotypic data (Meeussen et al., 2024). One robot is outfitted with a photosynthetically active radiation (par) meter (Apogee Instruments, 2025), measuring light intensity across the growspace (Figure 1). Using additional machine learning techniques (Riesselman & Meeussen, 2023), we impute sunlight and LED intensity at 15 minute intervals for each growing tray throughout its lifecycle. Ambient air temperature and humidity are also recorded at regular intervals across the growspace.

Phenotypic information is gathered using Intel RealSense D455 depth cameras (Intel Corporation) mounted on a mobile robotic platform that traverses the growing area daily (Figure 1). The images are automatically cropped to the boundaries of individual growing trays, treating each as a distinct sample. From the depth information, we extract plant height measurements, summarized as the median height in centimeters across the growing tray. Additionally, we featurize the height distribution across the tray by calculating 10 height deciles to capture the variability of plant canopy composition.

Paired environment and height measurement data are recorded daily for each growing tray. Environmental data (E) are summarized by calculating daily minimum, mean, and maximum values for air temperature and humidity. We also compute the daily light integral (DLI) separately for sunlight and LED lighting systems. These environmental features are paired with the corresponding daily height measurements (H). From transplant 0 to harvest D, we organize these observations as sequences $E = \{e_0, e_1, ..., e_D\}$ and $H = \{h_0, h_1, ..., h_D\}$. Finally, the mass of harvested leaf material (m_D) is recorded in grams by weighing the growing tray before and after harvest. This results in 657663 growing days of environmental data, 639352 phenotypic data points, and 28410 total harvested growing trays.

3.2. Height growth curve

We parameterize plant growth trajectories with a differentiable growth curve by summarizing observed plant height with two key parameters (Figure 2): seedling lag (β_{lag}) and growth rate (β_{gr}):

$$\hat{h}_d = \beta_{\text{gr}} \cdot \text{softplus}(\text{age}_d - \beta_{\text{lag}}) = \beta_{\text{gr}} \cdot \ln(1 + e^{(\text{age}_d - \beta_{\text{lag}})}) \tag{1}$$

Where $d \in \{0, 1, ..., D\}$ indexes the measurement day, age_d is the number of days after transplant, β_{lag} represents the initial lag period during which the plant establishes itself before detectable vertical growth begins, and β_{gr} determines the rate of height increase after this establishment phase. While height accumulation of lettuce and spinach does ultimately plateau, we harvest it before this point.

We regress out the effect of sun elevation on observed plant height, as it is a significant covariate in observed height measurements:

$$h_d^{\text{corrected}} = h_d - \beta \cdot \mathbb{1}_{\text{sun up}} \cdot h_d \tag{2}$$

Where β is a learned parameter and $\mathbb{1}_{\text{sun up}}$ is an indicator function that equals 1 when the sun is above the horizon.

3.3. Canopy mass

To estimate canopy mass, we learn a relative canopy density parameter that relates plant height to mass (Figure 2). Here we assume the density is the same at any point in its lifecycle, and that canopy composition is isotropic. Harvested leaf length l_D is calculated by subtracting the cut height c from the total plant height h_D:

$$l_D = h_D - c \tag{3}$$

Cut height c is set per growing tray for desired leaf product specifications. We predict harvest mass m_D as the product of l_D and a learned density parameter β_{mass}:

$$\hat{m}_D = l_D \cdot \beta_{\text{mass}} \tag{4}$$

This parameterization uses leaf length as a proxy for the volume of harvestable product, with the density β_{mass} only observed at harvest.

3.4. Parameter estimation

We can directly estimate these parameters individually for each growing tray when harvested using the collected plant heights $H_{0:D}$, ages of height observations $A_{0:D}$ and harvest mass m_D:

$$P(H_{0:D}, m_D | A_{0:D}, \phi) \tag{5}$$

where $\phi = \{\beta_{\text{lag}}, \beta_{\text{gr}}, \beta_{\text{mass}}\}$ and $\phi \geq 0$.

3.5. Deep generative model

While direct parameter estimation at harvest time is straightforward, our goal is to predict these parameters in the middle of a plant's lifecycle, which in turn can be used to project future growth. To accomplish this, we model the relationship between observed data and growth parameters as a nonlinear function f with parameters θ:

$$f_\theta(g, E_{0:d}, \tilde{H}_{0:d}) = [\beta_{\text{lag},d}, \beta_{\text{gr},d}, \beta_{\text{mass},d}] = \phi'_d \tag{6}$$

where g represents genotype information, $E_{0:d} = \{e_0, e_1, ..., e_d\}$ represents all environmental observations up to day d, and $\tilde{H}_{0:d} = \{\tilde{h}_0, \tilde{h}_1, ..., \tilde{h}_d\}$ represents all height feature observations up to day d. We parameterize the function f_θ using a unidirectional LSTM (Hochreiter & Schmidhuber, 1997; Merrill et al., 2024). All generated growth parameters ϕ' are constrained to be positive through a softplus transformation. At each time step d, the model takes all environmental $E_{0:d}$ and height feature $\tilde{H}_{0:d}$ data observed from the beginning of the cycle up to that point; missing height data are masked with zeros (Figure 1). Growth parameters ϕ' are generated at each step, $[\phi'_0 ... \phi'_d]$, conditioned on previously observed environmental, phenotypic, and genetic data.

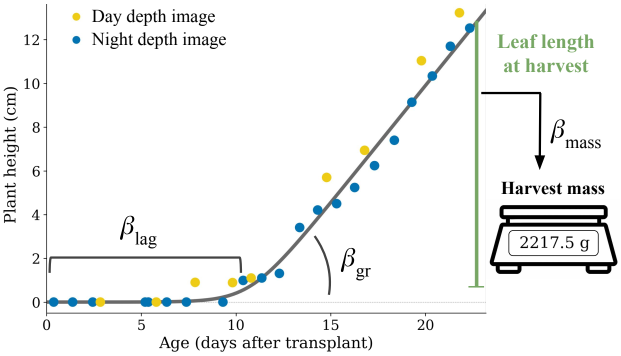

Figure 2. Phenotypic observations and growth parameterization of a harvested growing tray. β_{lag} and β_{gr} map plant age (age_d to estimated height. Harvest mass m_D is estimated from the leaf length at harvest and β_{mass}.

These parameters can then be used to transform plant age (days after transplant) at time t into predicted height $\hat{h}_{dt}$:

$$\hat{h}_{dt} = \beta_{\text{gr},d} \cdot \text{softplus}(\text{age}_t - \beta_{\text{lag},d}) \tag{7}$$

To predict leaf mass at harvest, we employ a more flexible vector-based approach to model variable leaf densities along different sections of the plant:

$$\hat{m}_D = \sum_{j=1}^{k} (\boldsymbol{\beta}_{\text{mass},d,j} \cdot \mathbf{I}_j) \tag{8}$$

Where $\boldsymbol{\beta}_{\text{mass},d} \in \mathbb{R}^k$ is a vector of k canopy density parameters output by the model at day d, and $\mathbf{I} \in \{0, 1\}^k$ is a vector of indicator variables representing the median harvested leaf length per tray l_{dt}. This formulation allows realization of variable canopy density across different sections of the leaf but allow for variable cut height c, which is not input directly into the model, to prevent overfitting.

Collectively, these components define a model that estimates the conditional probability distribution:

$$P(H_{0:D}, m_D | E_{0:d}, \tilde{H}_{0:d}, A_{0:d}, g; \theta) \tag{9}$$

where $H_{0:D}$ represents median plant height and $\tilde{H}_{0:d}$ represents height features (height deciles), respectively, for the entire cycle length, from time 0 to harvest day D. The model is end-to-end differentiable, optimizing LSTM parameters while also enforcing structure and interpretability in phenotype characterization through ϕ'.

3.6. Objective function and training

We implemented our model using a 3 layer residual LSTM architecture in Pytorch following (Beck et al., 2024). The LSTM consists of 3 layers with a dropout rate of 0.05. Input features from environment, height observations, and genotype are embedded into a common 128-dimensional representation space (Appendix).

The overall objective function for our model combines losses for plant height and harvest mass predictions with regularization terms that provide disambiguation when limited observations are available:

$$\mathcal{L} = \mathcal{L}_{\text{height}} + \mathcal{L}_{\text{mass}} + \mathcal{L}_{\text{prior}} \tag{10}$$

Where $\mathcal{L}_{\text{height}}$ and $\mathcal{L}_{\text{mass}}$ are Pseudo-Huber losses (Charbonnier et al., 1997) for plant height and harvest mass predictions respectively, and $\mathcal{L}_{\text{prior}}$ encompasses four regularization terms. We use Pseudo-Huber loss due to its robustness against outliers in our robot-collected data. The complete mathematical formulation and hyperparameters of all six individual loss terms are provided in the Appendix.

We trained the model using the AdamW optimizer (Loshchilov & Hutter, 2017) with a learning rate of 0.001 and early stopping with a patience of 10 epochs.

4. Experiments

A key operational challenge in hydroponic production is accurately forecasting harvest mass with sufficient lead time to meet customer fulfillments. In our system, we need to predict the harvest mass 5 days in advance of the actual harvest date to ensure accurate order fulfillment.

For each growing tray in the growspace, we collate all environmental and phenotypic data available up until the planning date, and use these data in conjunction with the genotype features to predict plant growth and relative canopy density parameters. We then project height and harvest mass 5 days into the future. As a baseline, we compare the HINTS-predicted harvest height and mass estimates to those from rolling N-day genotype-specific historic parameter averages for β_{lag}, β_{gr}, and m_D for $N \in \{10, 30, 90\}$. We evaluated the performance of HINTS and each baseline using 1989 lettuce growing trays harvested in April of 2025.

As shown in Figure 3, HINTS outperforms each of the N-day rolling average parameter baselines. Specifically, HINTS outperforms the 10-, 30-, and 90-day baseline predictions for plant height at harvest by 29.92%, 45.92%, and 67.37% mean absolute error. It similarly outperformed the baseline predictions for harvest mass (g) by 33.93%, 45.99%, and 67.98% respectively. This demonstrates HINTS's ability to successfully recover growth dynamics and harvest characteristics by combining partial growth trajectories with environmental features.

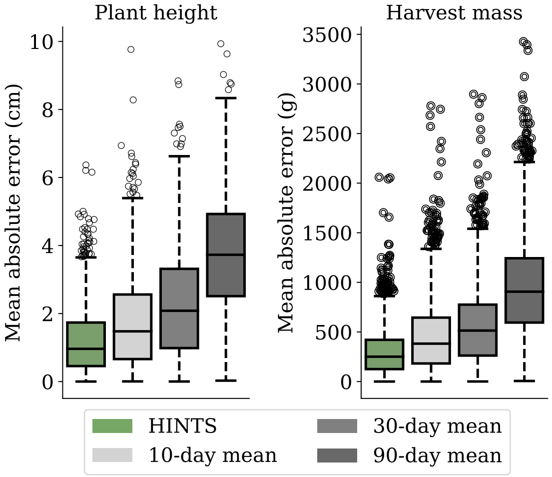

Figure 3. Absolute error comparison of HINTS and a baseline using N-day rolling average parameters. Performance is evaluated using $1,989$ lettuce growing trays

5. Conclusion

We present a generative, self-supervised model, HINTS, that accurately forecasts plant height and harvest mass in a large-scale hydroponic system by leveraging environmental, phenotypic, and genotypic data. This approach outperforms conventional heuristic baselines using rolling averages for parameters, achieving at least a 29.92% improvement in harvest height prediction, and 33.93% harvest mass prediction. Accurately predicting these phenotypes enables successful production planning while simultaneously minimizing wasted resources and food.

Importantly, HINTS maintains interpretability through biologically meaningful parameters (e.g. β_{lag}, β_{gr}), allowing operators to reason about predictions and optimize growing conditions proactively. These findings demonstrate the utility of combining automated phenotyping with structured deep learning for controlled environment agriculture.

Impact Statement

This work addresses critical challenges in food security and climate change by improving our ability to predict and optimize crop yields in controlled environment agriculture. Methods herein enable more accurate predictions of plant growth and resource allocation as well as reduce food waste.

Acknowledgements

We would like to thank the entire Hippo Harvest team for their contributions to this work, particularly Eitan Marder-Eppstein, Chelsea Wirth, Kevin Hannegan, Zack Lewis, Lucas Becker-Lipton, and Teka Nicholas.

References

Abebe, A. M., Kim, Y., Kim, J., Kim, S. L., and Baek, J. Image-based high-throughput phenotyping in horticultural crops. *Plants*, 12(10):2061, 2023.

Apogee Instruments. Sq-100x-ss original quantum sensor. https://www.apogeeinstruments.com/sq-100x-ss-original-quantum-sensor/, 2025. Accessed: 2025-05-12.

Atefi, A., Ge, Y., Pitla, S., and Schnable, J. Robotic technologies for high-throughput plant phenotyping: Contemporary reviews and future perspectives. *Frontiers in plant science*, 12:611940, 2021.

Beck, M., Pöppel, K., Spanring, M., Auer, A., Prudnikova, O., Kopp, M., Klambauer, G., Brandstetter, J., and Hochreiter, S. xlstm: Extended long short-term memory. *arXiv preprint arXiv:2405.04517*, 2024.

Brechner, M., Both, A., and Staff, C. Hydroponic lettuce handbook. *Cornell Controlled Environment Agriculture*, 834:504–509, 1996.

Brixi, G., Durrant, M. G., Ku, J., Poli, M., Brockman, G., Chang, D., Gonzalez, G. A., King, S. H., Li, D. B., Merchant, A. T., et al. Genome modeling and design across all domains of life with evo 2. *BioRxiv*, pp. 2025–02, 2025.

Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on image processing*, 6(2):298–311, 1997.

de Réaumur, R. Observation du thermometer, faites à paris pendant l'année 1735, compares avec celles qui ont été faites sous la ligne, à l'isle de france, à alger et en quelques-unes de nos isles de l'amérique. *Mémoires de l'Académie des Sciences de Paris (1735)*, 1735.

Fay, M. M., Kraus, O., Victors, M., Arumugam, L., Vuggumudi, K., Urbanik, J., Hansen, K., Celik, S., Cernek, N., Jagannathan, G., et al. Rxrx3: Phenomics map of biology. *Biorxiv*, pp. 2023–02, 2023.

Hatfield, J. L., Boote, K. J., Kimball, B. A., Ziska, L., Izaurralde, R. C., Ort, D., Thomson, A. M., and Wolfe, D. Climate impacts on agriculture: implications for crop production. *Agronomy journal*, 103(2):351–370, 2011.

Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Hunt, R. *Plant growth curves. The functional approach to plant growth analysis.* 1982.

Intel Corporation. Intel® realsense™ depth camera d455. https://www.intelrealsense.com/depth-camera-d455/. Accessed: 2025-05-12.

Li, D., Quan, C., Song, Z., Li, X., Yu, G., Li, C., and Muhammad, A. High-throughput plant phenotyping platform (ht3p) as a novel tool for estimating agronomic traits from the lab to the field. *Frontiers in Bioengineering and Biotechnology*, 8:623705, 2021.

Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., and Tang, J. Self-supervised learning: Generative or contrastive. *IEEE transactions on knowledge and data engineering*, 35(1):857–876, 2021.

Lopez, R., Regier, J., Cole, M. B., Jordan, M. I., and Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nature methods*, 15(12):1053–1058, 2018.

Loshchilov, I. and Hutter, F. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

Lu, A. X., Kraus, O. Z., Cooper, S., and Moses, A. M. Learning unsupervised feature representations for single cell microscopy images with paired cell inpainting. *PLoS computational biology*, 15(9):e1007348, 2019.

Meeussen, W., Marder-Eppstein, E., and Boenig, A. Grow method and system, August 26 2021. US Patent App. 17/182,214.

Meeussen, W., Wirth, C., Riesselman, A., Jones, E., and LaRue, T. Grow space integration for mobile robots, October 31 2024. US Patent App. 18/771,643.

Merrill, W., Petty, J., and Sabharwal, A. The illusion of state in state-space models. *arXiv preprint arXiv:2404.08819*, 2024.

Miller, P., Lanier, W., and Brandt, S. Using growing degree days to predict plant stages. *Ag/Extension Communications Coordinator, Communications Services, Montana State University-Bozeman, Bozeman, MO*, 59717(406):994–2721, 2001.

Notin, P., Rollins, N., Gal, Y., Sander, C., and Marks, D. Machine learning for functional protein design. *Nature biotechnology*, 42(2):216–228, 2024.

Riesselman, A. and Meeussen, W. High-resolution environmental sensor imputation using machine learning, November 30 2023. US Patent App. 18/325,742.

Sharmin, S., Hossan, M. T., and Uddin, M. S. A review of machine learning approaches for predicting lettuce yield in hydroponic systems. *Smart Agricultural Technology*, pp. 100925, 2025.

Sivanandan, S., Leitmann, B., Lubeck, E., Sultan, M. M., Stanitsas, P., Ranu, N., Ewer, A., Mancuso, J. E., Phillips, Z. F., Kim, A., et al. A pooled cell painting crispr screening platform enables de novo inference of gene function by self-supervised deep learning. *bioRxiv*, pp. 2023–08, 2023.

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., et al. Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972):47–60, 2023.

Xu, Y., Zhang, X., Li, H., Zheng, H., Zhang, J., Olsen, M. S., Varshney, R. K., Prasanna, B. M., and Qian, Q. Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. *Molecular Plant*, 15(11):1664–1695, 2022.

A. Appendix

A.1. Growspace facility description

The growspace consists of greenhouses outfitted with LEDs, shades, and climate control systems. Growing trays with plants are organized into a grid and occupy as much growspace real estate as possible. Robots travel beneath the trays: gaps between columns give space for robot components to be mounted on a frame above the plants. This configuration is built for in situ robot operations, like imaging, weighing, or environmental sensing.

Figure 4. A view of a hydroponic leafy greens growspace.

A.2. Detailed Loss Function Formulation

The model's objective function consists of six distinct terms, with loss weights applied to the prior terms as follows:

$$\mathcal{L} = \mathcal{L}_{\text{height}} + \mathcal{L}_{\text{mass}} + 0.01 \cdot \mathcal{L}_{\text{embed}} + 0.1 \cdot \mathcal{L}_{\text{density}} + 0.01 \cdot \mathcal{L}_{\text{lag}} + 0.1 \cdot \mathcal{L}_{\text{growth}} \tag{11}$$

The individual terms are defined as follows:

1. $\mathcal{L}_{\text{height}}$ is the Pseudo-Huber loss for plant height predictions:

   $$\mathcal{L}_{\text{height}} = \sum_d \mathcal{H}_{\delta_h}(h_d^{\text{corrected}} - \hat{h}_d) \tag{12}$$

   with $\delta_h = 3$ cm (transition point from quadratic to linear loss)

2. $\mathcal{L}_{\text{mass}}$ is the Pseudo-Huber loss for harvest mass predictions:

   $$\mathcal{L}_{\text{mass}} = \mathcal{H}_{\delta_m}(m_D - \hat{m}_D) \tag{13}$$

   with $\delta_m = 2$ kg

   The Pseudo-Huber loss $\mathcal{H}_\delta$ is defined as:

   $$\mathcal{H}_\delta(x) = \delta^2 \left(\sqrt{1 + \left(\frac{x}{\delta}\right)^2} - 1 \right) \tag{14}$$

   This differentiable approximation of the Huber loss provides robustness against outliers while maintaining differentiability at all points. We use it for both height and mass predictions due to occasional spurious measurements in our robot-collected data.

3. $\mathcal{L}_{\text{lag}}$ is the L1 prior loss on the seedling lag parameter:

   $$\mathcal{L}_{\text{lag}} = |\beta_{\text{lag}} - 12| \tag{15}$$

   where β_{lag} represents the seedling lag parameter with prior mean 12

4. $\mathcal{L}_{\text{growth}}$ is the L1 prior loss on the growth rate parameter:

$$\mathcal{L}_{\text{growth}} = |\beta_{\text{gr}} - (-0.5)| \tag{16}$$

   where β_{gr} represents the growth rate parameter with prior mean -0.5

5. $\mathcal{L}_{\text{density}}$ is the L1 prior loss on the canopy density parameter:

$$\mathcal{L}_{\text{density}} = |\beta_{\text{mass}} - (-2)| \tag{17}$$

   where β_{mass} represents the canopy density parameter with prior mean -2

6. $\mathcal{L}_{\text{embed}}$ is the L1 regularization on the genotype embedding weights that project the genotype embedding to the hidden unit size:

$$\mathcal{L}_{\text{embed}} = \frac{1}{|\mathbf{W}|} \sum_{i,j} |W_{i,j}| \tag{18}$$

   where $\mathbf{W}$ is the weight matrix of the genotype embedding layer

In particular, priors on generated parameters β_{gr} and β_{lag} provide disambiguation when limited height observations are available, or all zeros.

Parameter	**Weight Coefficient**	**Prior Mean**	**Description**
Embedding prior	0.01	0	Genotype embedding regularization
Canopy density prior	0.1	-2	Weight for the canopy density parameter
Seedling lag prior	0.01	12	Weight for the seedling lag parameter
Growth rate prior	0.1	-0.5	Weight for the growth rate parameter

Table 1. Weight coefficients for the L1 priors in the objective function

A.3. Genotype Encoding

Genotype information is included into the model by MD5 hash-based encoding, where each plant variety string is hashed to produce a stable 16-byte representation. This vector is then normalized to the range [0,1] before being projected to the model's embedding space.