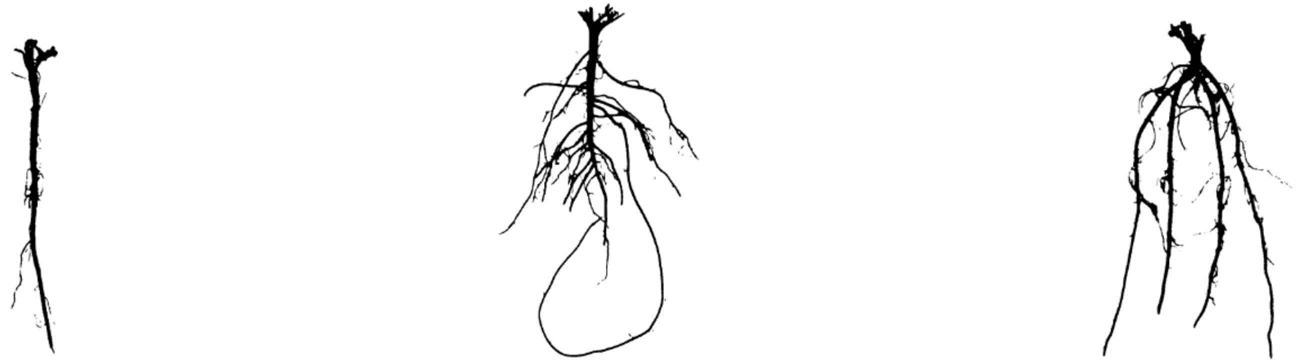
Phenotyping Alfalfa (*Medicago sativa* L.) Root Structure Architecture via Integrating Confident Machine Learning with ResNet-18

Brandon J. Weihs, Zhou Tang, Zezhong Tian, Deborah Jo Heuschele, Aftab Siddique, Thomas H. Terrill, Zhou Zhang, Larry M. York , Zhiwu Zhang, and Zhanyou Xu



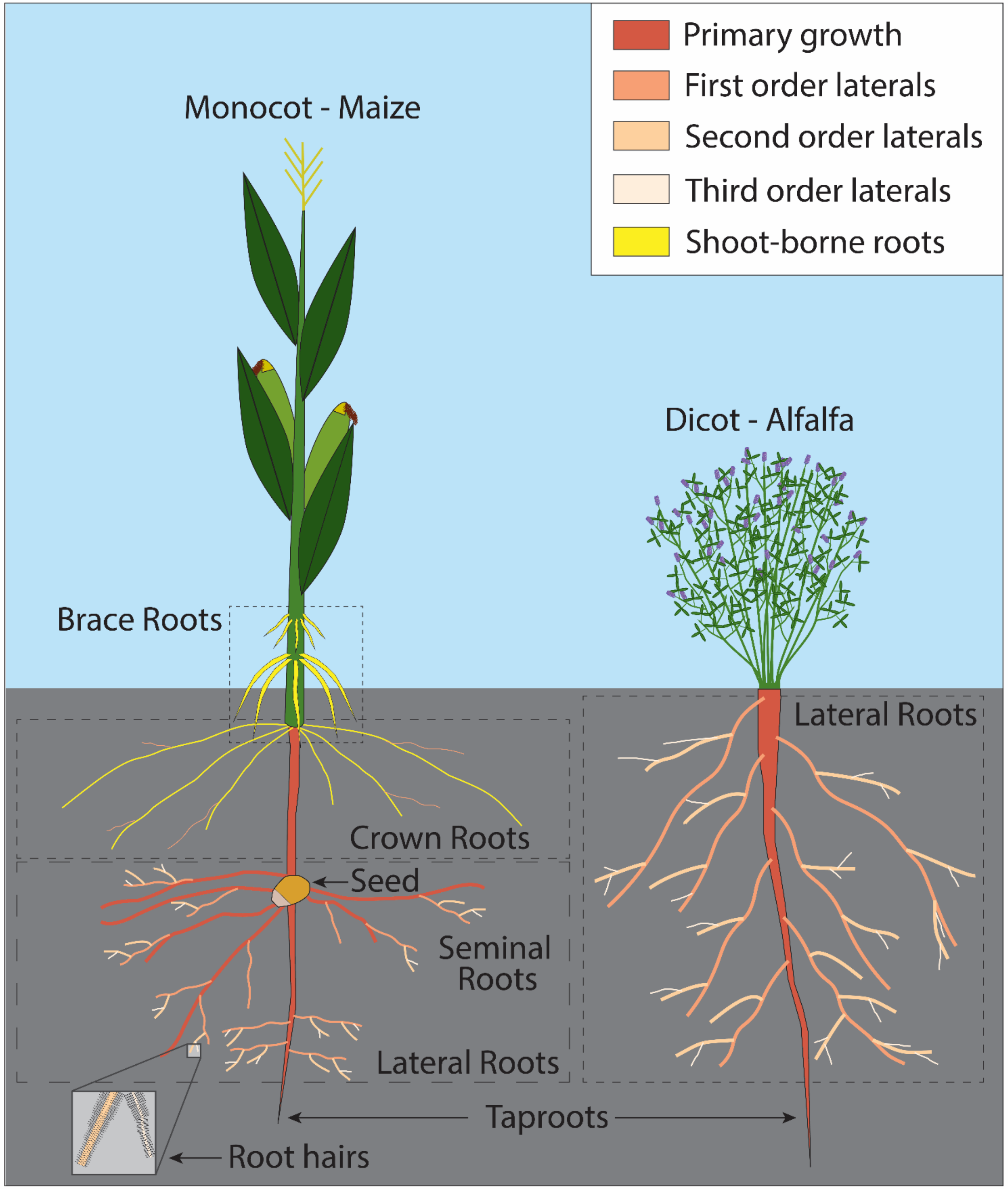
Why Alfalfa?

- Third most widely produced crop in U.S.
  - ~50 million tons harvested from 6.1 million hectares in 2021.
- Global dairy and beef industries rely on alfalfa for high protein feed.
- Atmospheric CO_2 and N fixation:
  - Deep roots (6-15 m) sequester CO_2 (helps mitigate GHG-caused climate change).
  - Symbiosis with soil bacteria fixes nitrogen.
    - Restores N-depleted soils.
    - Reduces the need for fossil fuel-based fertilizers.
- It's a perennial that typically persists/can be maintained for 4-6 years.
- 2nd Green Revolution goals of improving plants through their abiotic stress tolerance and nutrient/water acquisition/useability and efficiency…
  - RSA Ideotypes achieve these goals.
  - Some RSA research has already shown gains in yield, winter survival, and P uptake as a function of RSA type.

Why Root Structure Architecture?

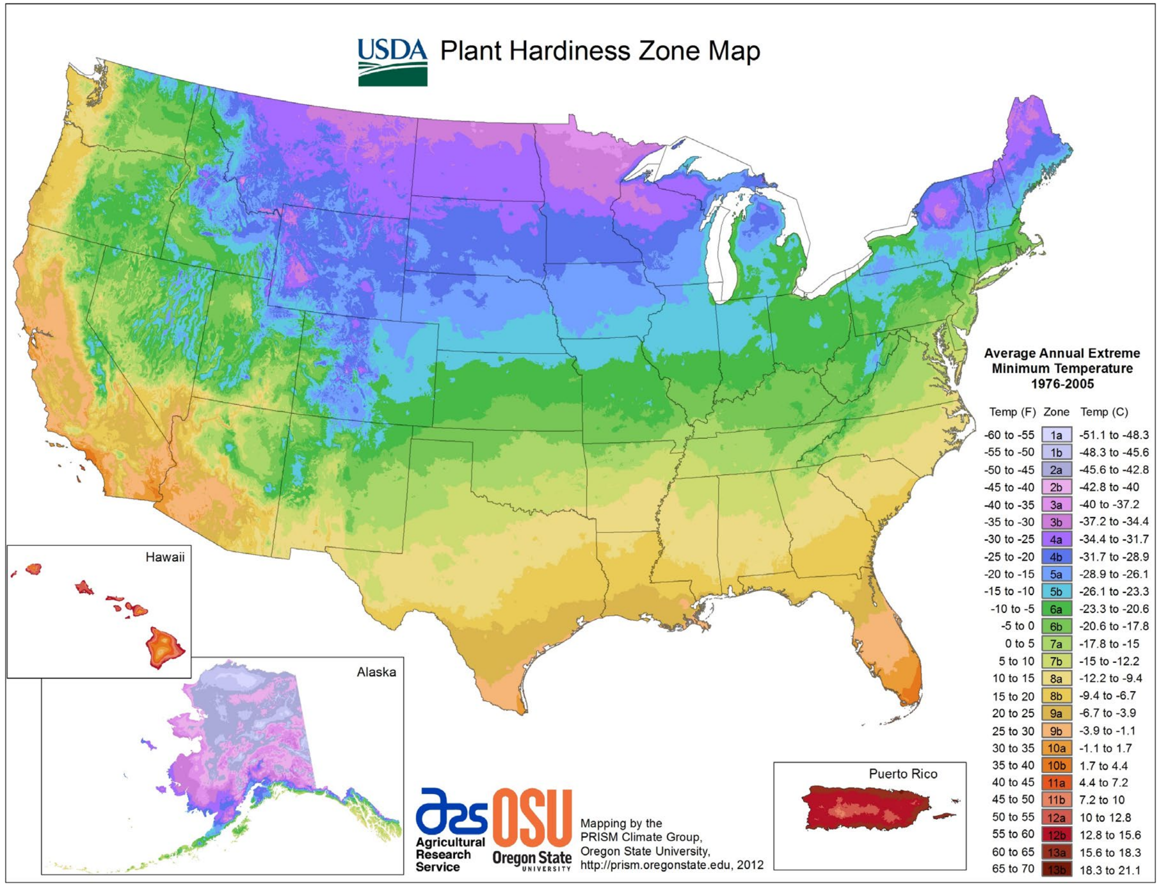
- Roots serve many functions for plants:
  - Structural support.
  - Water & nutrient acquisition organs.
  - Storage areas.
  - Symbiotic interfaces for relationships with other organisms.

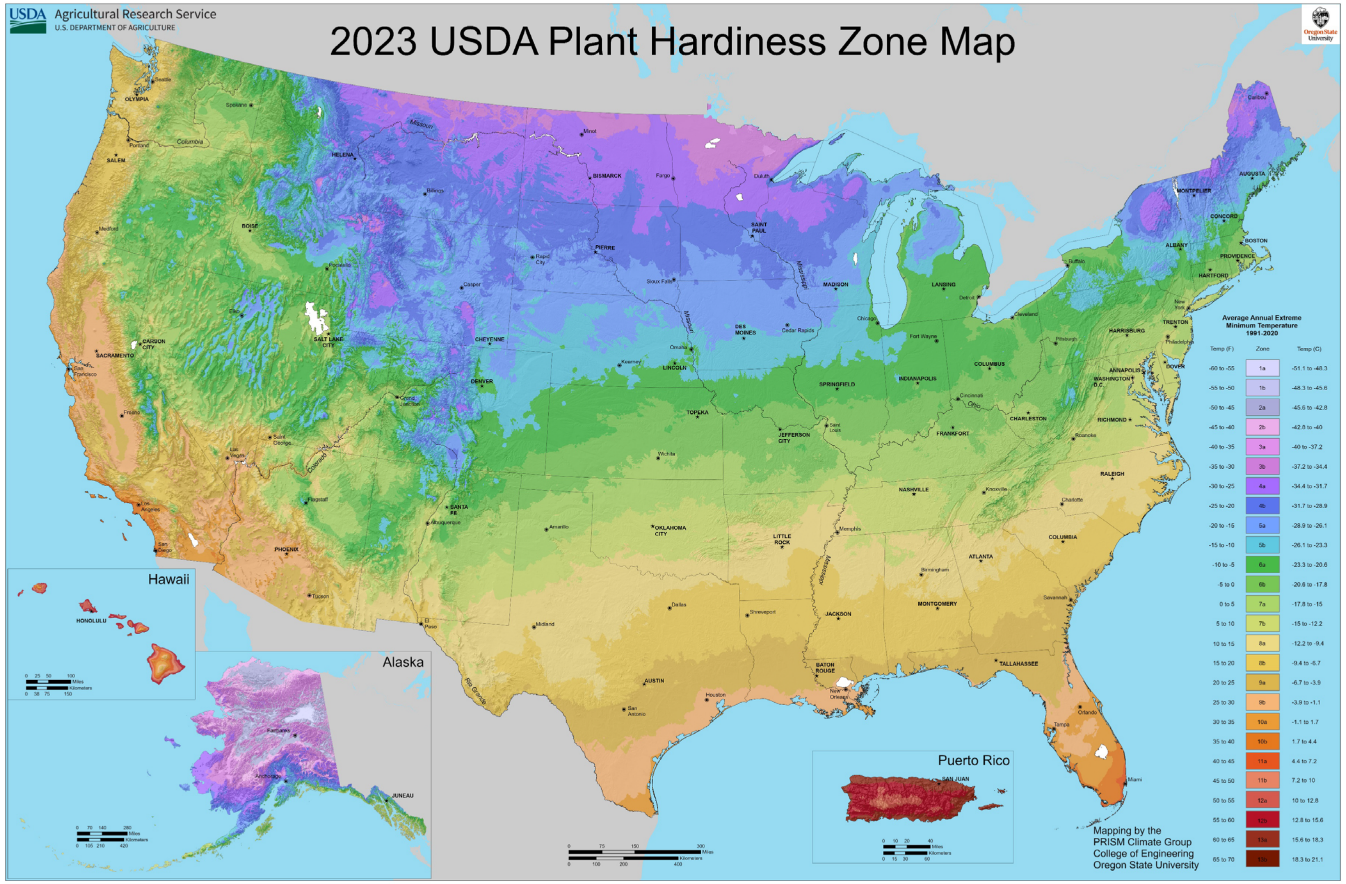
- Roots that perform optimally for a given set of conditions are the goal…Designer roots (ideotypes):
  - For improved yield.
  - For H_2O and N acquisition.
  - For P acquisition.
  - CO_2 sequestration.
  - Drought resistance.
  - Winter survival.
  - Fall dormancy.
  - Pest/disease/environmental conditions.
  - … and more.



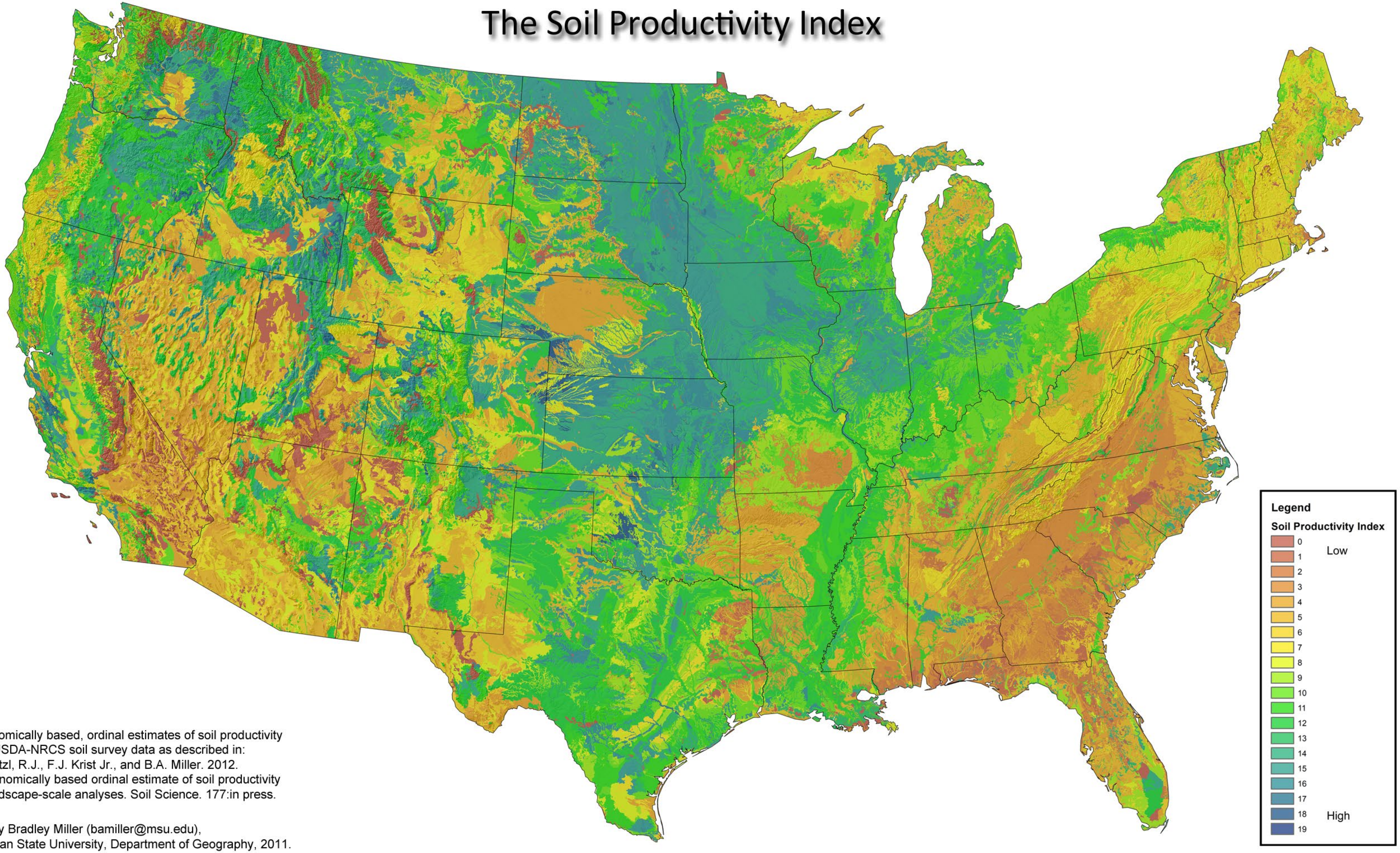
From Weihs et al. (2024)

Root Ideotypes are Needed to Address Climate Changes and Shifting Land Potential…





The Soil Productivity Index



Taxonomically based, ordinal estimates of soil productivity from USDA-NRCS soil survey data as described in:
Schaetzl, R.J., F.J. Krist Jr., and B.A. Miller. 2012.
A taxonomically based ordinal estimate of soil productivity for landscape-scale analyses. Soil Science. 177:in press.

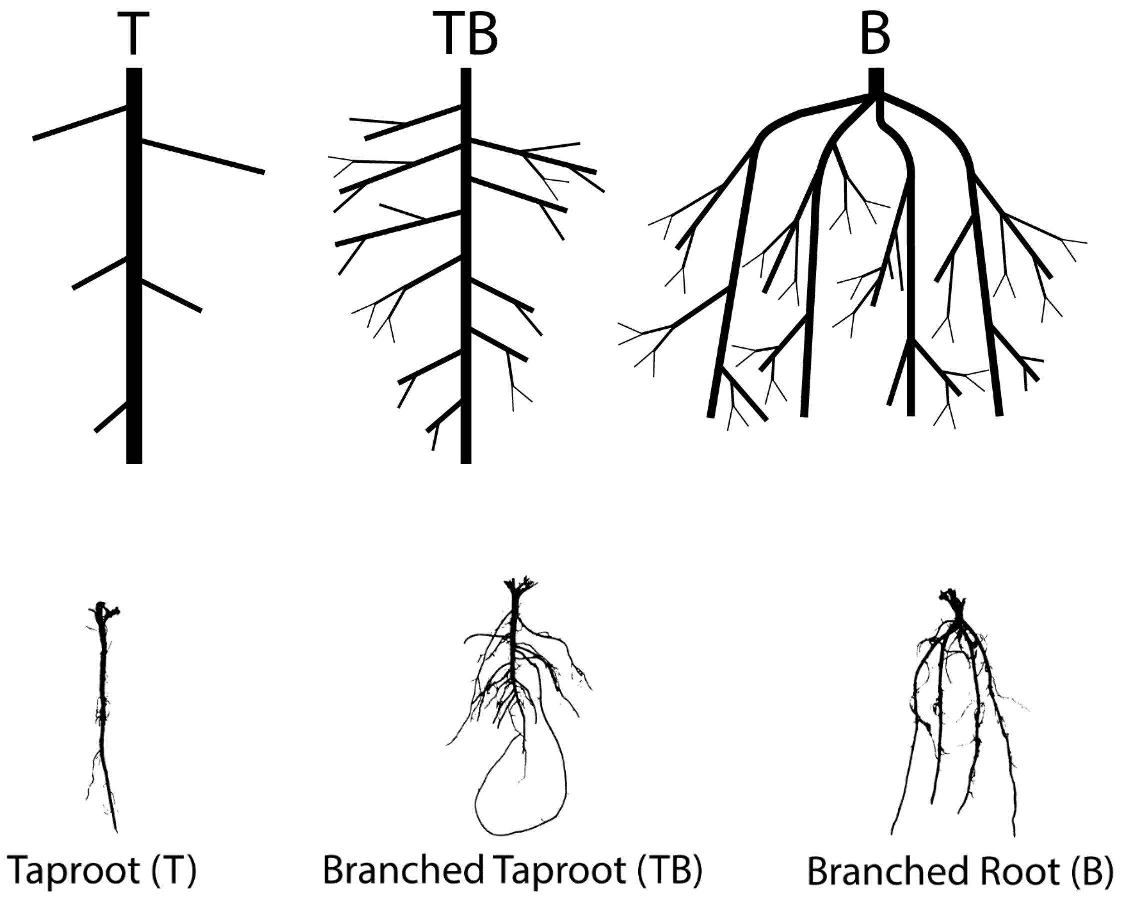
Map by Bradley Miller (bamiller@msu.edu),
Michigan State University, Department of Geography, 2011.

What is Root Structure Architecture (RSA)?

- RSA refers to the spatial configuration of the entire root system.
- RSA is composed of:
  - Morphology
    - The surface shape, pattern, and size of individual plant root parts:
      - Primary, secondary etc. growth.
      - Root epidermis characteristics including root hairs.
  - Topology
    - How individual plant root parts are connected in terms of axes and branching.
  - Distribution
    - How individual plant root parts are distributed within a root system.
      - Can be used to study biomass or length as a function of:
        - Soil depth
        - Distance from the stem…and more.

Study Objectives

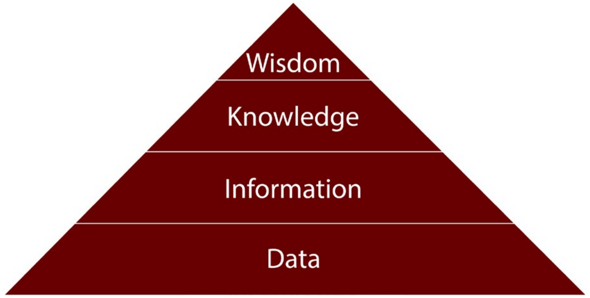
- Compare AI model predictions:
  - Random Forest + feature traits.
  - ResNet-18 + segmented images.
- Test the ability of confident machine learning (CL) and reactive machine learning (RL) to:
  - Minimize subjective labeling errors.
  - Improve labeling and prediction accuracies.



Root phenotypes from UMN4561 and UMN4563 fourth cycle progenies.

Why Artificial Intelligence?

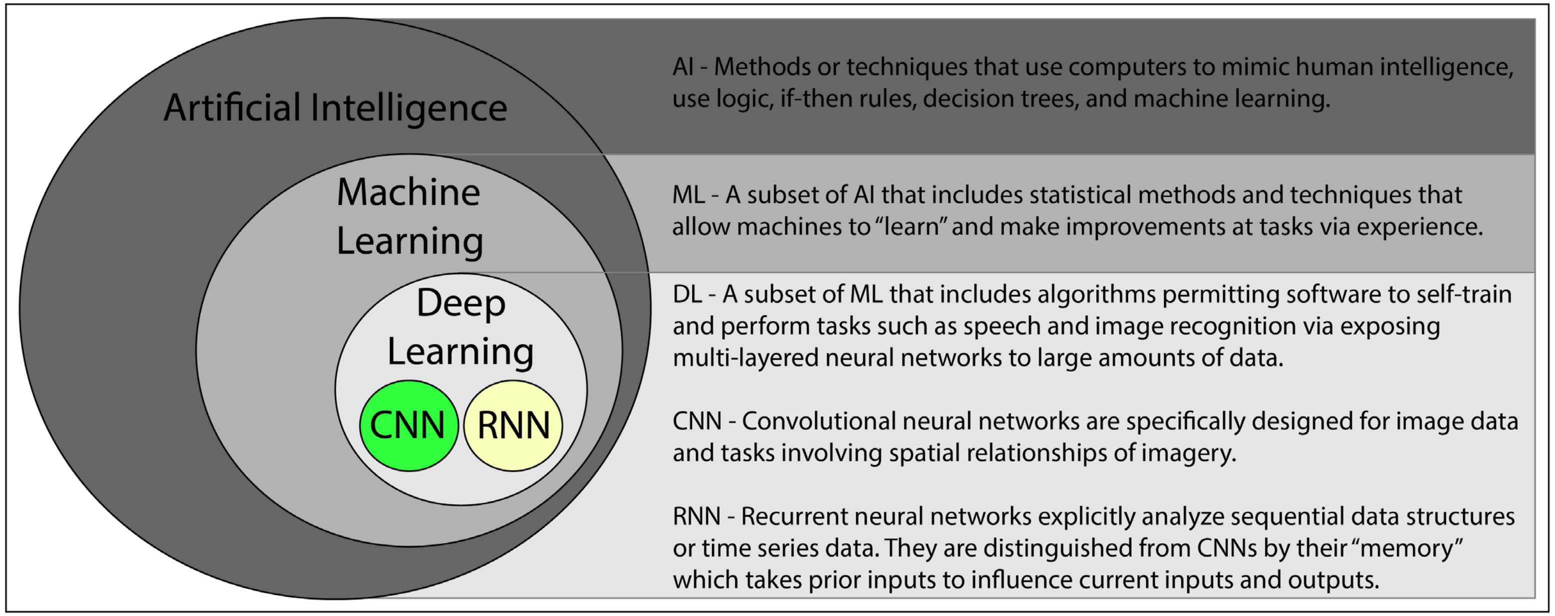
- Benefits of AI:
  - Speed.
  - Improved accuracy.
    - Reduced error (automation).
  - Capable of handling large amounts of data.
  - Ability to reveal patterns in data.
  - Mimic human intelligence.
  - Reduced subjectivity (human bias).
  - Can increase phenotypic selection speeds with early/rapid RSA analyses (~2-week-old plants, Bucciarelli et al., 2021).



Root structure architecture analyses coupled with AI is a push toward faster, more accurate, and less subjective phenotyping…

Experimental Design and Plant Materials

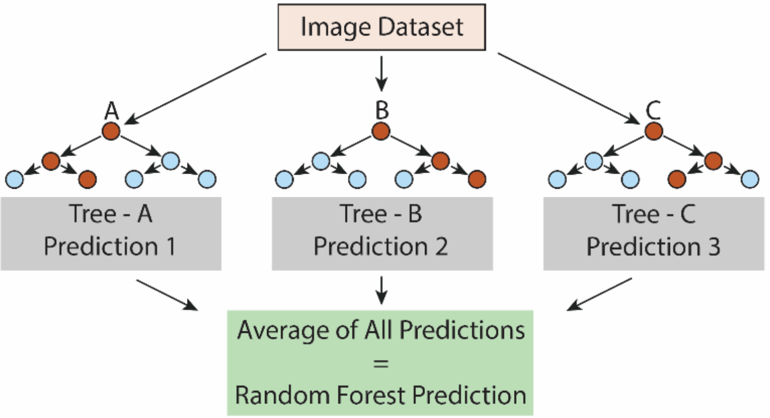
- Two image datasets:
  - St. Paul, Minnesota populations (617 images).
    - Bred for taproot type and branched type RSAs (UMN3233 & UMN3234).
  - Burneyville, Oklahoma population (264 images).
    - Commercial line (America’s Alfalfa Alfagraze 600 RR).
      - Sampled/studied by Mattupauli et al. (2019) for RSA changes regarding root rot disease.
  - Labeled by three experts using a protocol.
    - Taproot (T), Intermediate (TB), and Branched (B).
  - Segmented into binary images.
  - Image augmentation applied (881*10 images).



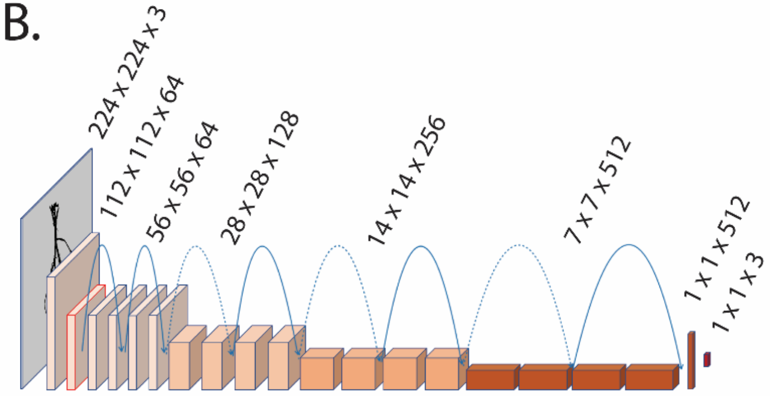
Models Tested: ResNet-18 and Random Forest

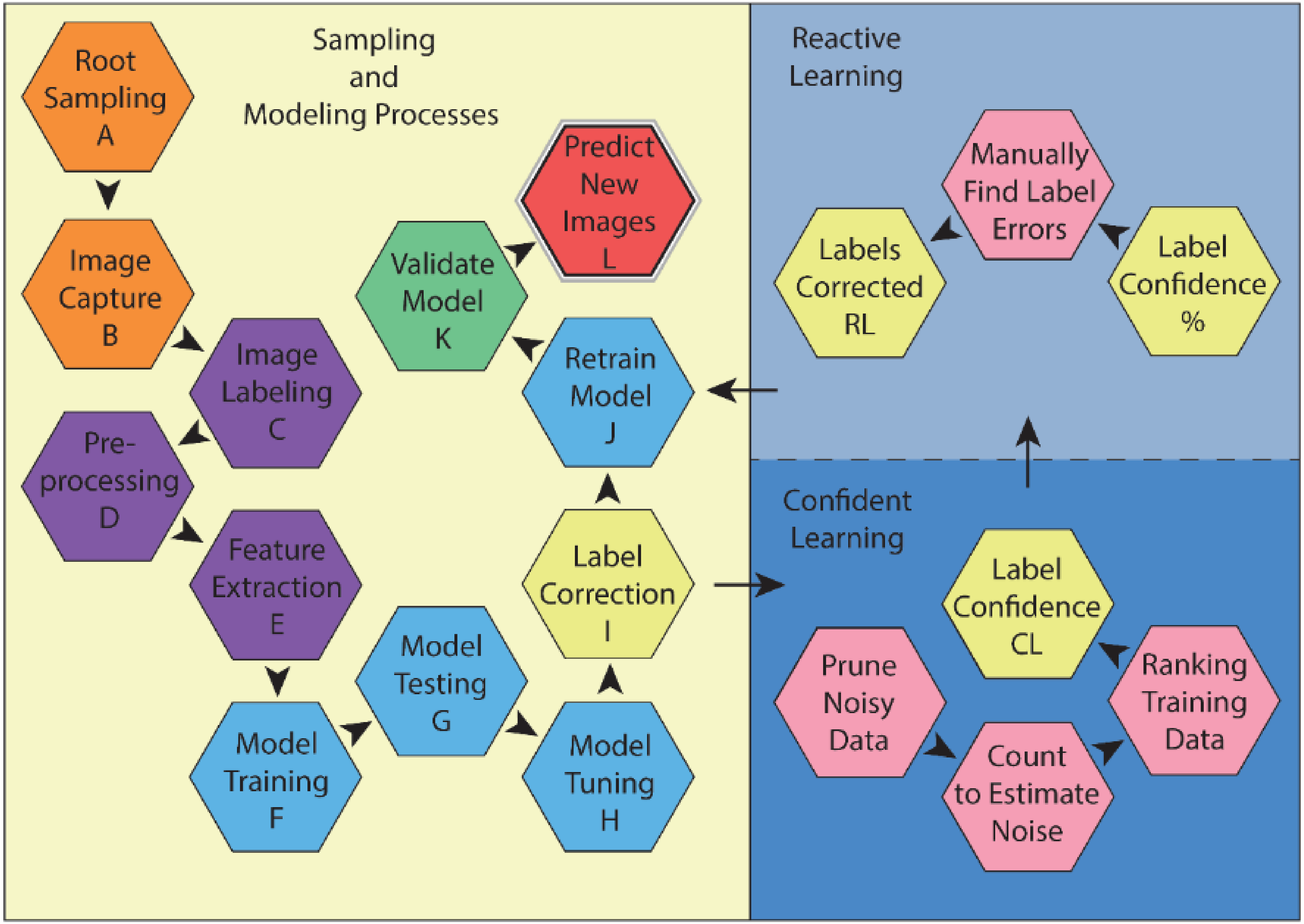
- ResNet:
  - Deep CNN (DNN)
  - Several model sizes to choose from (18, 34, 50, 101, and 152).
  - Residual "Res" network "Net".
    - Identity shortcut connections can bypass intermediate layers (solid lines in adjacent figure B).
  - Uses images (pixels) as input.
- Random Forest:
  - Supervised ML algorithm.
  - Randomly samples data and builds series of decision trees.
  - Uses an ensemble of decision trees to make predictions.
  - Can help reduce overfitting and bias.
  - Uses feature data (tabular data) as input.

A.



B.





Sampling
and
Modeling Processes
Root Sampling A
Image Capture B
Image Labeling C
Pre-processing D
Feature Extraction E
Model Training F
Model Testing G
Model Tuning H
Label Correction I
Retrain Model J
Validate Model K
Predict New Images L
Reactive Learning
Manually Find Label Errors
Labels Corrected RL
Label Confidence %
Confident Learning
Label Confidence CL
Prune Noisy Data
Ranking Training Data
Count to Estimate Noise

Confident Learning (CL) and Reactive Learning (RL)

- Confident learning (Northcutt et al., 2021) involves using the original model class prediction accuracies to determine label confidence:
  - Uses a class's probability threshold to determine label accuracy (confidence).
  - Labels identified as low confidence are subjected to Reactive Learning.
  - CL is a 3-step process:
    - Pruning noisy data (searching for mislabeled data)
    - Counting with probabilistic thresholds (training on clean data)
    - Ranking which data to use during training (training with confidence)
- Reactive Learning:
  - Image label error corrections based on CL analyses.
- Why CL+RL Methods?
  - We want **CLEAN/TRUE DATA!!!**
  - Garbage in = garbage out (model or data labels)

Datasets and Label Correction Combinations

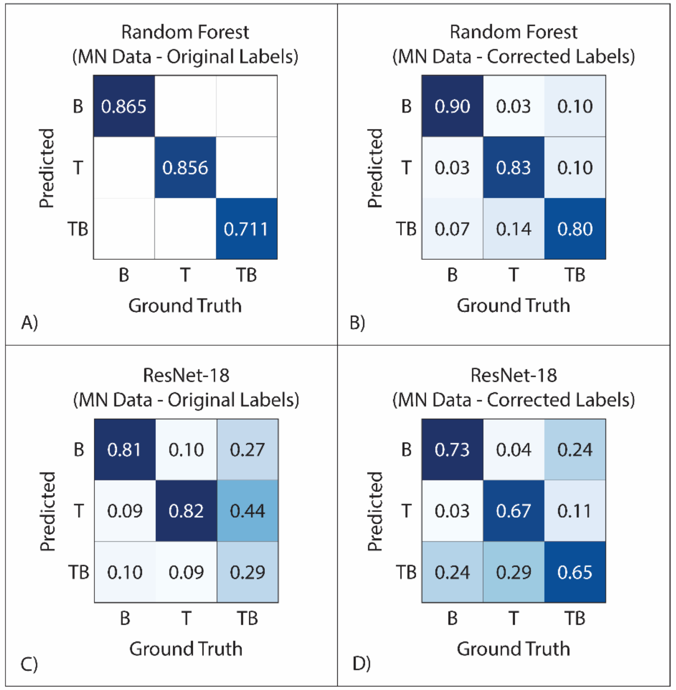
- Minnesota data only
  - Original labels
    - ResNet-18
    - Random Forest
  - Corrected labels
    - ResNet-18
    - Random Forest

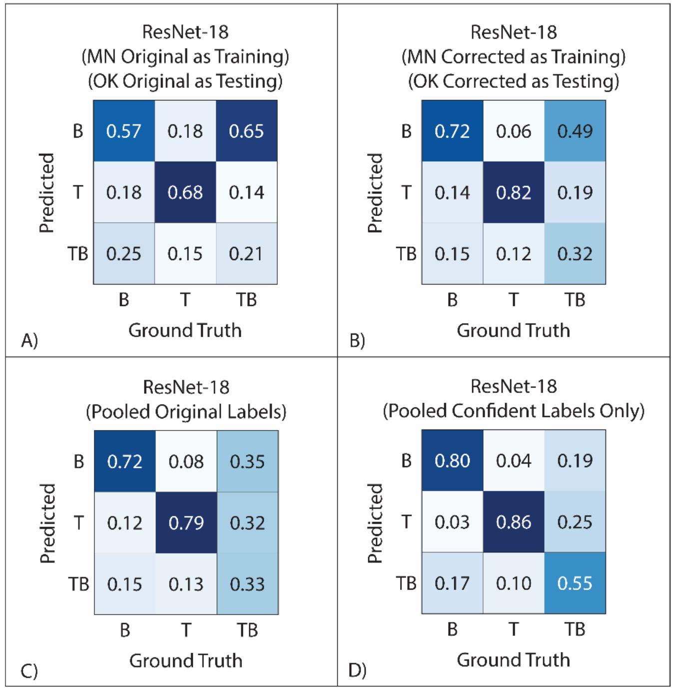
- ResNet-18 Cross-population
  - Eight permutations
    - Training/Testing
      - OK original/MN original
      - OK corrected/MN original
      - OK original/MN corrected
      - OK corrected/MN corrected
      - MN original/OK original
      - MN original/OK corrected
      - MN corrected/OK original
      - MN corrected/OK corrected

- ResNet-18 Pooled
  - Two permutations
    - Training/Testing
      - Pooled original labels
        - 881 images
      - Pooled original labels
        - Only confident labels
        - 608 images
        - Highest ResNet-18 overall accuracy in study (~75%)

Results

- CL algorithm:
  - Minorly improved the Random Forest prediction accuracies (~1%).
  - ResNet-18
    - Cross population prediction accuracy improved ~8-13%
- Highest accuracy data combinations:
  - CL/RL corrected datasets for predicting taproots (~86%).
  - Pooled dataset + CL (~75%).





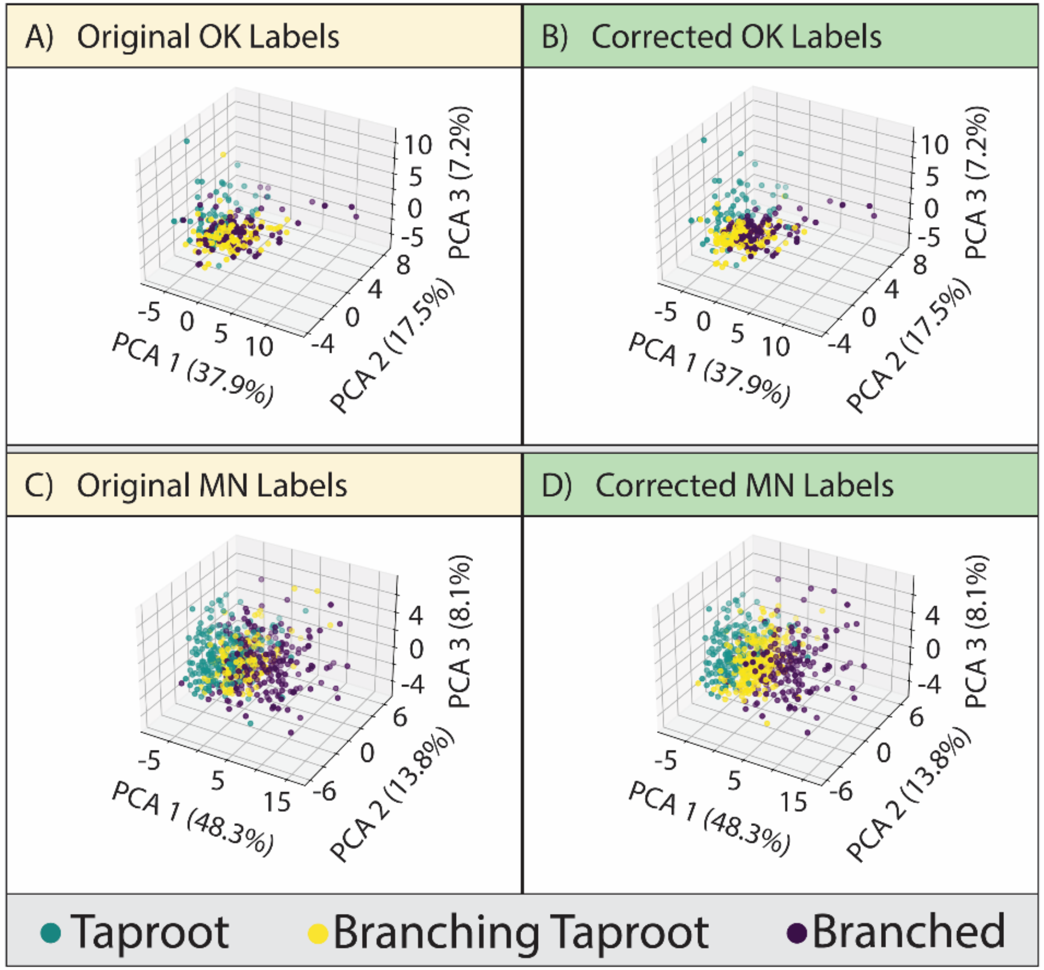
Model Prediction Accuracies

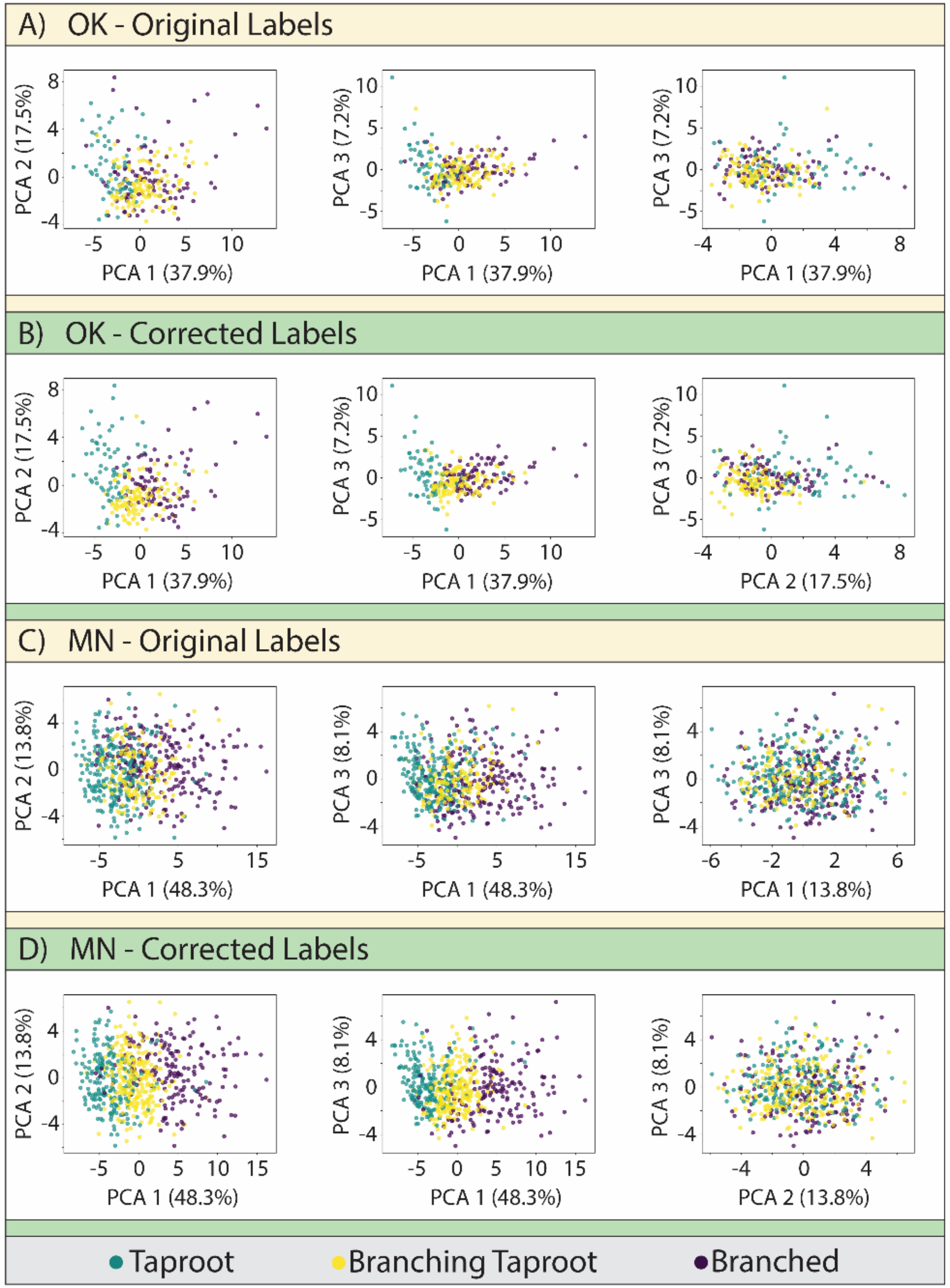
- Confident pooled data:
  - Highest overall accuracy using ResNet-18 (~75%).
  - Highest prediction accuracy for Taproot class RSA (86%).
- ResNet-18 and MN corrected label dataset.
  - Highest Intermediate class RSA (65%).

ID #	Input	Model and Dataset Combinations	Training Data	Testing Data	Branching (B)	Taproot (T)	Intermediate (TB)	Overall Accuracy	CL Applied	RL Applied	Analysis Level
1	Feature Data (38)	Random Forest - MN data only (n=617)	MN original labels	MN original labels	0.865	**0.856**	0.711	0.828	No	No	I
2	Feature Data (38)	Random Forest - Current Study - MN data (n=617)	MN corrected labels	MN corrected labels	**0.9**	0.83	**0.8**	0.838	Yes	Yes	II
3	Image	ResNet-18 - MN data only (n=617)	MN original labels	MN original labels	0.81	0.82	0.29	0.700	No	No	I
4	Image	ResNet-18 - MN data only (n=617)	MN corrected labels	MN corrected labels	0.73	0.67	**0.65**	0.679	Yes	Yes	II
5	Image	ResNet-18 - Cross-population MN (n=617) and OK (n=264)	MN original labels	OK original labels	0.57	0.68	0.21	0.447	No	No	I
6	Image	ResNet-18 - Cross-population MN (n=617) and OK (n=264)	OK original labels	MN original labels	0.58	0.53	0.27	0.492	No	No	I
7	Image	ResNet-18 - Cross-population MN (n=617) and OK (n=264)	MN corrected labels	OK original labels	0.73	0.55	0.17	0.458	Yes	Yes	II
8	Image	ResNet-18 - Cross-population MN (n=617) and OK (n=264)	MN original labels	OK corrected labels	**0.83**	0.57	0.06	0.491	Yes	Yes	II
9	Image	ResNet-18 - Cross-population MN (n=617) and OK (n=264)	OK corrected labels	MN corrected labels	0.54	0.85	0.24	0.556	Yes	Yes	II
10	Image	ResNet-18 - Cross-population MN (n=617) and OK (n=264)	MN corrected labels	OK corrected labels	0.72	0.82	0.32	0.576	Yes	Yes	II
11	Image	ResNet-18 - Cross-population MN (n=617) and OK (n=264)	OK original labels	MN corrected labels	0.65	0.55	0.307	0.480	Yes	Yes	II
12	Image	ResNet-18 - Cross-population MN (n=617) and OK (n=264)	OK corrected labels	MN corrected labels	0.54	0.85	0.24	0.513	Yes	Yes	II
13	Image	ResNet-18 - Combined MN and OK data (n=881)	Pooled original labels	Pooled original labels	0.72	0.79	0.33	0.637	Yes	No	III
14	Image	ResNet-18 - Combined MN and OK data (n=608)	Confident pooled original labels	Confident pooled original labels	0.8	**0.86**	0.55	**0.748**	Yes	No	III

Principal Component Analysis

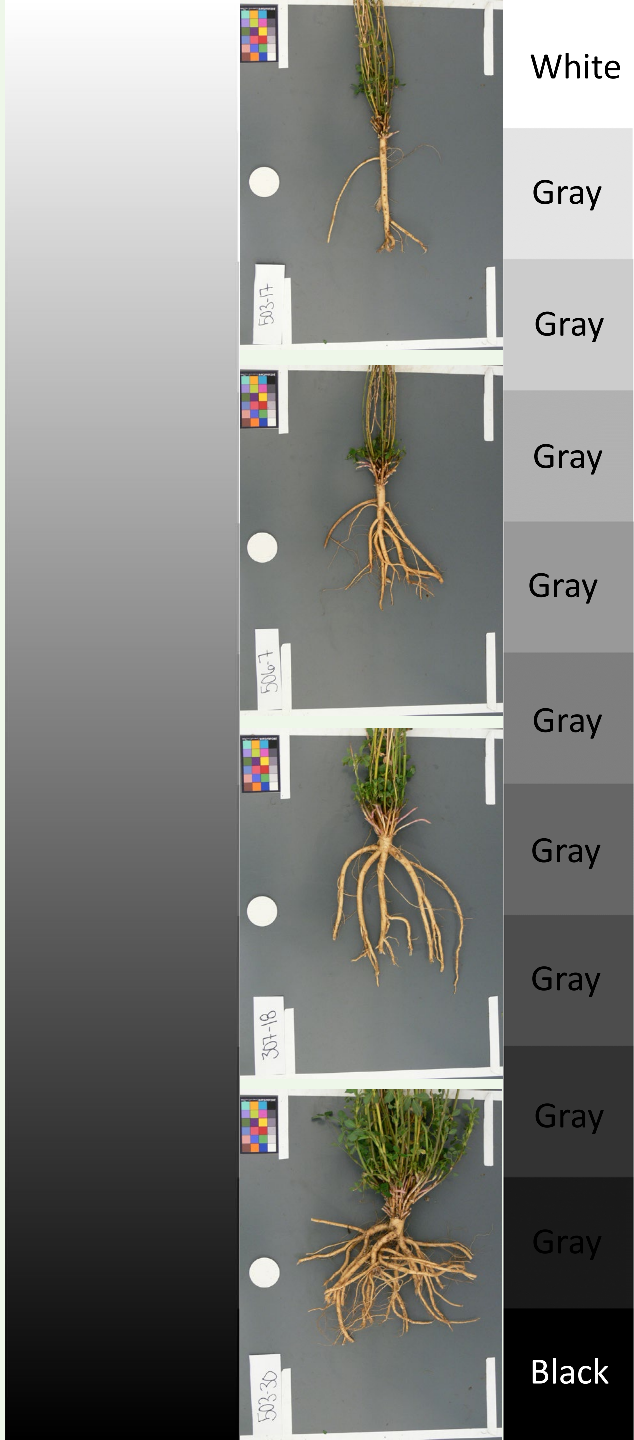
- Used to visualize the label correction process…





Conclusions

- RSA images as direct inputs into Deep Neural Networks:
  - Suitable replacement for traditional methods:
    - Level I (manual measurements).
    - Level II (features as input).
  - Less error-prone than Level I and II.
    - Less human input = less human bias.
- Confident Learning and Reactive Learning:
  - Low-cost and time-efficient.
  - Improve performance and may reduce overparameterization.
  - Improved ResNet-18 prediction accuracies ~11-13%.
- Model refinements are needed to address shortfalls in prediction accuracy of intermediate RSA class.
  - Endmember RSAs of the distribution are easier to classify than the middle…
    - Continuous traits can be difficult to classify…
    - Human labelers also have difficulty with intermediate RSA labels…



Taproot | Hybrid Taproot | Hybrid Branched | Branched

Fibrous F-1

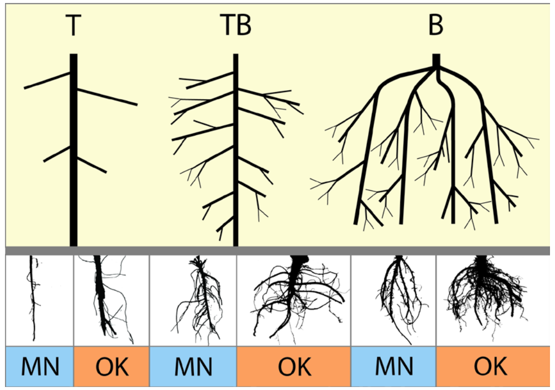
Fibrous F-2

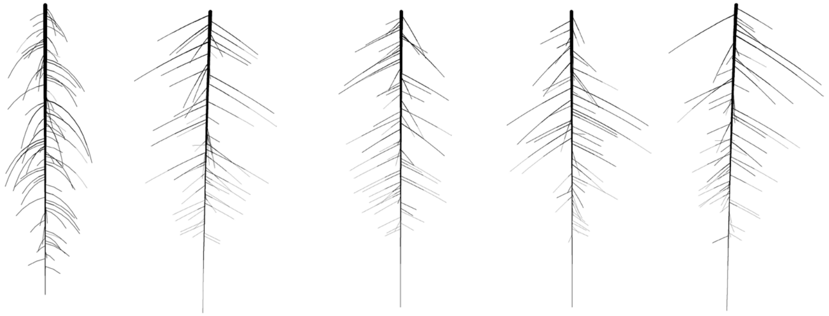
Fibrous F-3

Future Directions…

- Deploy ResNet model in mobile device RSA application (a mobile App)
    - Put AI-driven RSA image analysis into farmers and stakeholder hands.
- Test methods to create synthetic data used for model training.
- Couple AI RSA analyses results with downstream traits such as yield, winter survival, etc.
- Investigate the possible relationship between RSA and yield via root/shoot allometry.
- Include digestibility, maturity, and biomass estimations based on image analyses into the mobile App functions.

T-1	T-2	T-3
HT-1	HT-2	HT-3
HB-1	HB-2	HB-3
B-1	B-2	B-3





Simulated roots – Lobet et al. (2016)

Acknowledgements

- Thanks to everyone that supported and contributed to our efforts on this work! We appreciate you, and all your help!
  - Special thanks to Zhanyou Xu, Jo Heuschele, and Deb Samac!
- Funding:
  - US Department of Agriculture (Agricultural Research Service project 5062-12210-004-D and 2020-67021-32460)
  - Center for Bioenergy Innovation (CBI), U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number ERKP886
- The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The USDA is an equal opportunity provider and employer. All experiments complied with the current laws of the United States, the country in which they were performed.

References Cited

- Bucciarelli, B., et al., *Phenotyping seedlings for selection of root system architecture in alfalfa (Medicago sativa L.).* Plant Methods, 2021. **17**(1): p. 125.
- Lobet, G., Koevoets, I. T., Tocquin, P., Pagès, L., & Périlleux, C. (2016). Library of simulated root images [Data set]. Zenodo. https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.5281%2Fzenodo.61739&data=05%7C02%7CBrandon.Weihs%40usda.gov%7C83ade43d28f541d27e6b08dc0311a0cd%7Ced5b36e701ee4ebc867ee03cfa0d4697%7C1%7C0%7C638388621450521231%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=BaiGOIOE1TXMcunBGG%2BjfU5Z7OqhDKfWPC23yJdeqzY%3D&reserved=0
- Mattupalli, C., et al., *Digital Imaging to Evaluate Root System Architectural Changes Associated with Soil Biotic Factors.* Phytobiomes Journal, 2019. **3**(2): p. 102-111.
- Northcutt, C., L. Jiang, and I. Chuang, *Confident learning: Estimating uncertainty in dataset labels.* Journal of Artificial Intelligence Research, 2021. **70**: p. 1373-1411.
- Weihs, B.J., et al., *The State of the Art in Root System Architecture Image Analysis Using Artificial Intelligence: A Review.* Plant Phenomics, 2024.
- Xu, Z., et al., *Objective Phenotyping of Root System Architecture Using Image Augmentation and Machine Learning in Alfalfa (Medicago sativa L.).* Plant Phenomics, 2022.


Thanks for listening! ☺


Questions
and/or
Comments?