



Harmonized Multispectral Image Reconstruction using Generative Adversarial Networks with Multimodal Data Fusion for Alfalfa Monitoring

Tong Yu¹, Jiang Chen¹, Jerome H. Cherney² and Zhou Zhang^{1, *}

¹ Department of Biological Systems Engineering, University of Wisconsin-Madison,

² Soil and Crop Sciences Section, School of Integrative Plant Science, Cornell University



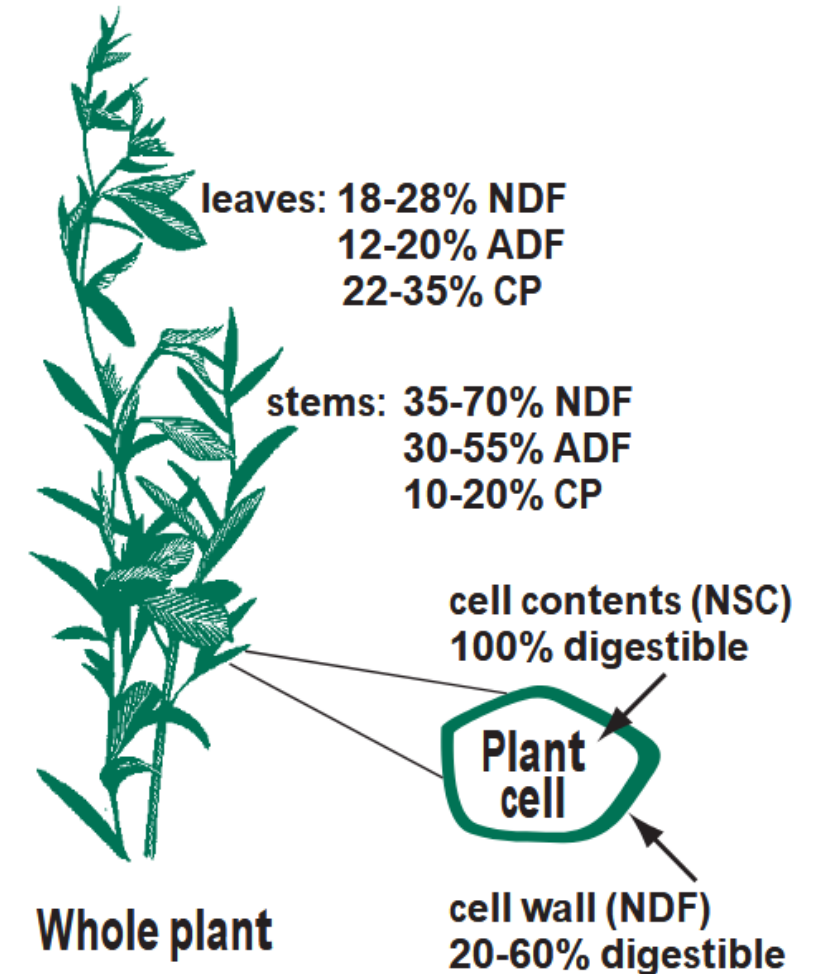
Background



Background

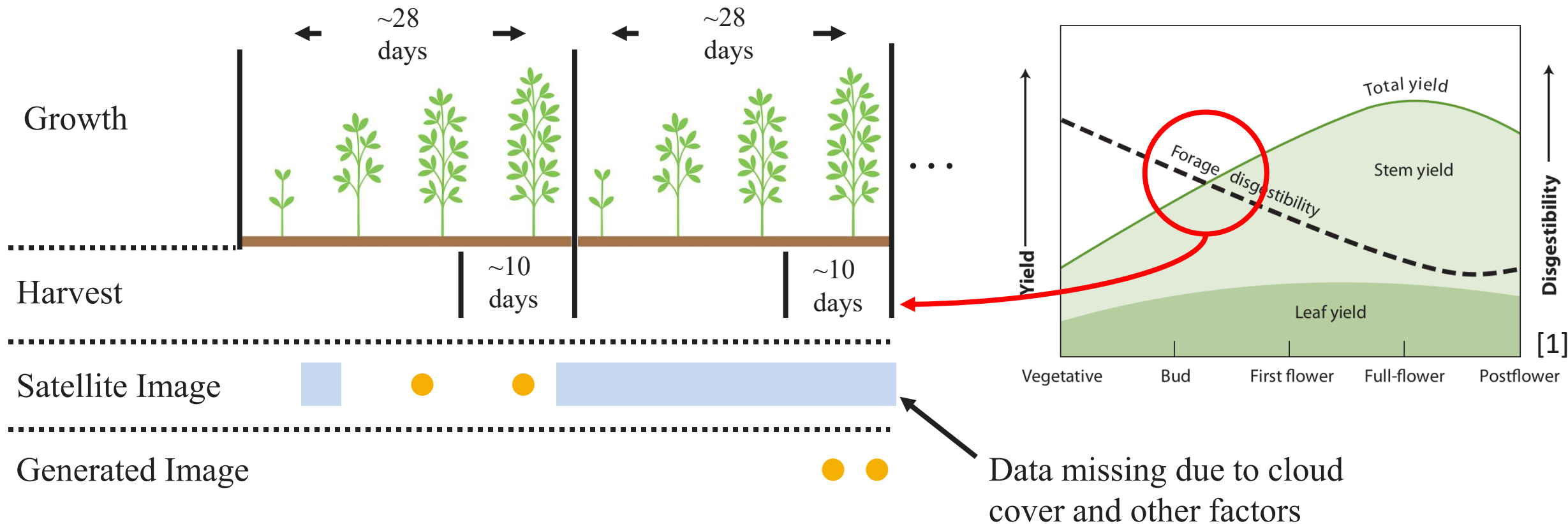
Alfalfa-"Queen of Forages"

- A perennial leguminous crop widely cultivated as animal feed.
- Impressive yield, rich nutritional profile, and significant protein content
- One of the largest and most valuable forage crops

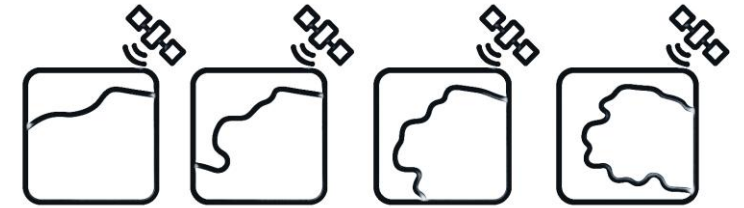
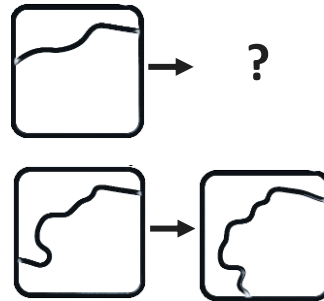


Adapted from: Putnam, Dan, 2000. Producing high quality alfalfa: Factors that influence alfalfa forage quality. Proc. CA Plant and Soil Conference, Jan. 19-20. Stockton, CA.

Background



Background



UAV-based methods

- Highly affected by wind and precipitation
- Require certain hardware support and manual labor

Spatiotemporal Fusion methods

- Require several image pairs to generate high-resolution and high-frequency imagery

Time-series reconstruction methods

- Suitable for surfaces with slow changes
- Require continuous data for model construction.

Not suitable for frequently harvested crops such as alfalfa

Background

Spectral Reconstruction Methods (e.g. MST++, HSGAN)

- Single-Temporal Dependency
- Spectral Super-Resolution



Gaps

Limited temporal generalization

One-to-one mapping

Neglect of geographic and environmental context



Required model capabilities

Beyond the timestamps of observations

Multisource data integration

Geographic and environmental factors incorporation



Methods and Material





Data Source-Image Reconstruction

- High data quality and is widely used in related studies.
- Long-term and stable observations.
- Rich spectral information, especially in bands relevant to vegetation growth.
- Easily accessible.
- Suitable resolution.

The HLS 30m Sentinel product (S30)

Output



The HLS 30m Landsat product (L30)

- Specifically harmonized to match the Sentinel-2

Sentinel-1 Ground Range Detected (GRD)

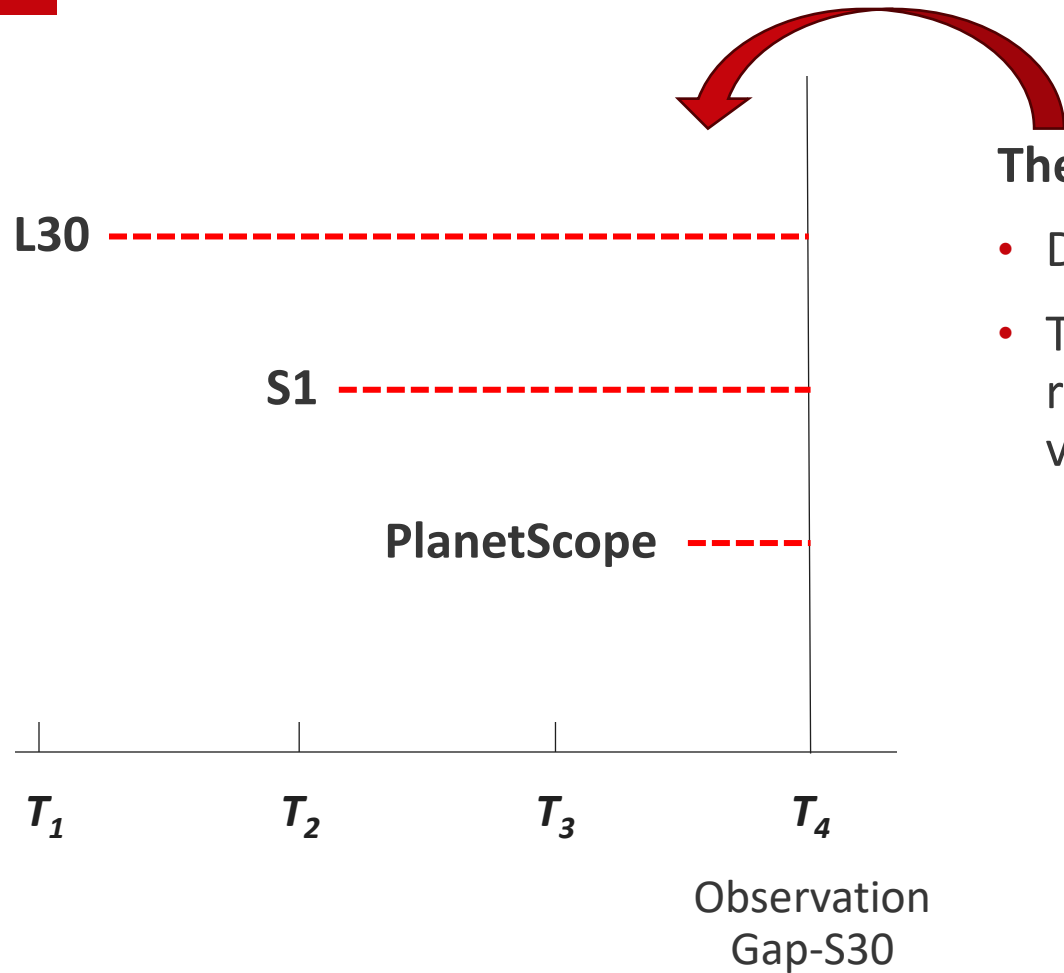
- All-weather, day-night monitoring

PlanetScope

- Near-daily global coverage
- Eight spectral bands that are largely interoperable with Sentinel-2 bands

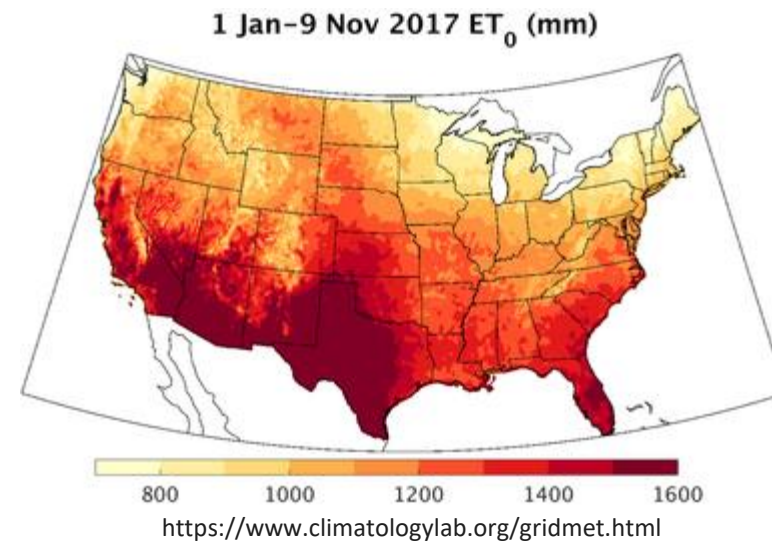
Input

Data Source-Image Reconstruction



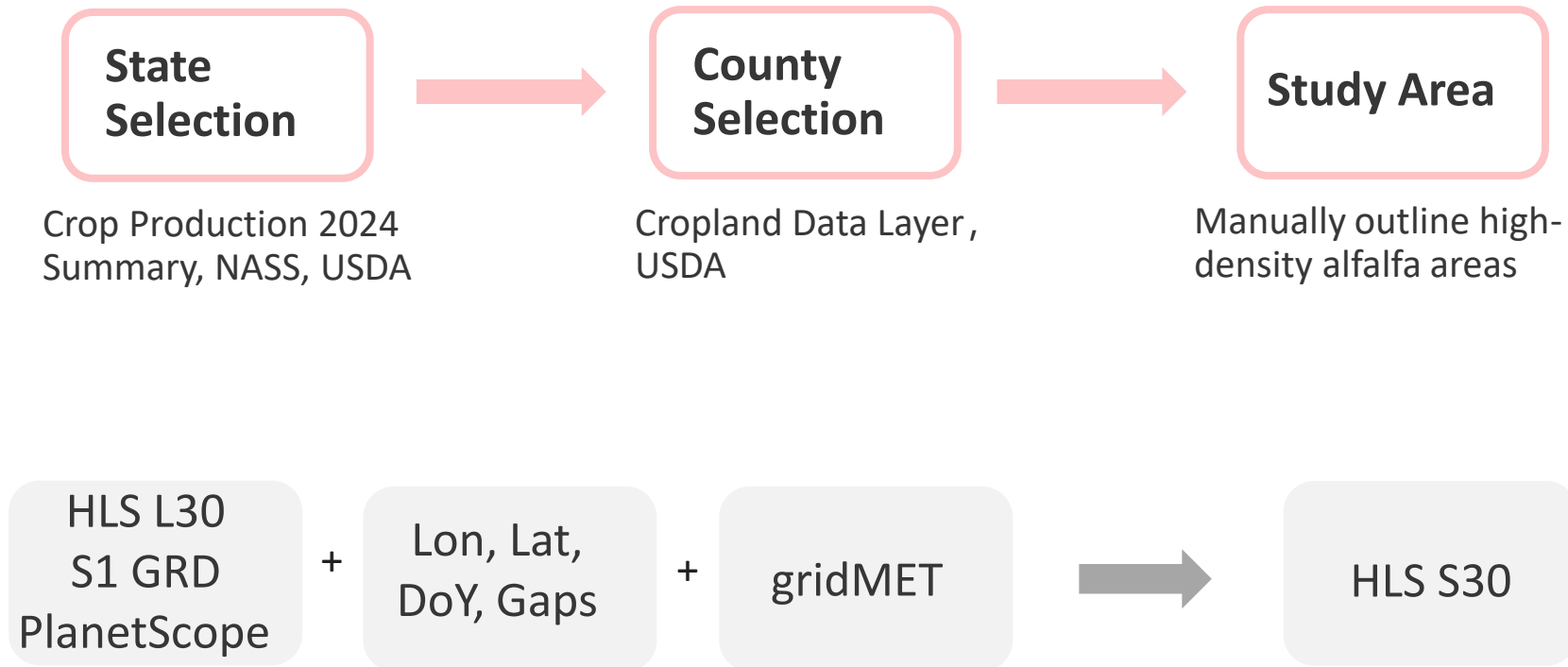
The Gridded Surface Meteorological (gridMET) dataset

- Daily surface meteorological dataset for the United States
- Temperature, precipitation, humidity, wind speed, reference evapotranspiration and other meteorological variables





Dataset-Image Reconstruction



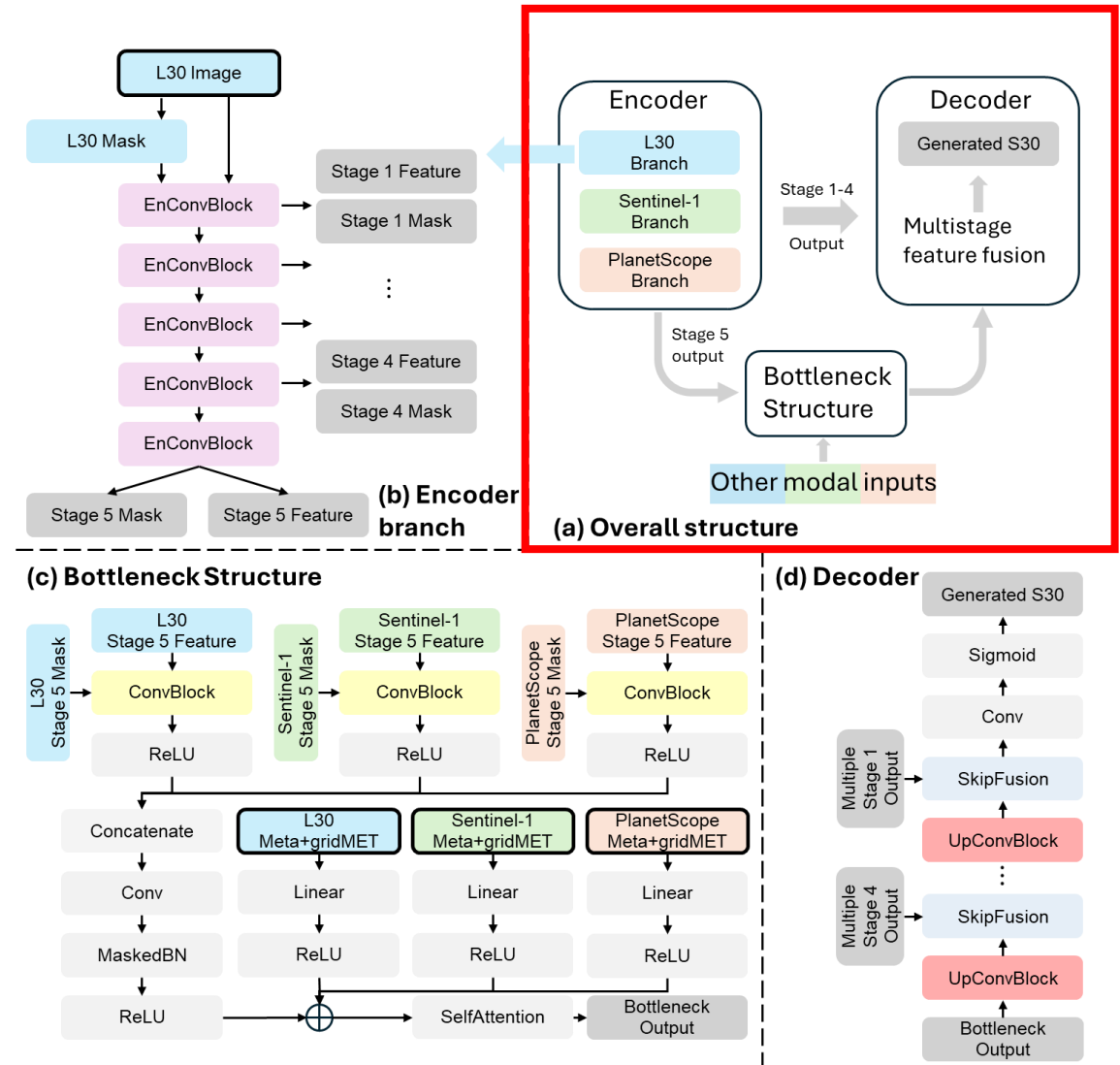
4138 training patches, **512** validation patches, and **532** testing patches



Model Structure-AMGAN

Generator:

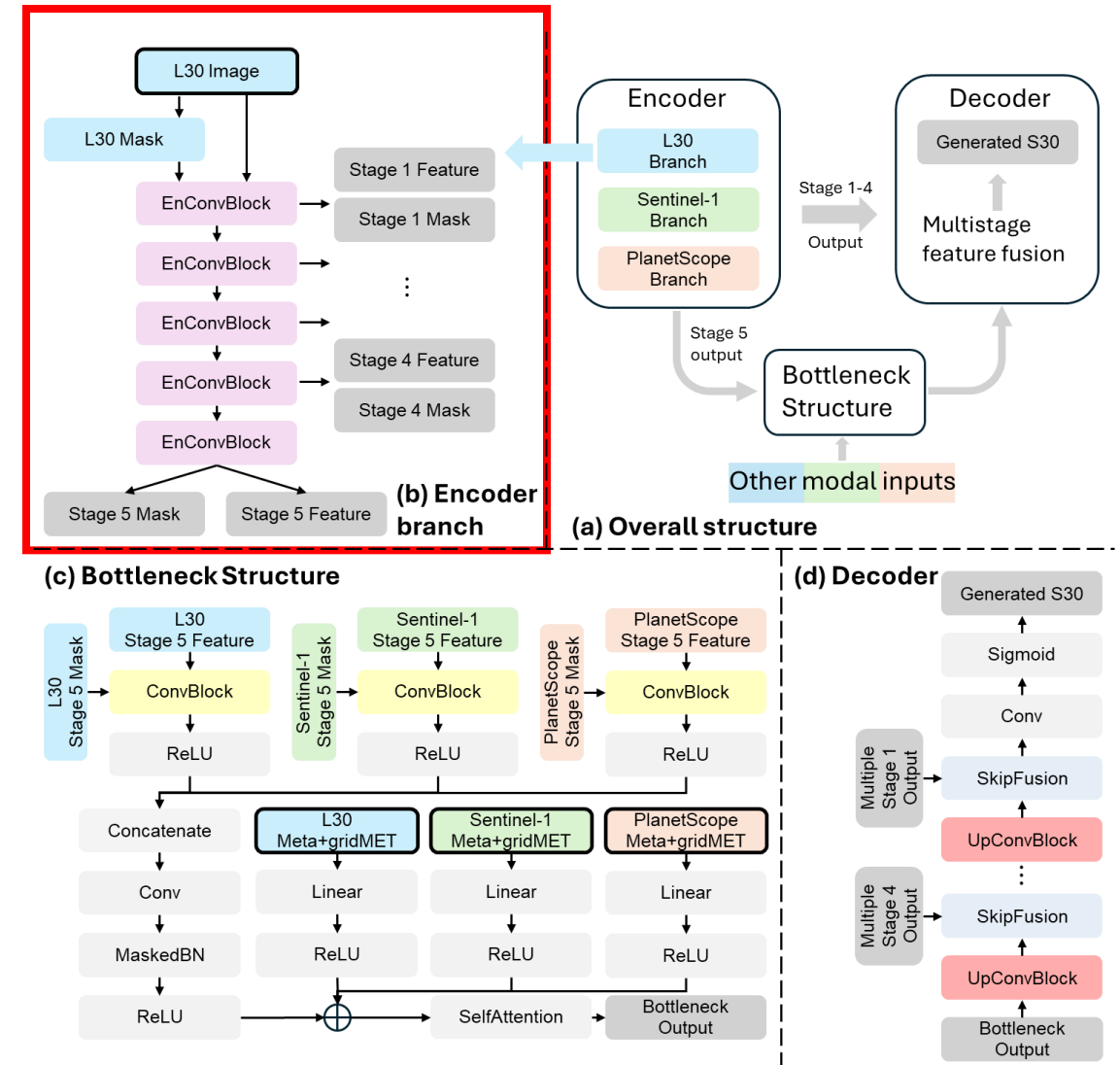
- Encoder–decoder structure



Model Structure-AMGAN

Generator:

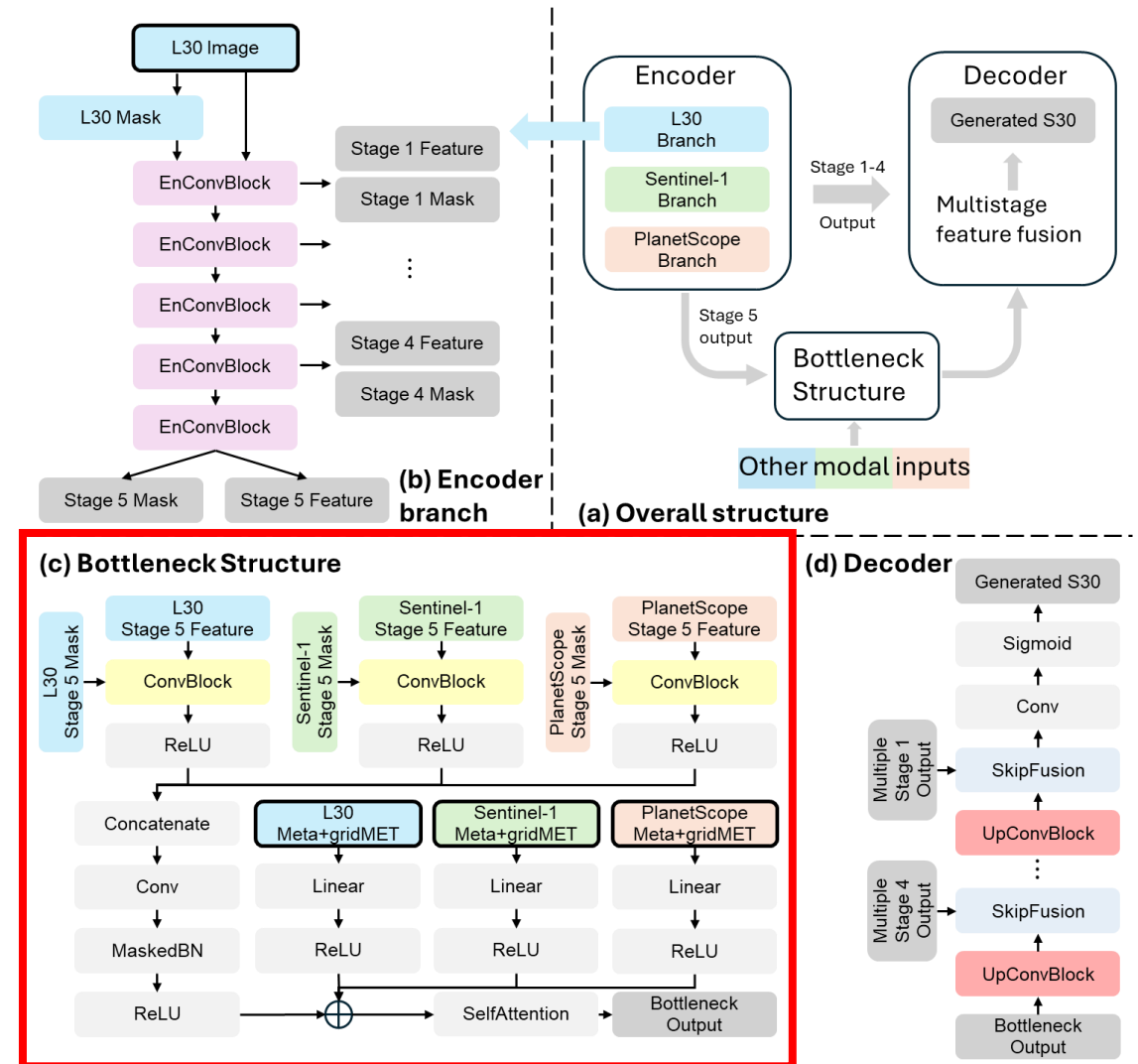
- Encoder–decoder structure
- Each image input has a separate encoder branch



Model Structure-AMGAN

Generator:

- Encoder–decoder structure
- Each image input has a separate encoder branch
- Other modal inputs are fed into the model through the bottleneck section

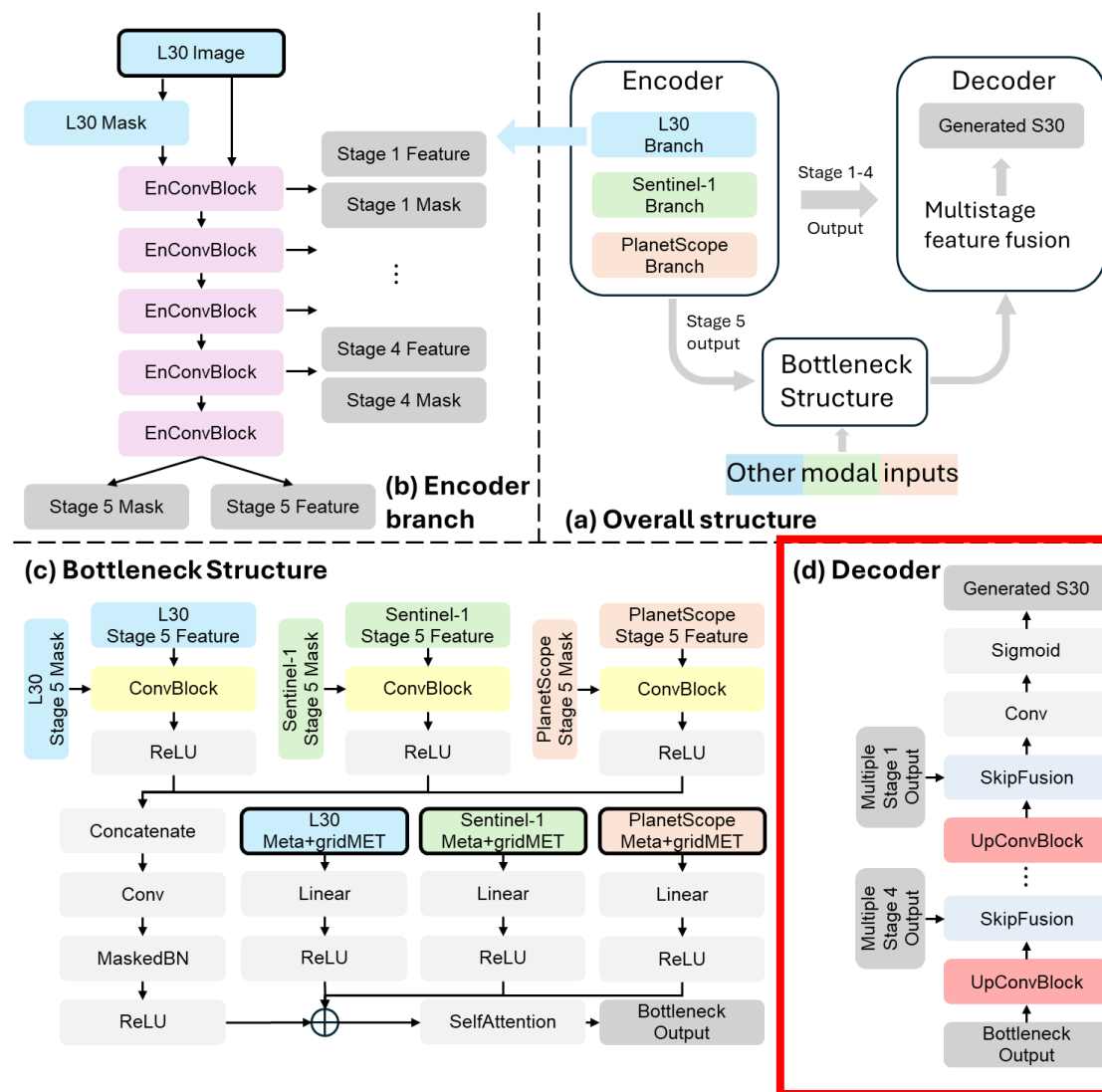




Model Structure-AMGAN

Generator:

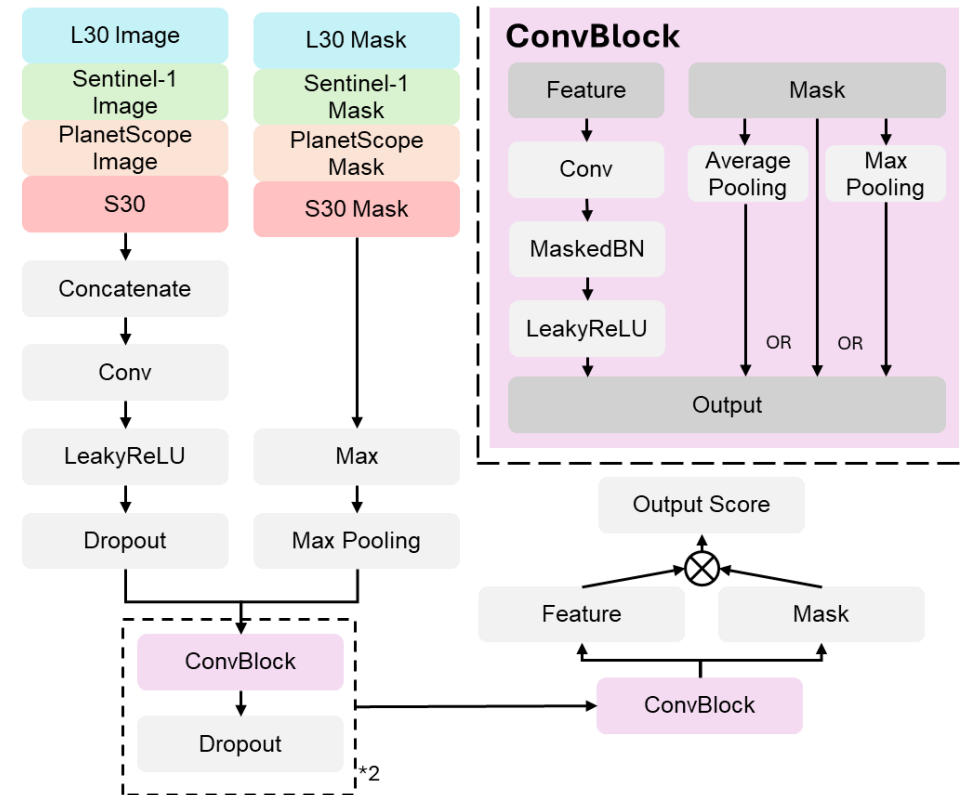
- Encoder–decoder structure
- Each image input has a separate encoder branch
- Other modal inputs are fed into the model through the bottleneck section
- Adaptive weighted skip connection module



Model Structure-AMGAN

Discriminator:

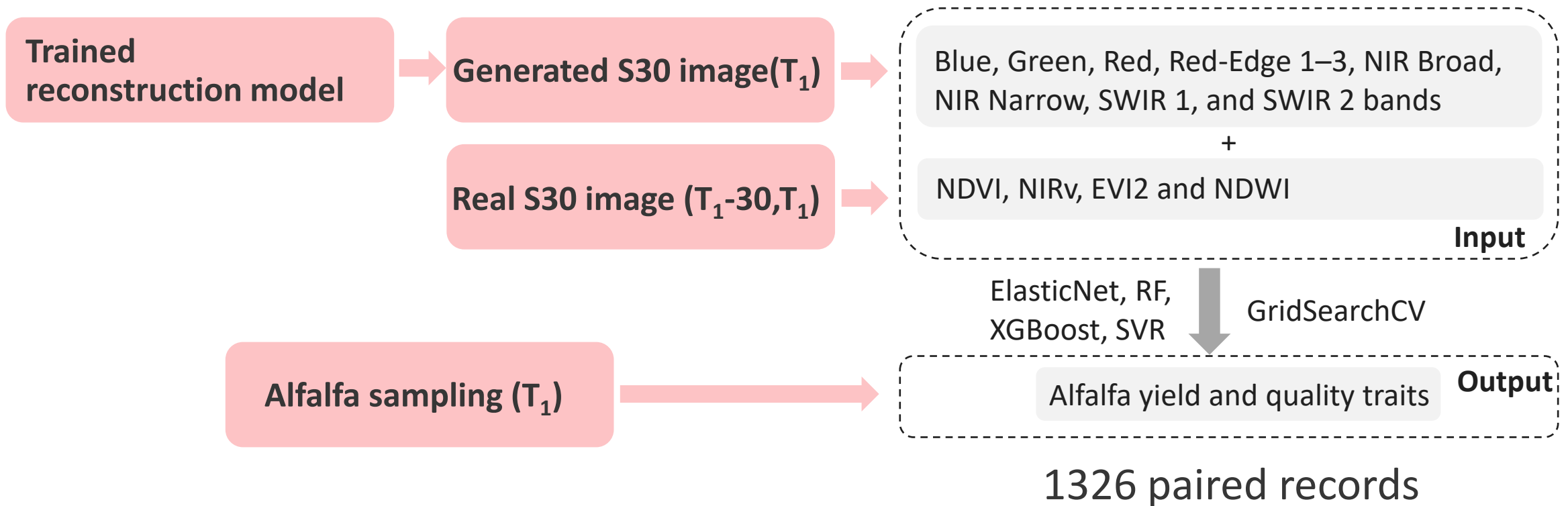
- A three-scale conditional design
- Not only input the real/generated S30 images, but the input to generate the S30 images





Dataset-Yield and quality prediction

To validate the effectiveness of the reconstructed data, multiple alfalfa sampling fields in Wisconsin and New York were selected as study fields from 2022 to 2024.





Results & Discussion





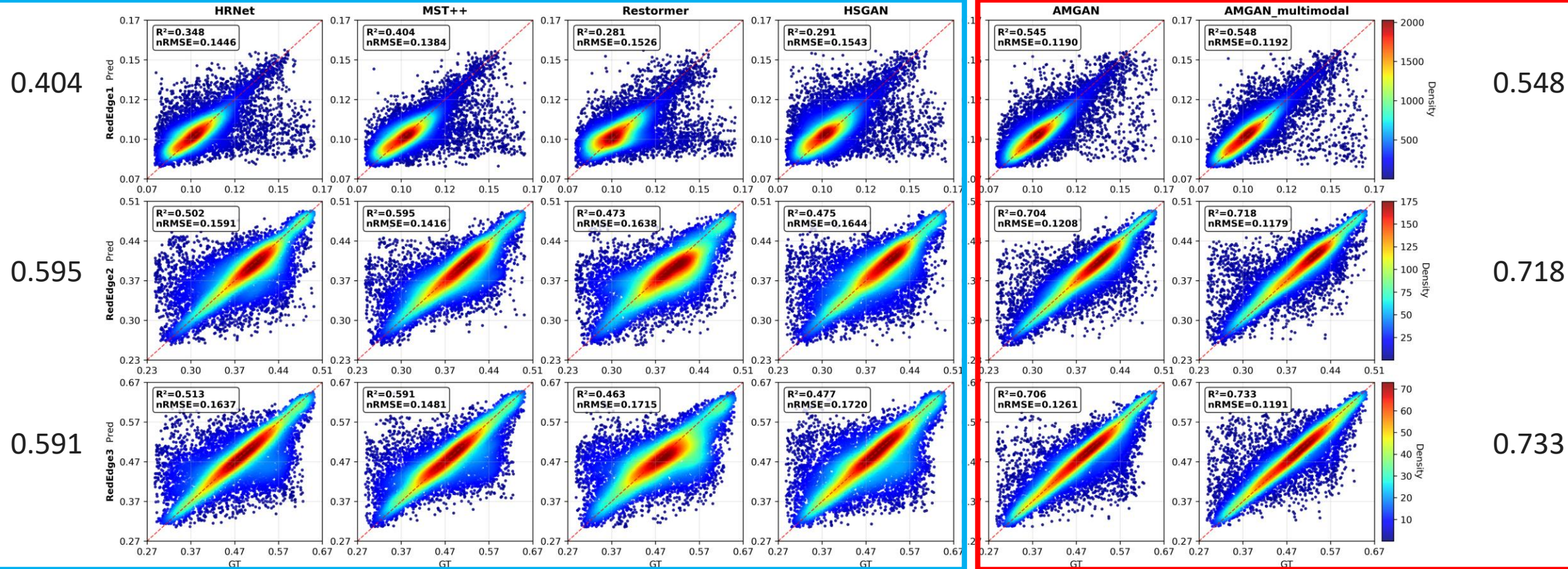
Reconstruction performance

- The addition of multimodal inputs reduced RMSE by 14.3% and 8.3% and SAM by 14.8% and 13.5%, while increasing PSNR by 3.2% and 2.2% and SSIM by 0.2% and 0.3% for the generator-only model and AMGAN, respectively.
- Adversarial training also yielded a significant overall improvement, even greater than that achieved by multimodal fusion alone
- AMGAN consistently outperformed all comparison models across all evaluation metrics

Table 3. Reconstruction performance comparison

Model	Inputs	RMSE ↓	PSNR ↑	SSIM ↑	SAM ↓
HRNet ^[2]	Image	0.030±0.024	33.84±8.03	0.957±0.044	3.26±3.21
Restormer ^[3]	Image	0.035±0.027	31.41±6.28	0.950±0.046	3.92±3.24
MST++ ^[4]	Image	0.028±0.021	33.66±6.75	0.958±0.042	3.11±2.90
HSGAN ^[5]	Image	0.031±0.024	32.63±6.96	0.952±0.044	3.50±3.15
Generator	Image	0.028±0.021	33.60±6.71	0.959±0.041	3.11±2.85
	Image + Meta + gridMET	0.024±0.019	34.66±6.46	0.961±0.040	2.65±2.57
AMGAN	Image	0.024±0.019	35.14±6.97	0.963±0.038	2.59±2.38
	Image + Meta + gridMET	0.022±0.019	35.90±6.87	0.966±0.038	2.24±2.54

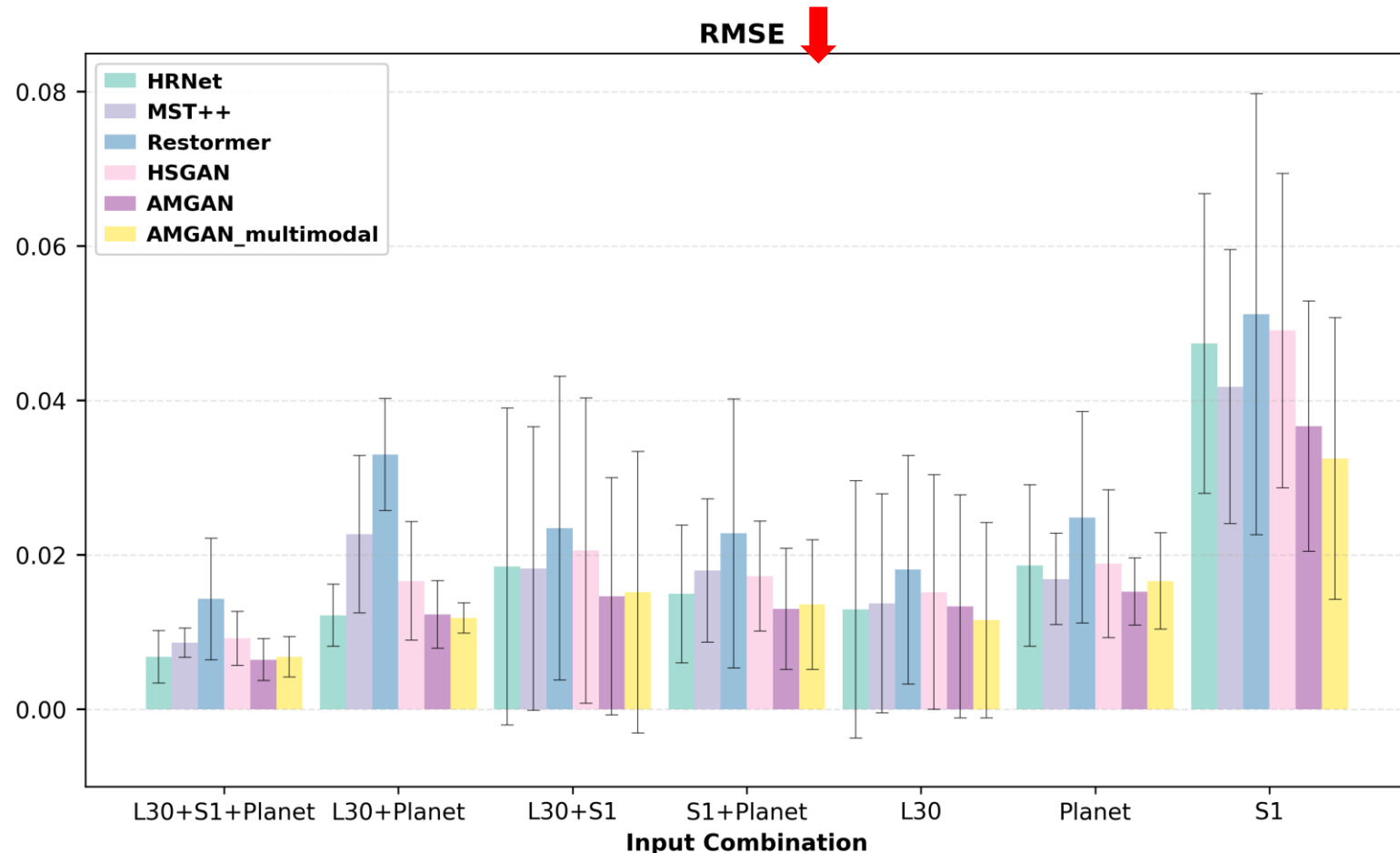
Reconstruction performance



Scatter density plots of reconstructed vs. reference reflectance for three red-edge bands using six models.

Reconstruction performance-Input combination

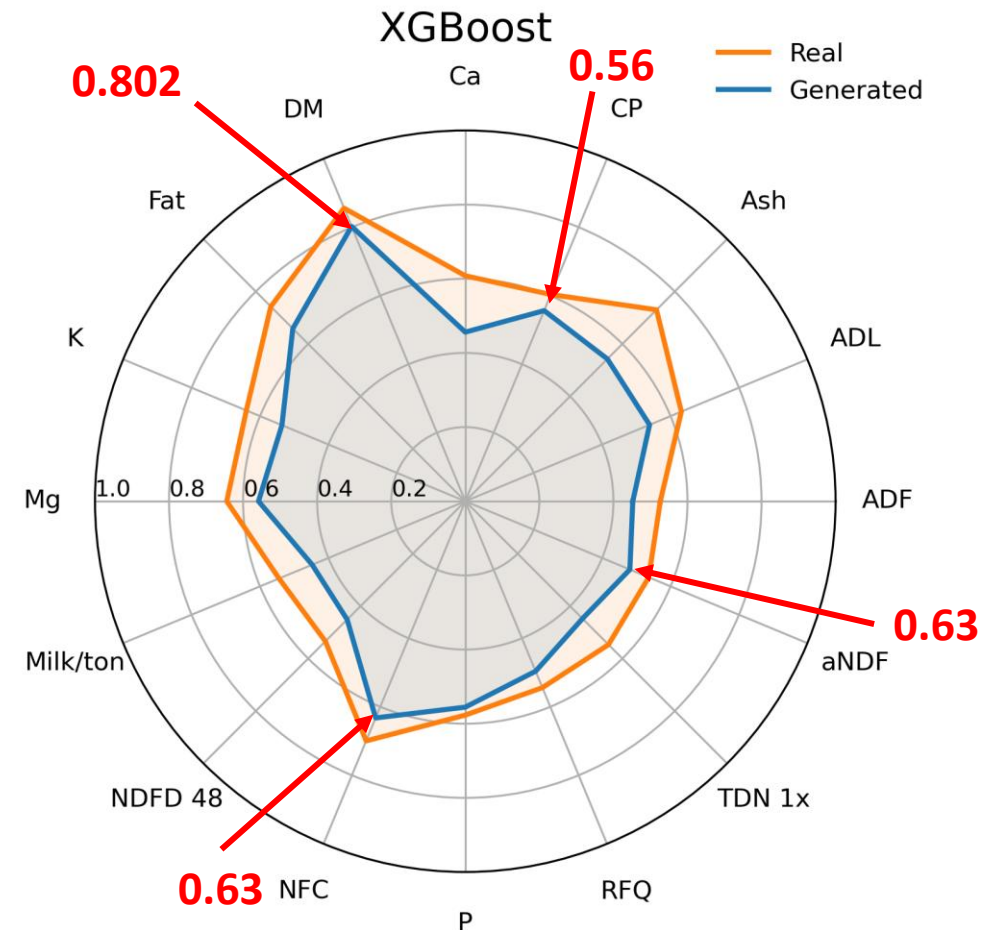
- L30+ S1+Planet achieved the best performance
- S1 alone produced the weakest results
- AMGAN outperformed the baseline models across all input combinations
- The contribution of each source is not strictly additive but depends on complementarity, redundancy, and noise



Comparison of reconstruction performance across models trained with different input combinations.

Yield and quality prediction performance

- The performance gap between real and reconstructed imagery remained small.
- DM is closely associated with canopy structure and biomass accumulation, both of which are well captured by reconstructed canopy spectral features.
- Most nutritional quality traits are not directly or strongly expressed in canopy spectra (broad bands -> narrow absorption features)



Comparison of yield and quality prediction performance (R^2) in repeated 5-fold CV experiment



Take-home message

- **Multimodal data substantially enhances spectral reconstruction**, with especially strong benefits for spectral prediction accuracy.
- **The proposed AMGAN model, with its flexible input design, delivers robust and reliable spectral estimates**, providing strong data support for alfalfa monitoring and decision-making during the harvest season.



Future Work

- **Spatiotemporal scalability**
Extend the framework to larger regions and longer time horizons.
- **Cross-crop applicability**
Evaluate the model's transferability to other crop types.
- **Integration of growth dynamics**
Incorporate process-based or phenology-driven constraints to improve temporal prediction and biological consistency.



Acknowledgement

- Yijia Xu
- Xiaoyu Wang
- Sadegh Ranjbar
- Jing Zhou
- Arlington Agricultural Research Station
- USDA, NIFA, Alfalfa Seed and Alfalfa Forage System Program(Award No. 2021-70005-35694)
- USDA, NIFA, Agriculture and Food Research Initiative Foundational Program (Award No. 2022-67021-36468)
- USDA Hatch Accession Number 7005141



Reference

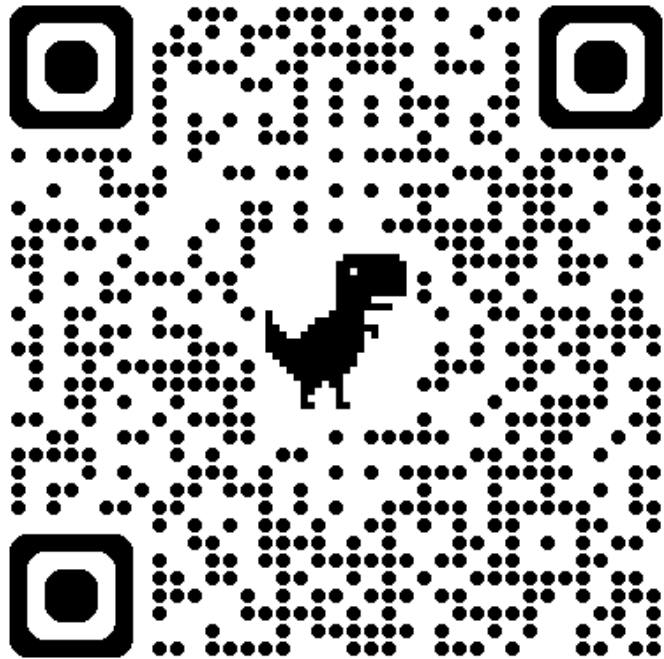
- [1] Mueller, S.C., Teuber, L.R., 2007. Alfalfa Growth and Development. IN (C. G. Summers and D. H. Putnam, eds.), Irrigated alfalfa management for Mediterranean and Desert zones. Chapter 3. University of California Agriculture and Natural Resources
- [2] Zhao, Y., Po, L.M., Yan, Q., Liu, W. and Lin, T., 2020. Hierarchical regression network for spectral reconstruction from RGB images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 422-423).
- [3] Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S. and Yang, M.H., 2022. Restormer: Efficient transformer for high-resolution image restoration. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 5728-5739).
- [4] Cai, Y., Lin, J., Lin, Z., Wang, H., Zhang, Y., Pfister, H., Timofte, R. and Van Gool, L., 2022. Mst++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 745-755).
- [5] Zhao, Y., Po, L.M., Lin, T., Yan, Q., Liu, W. and Xian, P., 2023. Hsgan: Hyperspectral reconstruction from rgb images with generative adversarial network. IEEE Transactions on Neural Networks and Learning Systems.



Thank you for your attention!

Open on the job market! Available to start in Fall 2026

<https://uwty.github.io/>



Personal website

Research Areas:

- AI + Remote Sensing + ?

Seeking Opportunities in:

- Postdoctoral positions
- Research scientist roles
- Applied AI / Geospatial engineering positions

