Mapping forest aboveground biomass in large areas

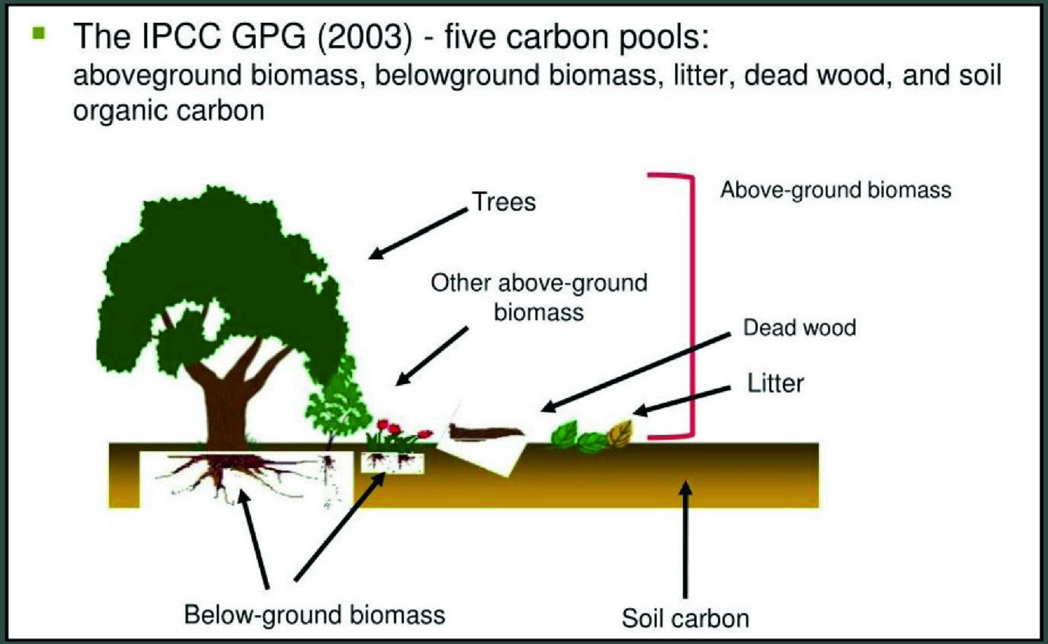
Global forest cover, 2022

We cannot tackle the nature and climate crises or provide sustainable jobs without forests. And yet we are destroying them.

----WWF

Source: *Esri / Sentinel-2 Land Cover Explorer*

What is forest aboveground biomass and why do we care?



Forest **a**bove**g**round **b**iomass (**AGB**):

[Weights of living vegetation above the soil]
(stem, branches, foliage, etc.)

Above**g**round **b**iomass **d**esnsity (**AGBD**):

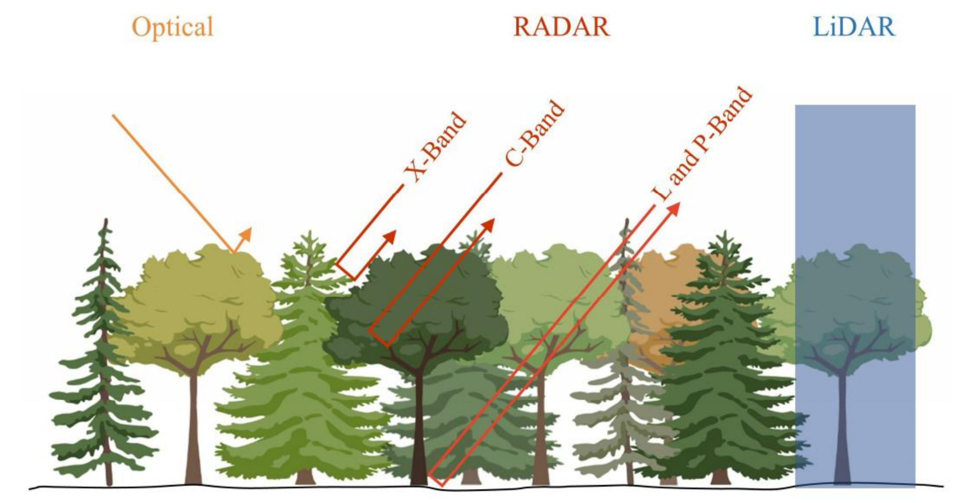
Mg / ha (1 Mg = 1000 kg)

AGB

- Indicator of forest status → Forest monitoring and management
- Calculation of carbon storage → Climate change mitigation

Source: *Ashton et al., 2012; Roy et al., 2020.*

How do we estimate forest aboveground biomass?



Traditional field survey

Harvest and weighing, allometric equations, etc. (**plot** scale)

Precise AGB measurement

↓

AGB estimation in large areas

↑

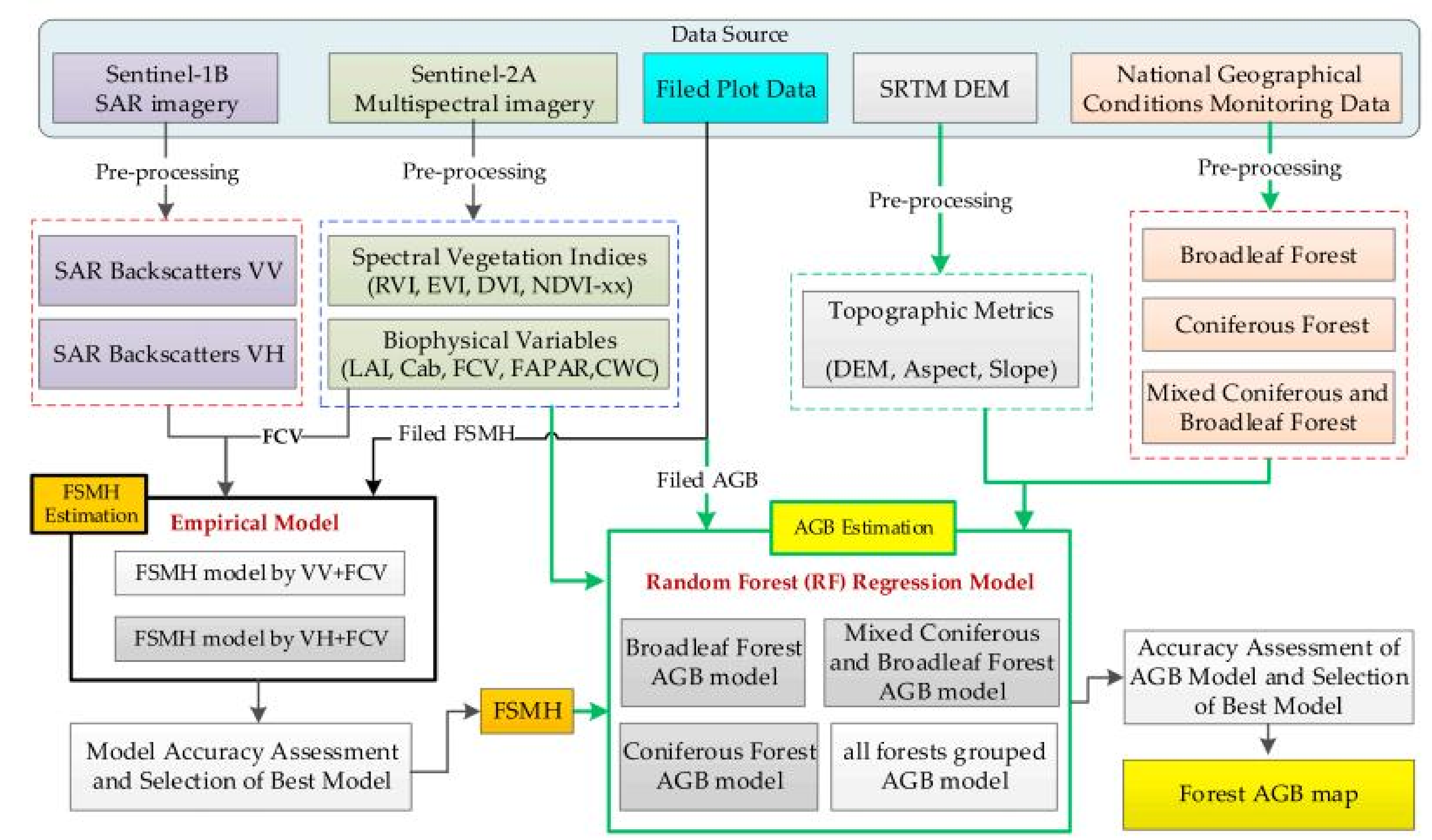
Vegetation index, backscatter coefficient, LiDAR metrics, etc.

Optical imagery, radar imagery, LiDAR (**region** scale)

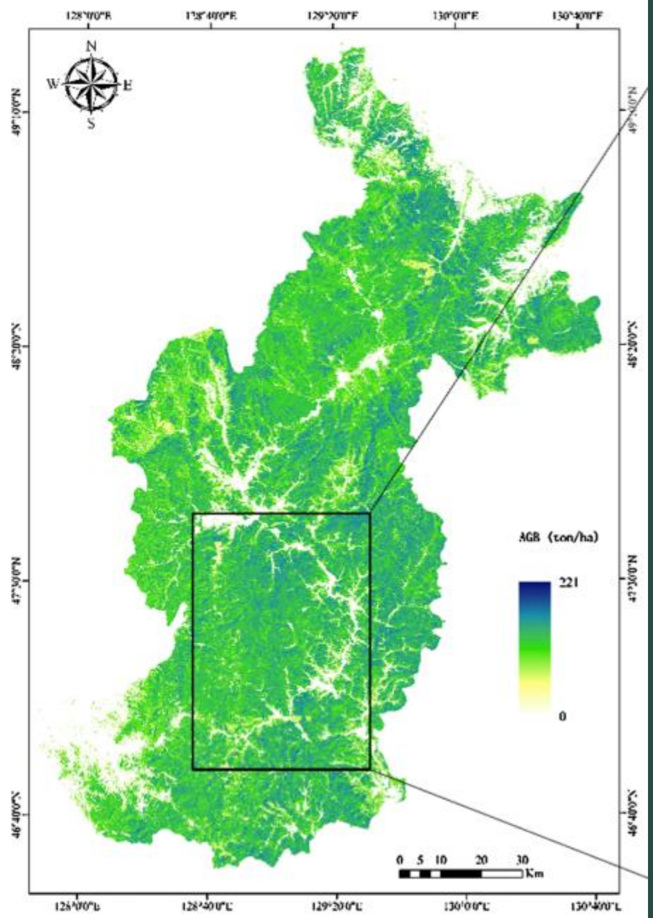
Remote sensing

__Source:__ https://sm.mashable.com/mashable_sea/photo/default/malaysia-deforestation-2023-header_mztm.png; https://s.wsj.net/public/resources/images/BN-RP616_NUMBER_M_20170112124801.jpg; Tian et al., 2023.

Example1: Sentinel–1 & Sentinel–2



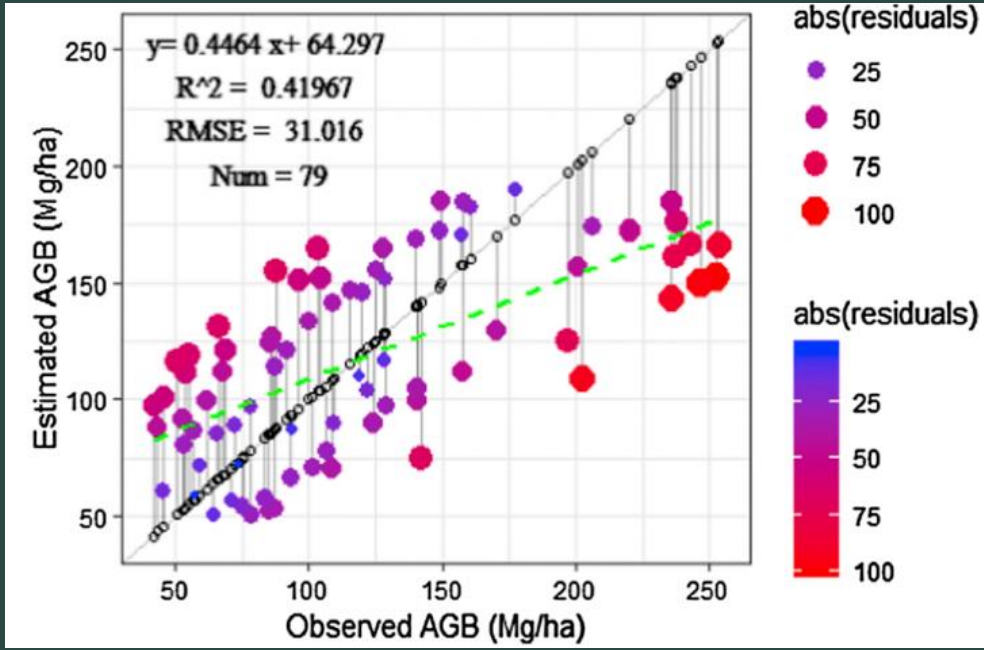
Overall workflow



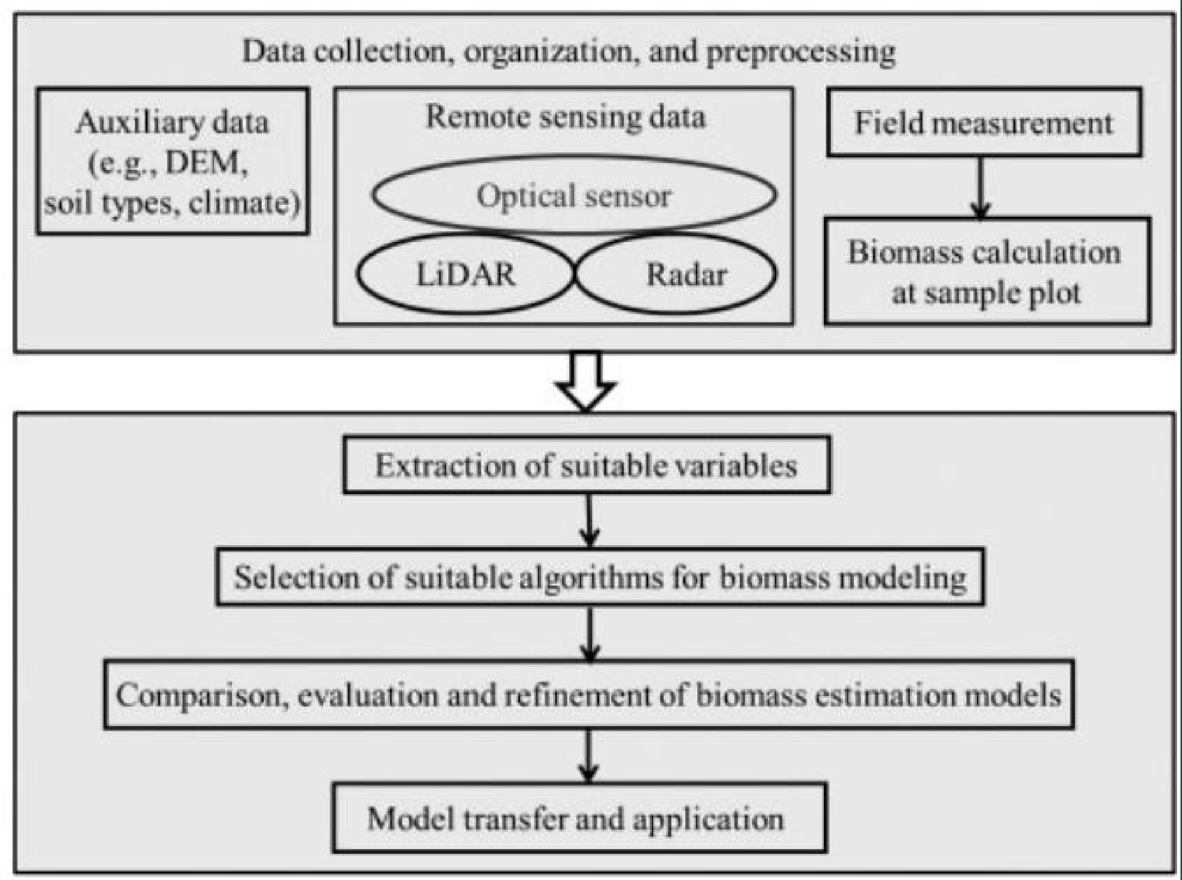
10-m resolution AGB map of Yichun

Source: *Liu et al., 2019.*

Example1: Sentinel–1 & Sentinel–2



Model performance evaluation

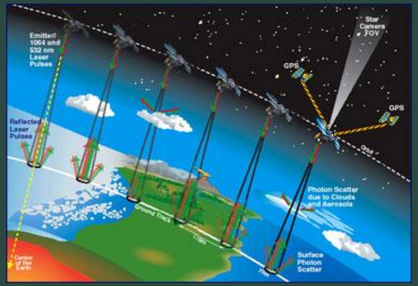


Common structure

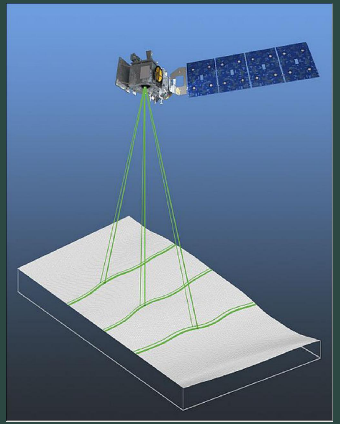
Source: *Lu et al., 2023; Liu et al., 2019.*

Enter the era of LiDAR…

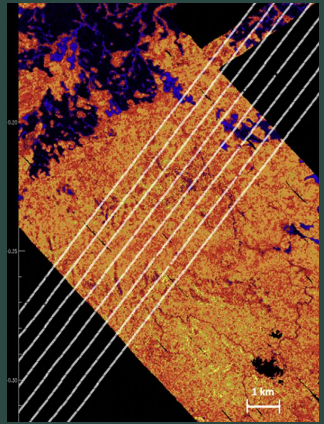
ICESat



ICESat-2



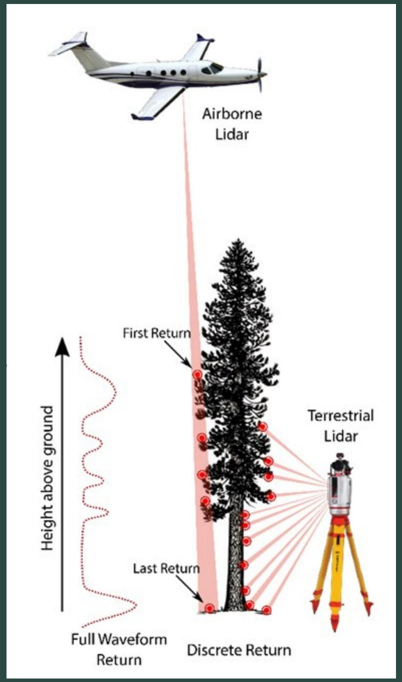
GEDI



LiDAR: Light Detection And Ranging

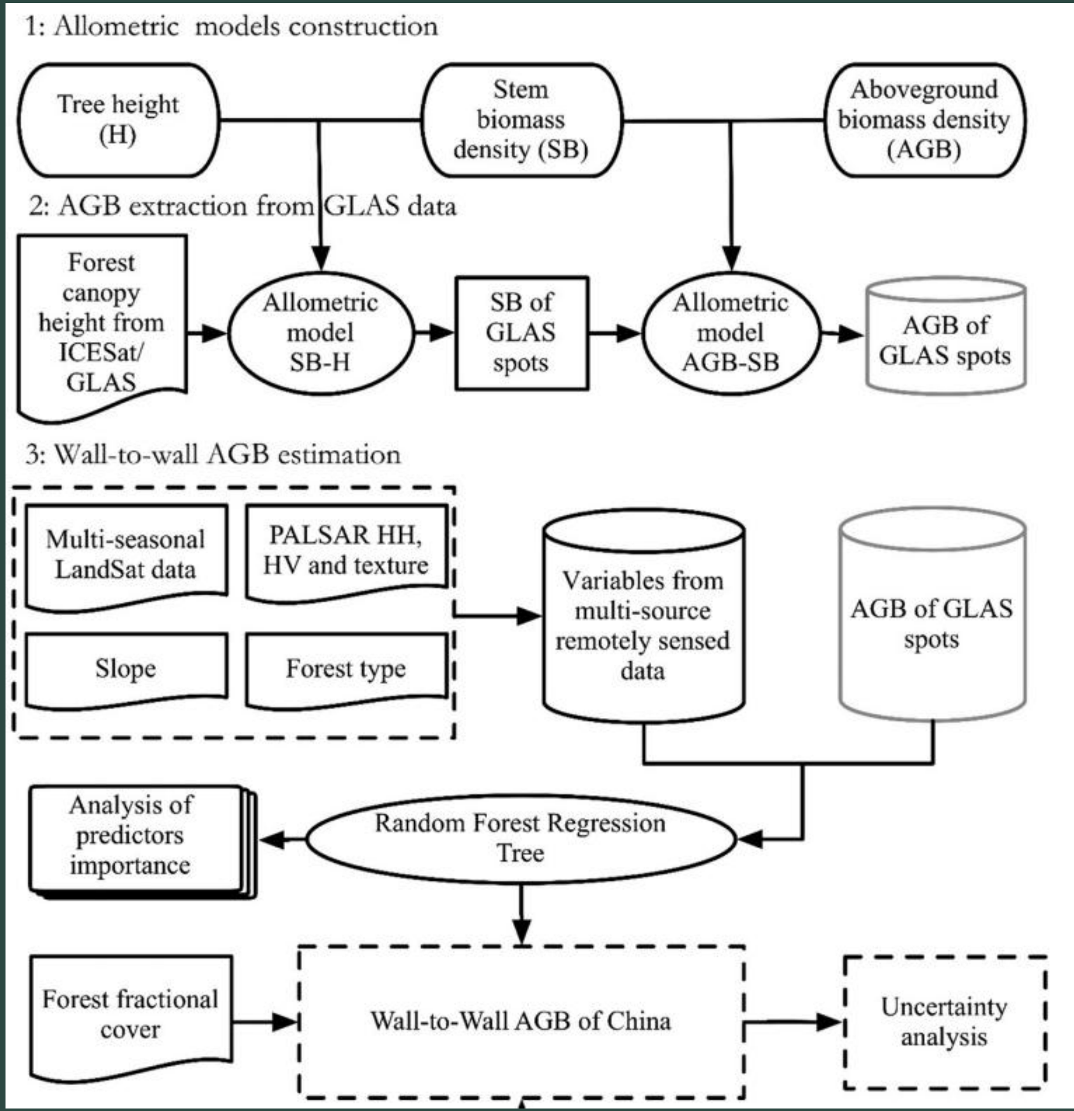
- Platform: terristrial, airborne, spaceborne
- High penetration, detecting vertical structure

	ICESat/GLAS	ICESat-2/ATLAS	GEDI
Host platform	Satellite	Satellite	International Space Station
Operation periode	1.2003-10.2009	9.2018-ongoing	11.2018-ongoing
Technology	Full waveform	Photon counting	Full waveform
Altitude	600km	500km	419km
Coverage area	86° S-86° N	88° S-88° N	51.6° S-51.6° N
Revisit period	183d	91d	-
Ground tracks	1	6(3 groups)	8
Footprint diameter	60-70m	17m	25m
Along-track spacing	170m	0.7m	60m
Cross-track spacing	-	90m(inter) / 3.3km(cross)	600m

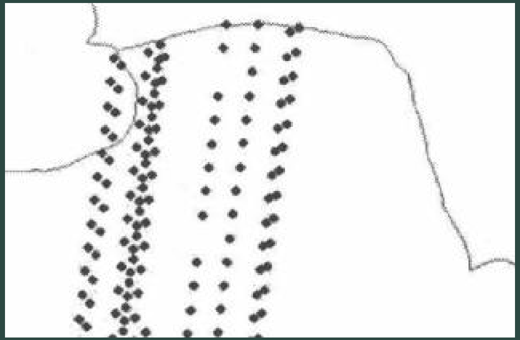


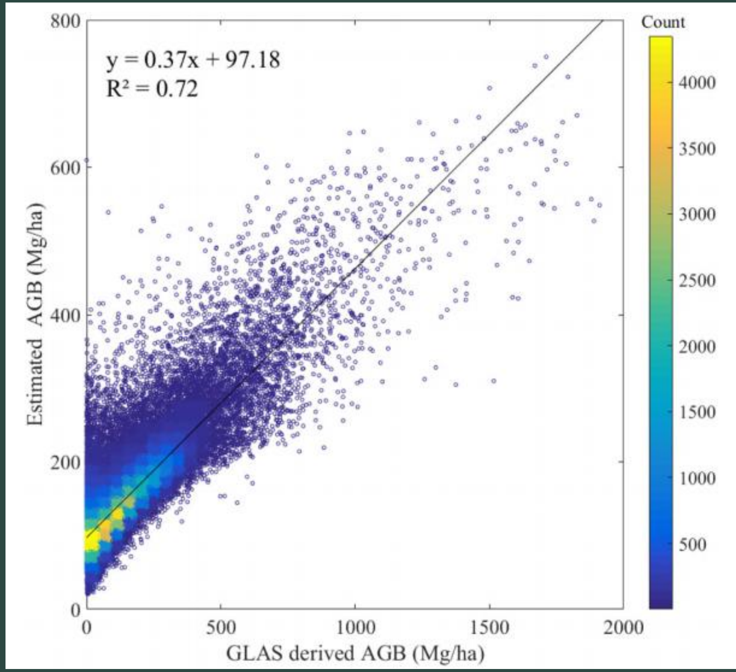
Source: *https://sweri.eri.nau.edu/wp-content/uploads/2019/11/FactSheet_LiDAR_Oct2019_FINAL.pdf, Dubayah et al., 2020; https://attic.gsfc.nasa.gov/glas/about.html; https://www.eoportal.org/satellite-missions/icesat-2#atlas-advanced-topographic-laser-altimeter-system; Wang et al., 2022.*

Example2: ICESat, Landsat, PALSAR



Overall workflow

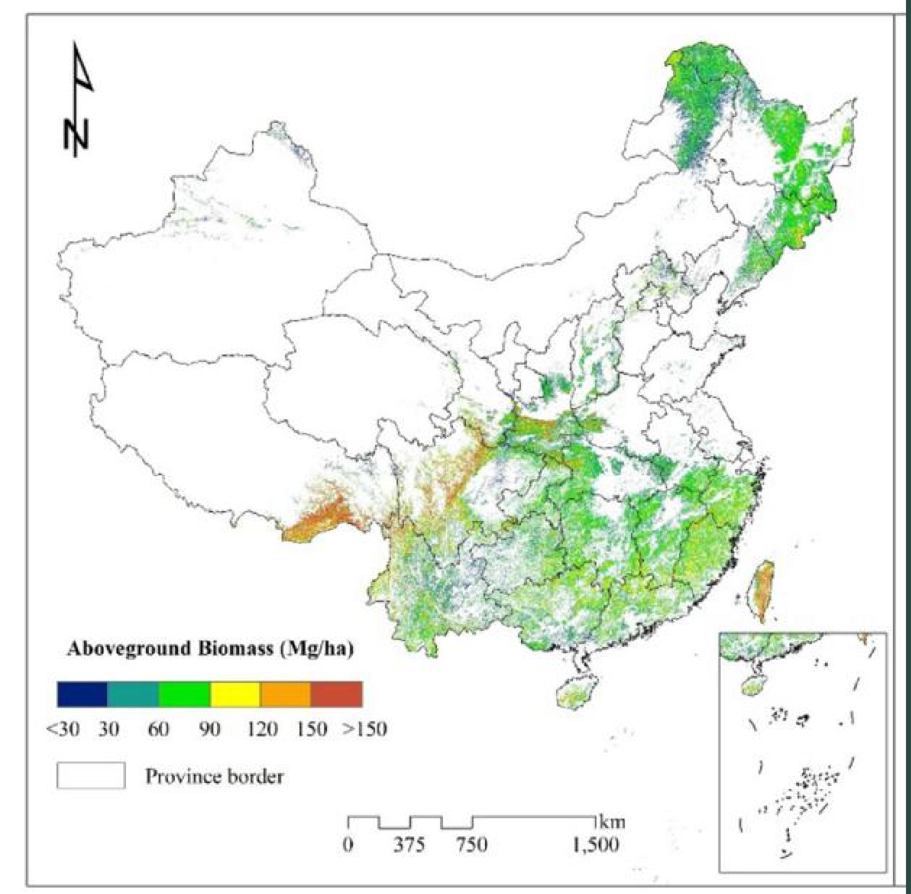




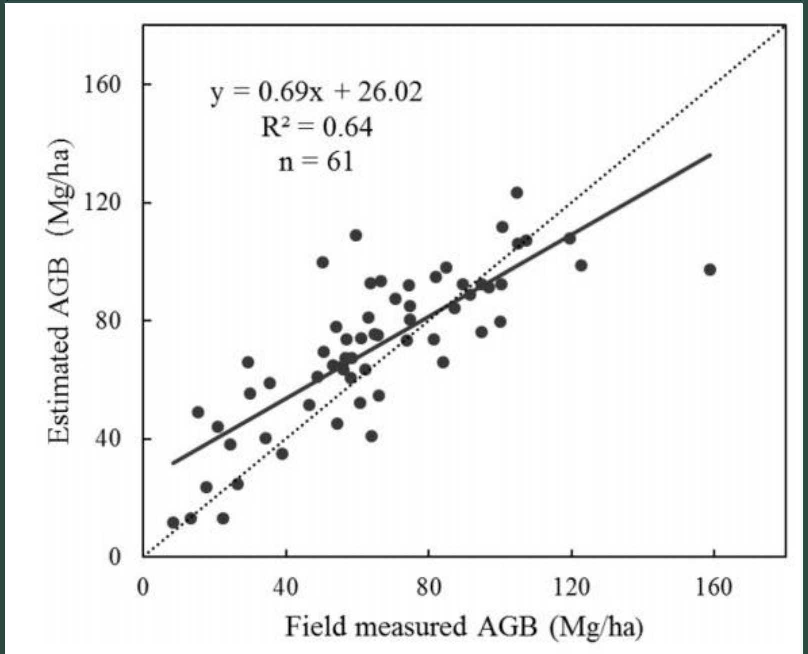
Footprint level AGB

__Source:__ Huang et al., 2019; XIng et al., 2016.

Example2: ICESat, Landsat, PALSAR



30-m resolution AGB map of China



Model performance evaluation

Footprint tree height → Footprint AGB → Continuous AGB

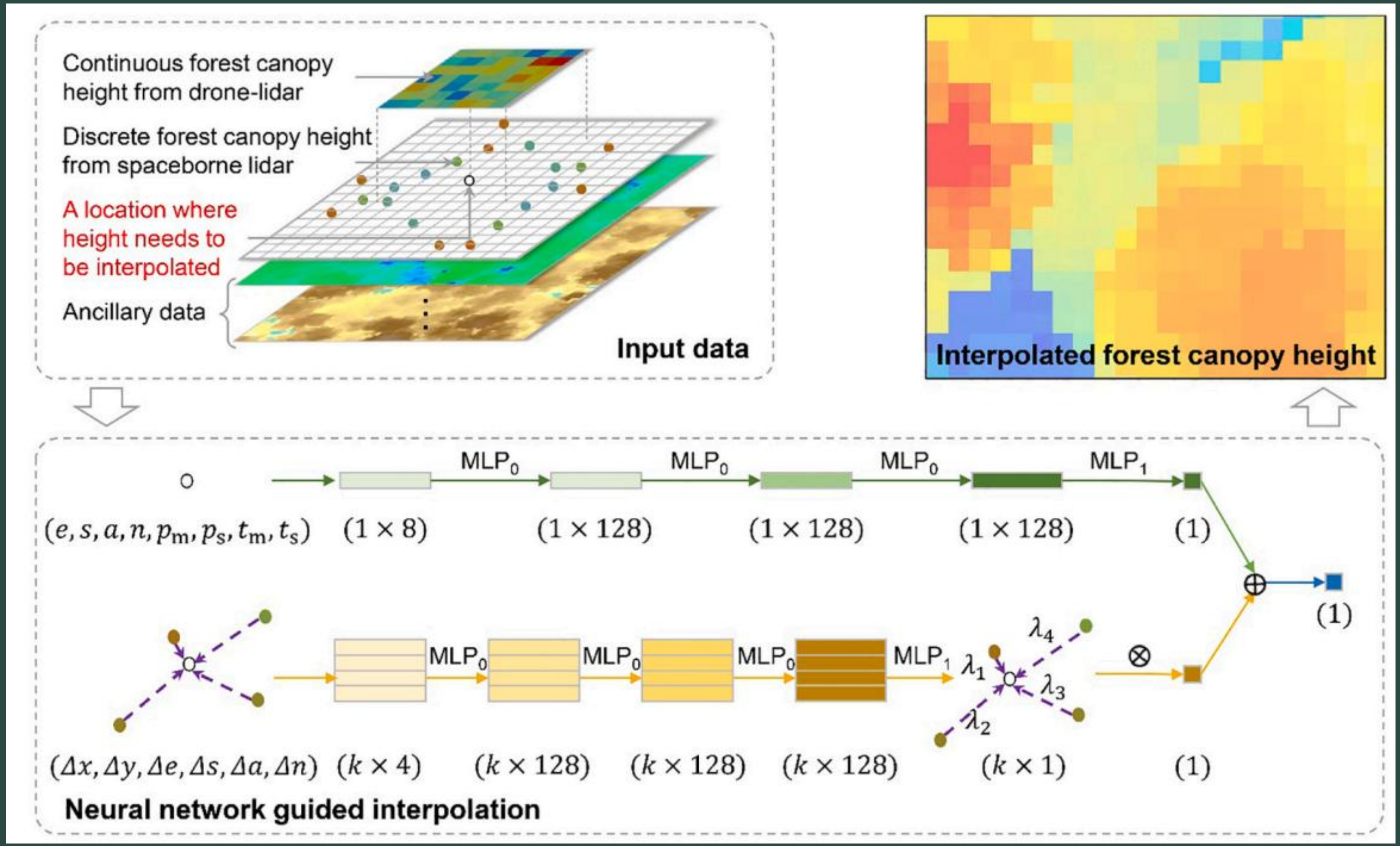
Allometric euaqtions

Continuous images

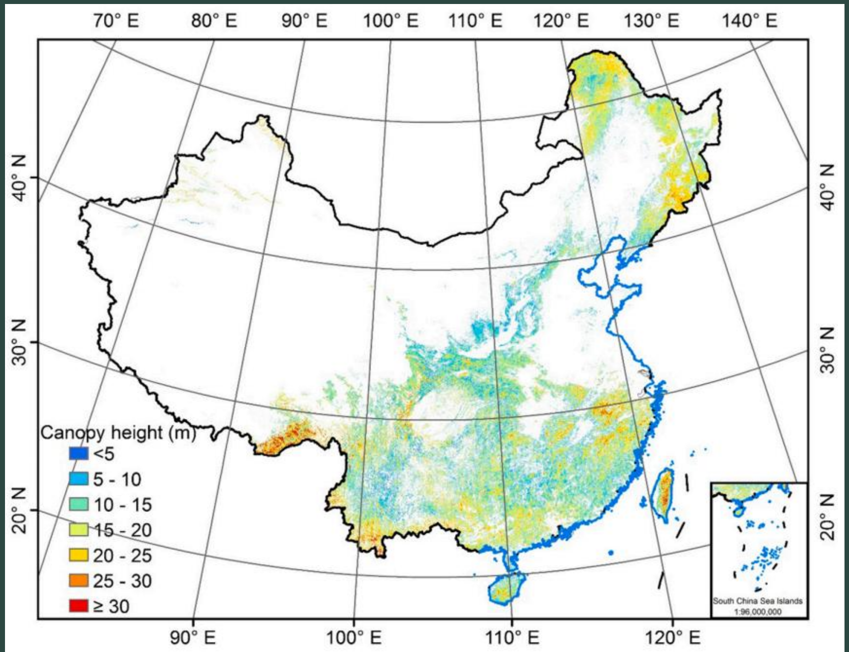
Source: *Huang et al., 2019.*

Example3: ICESat–2, GEDI, Sentinel–2

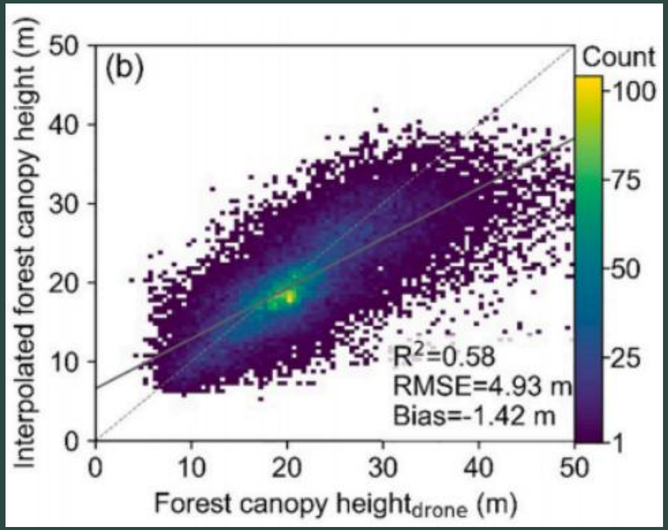
- First step: Canopy height mapping based on interpolation



Overall workflow



30m resolution canoopy height map of China



Model performance

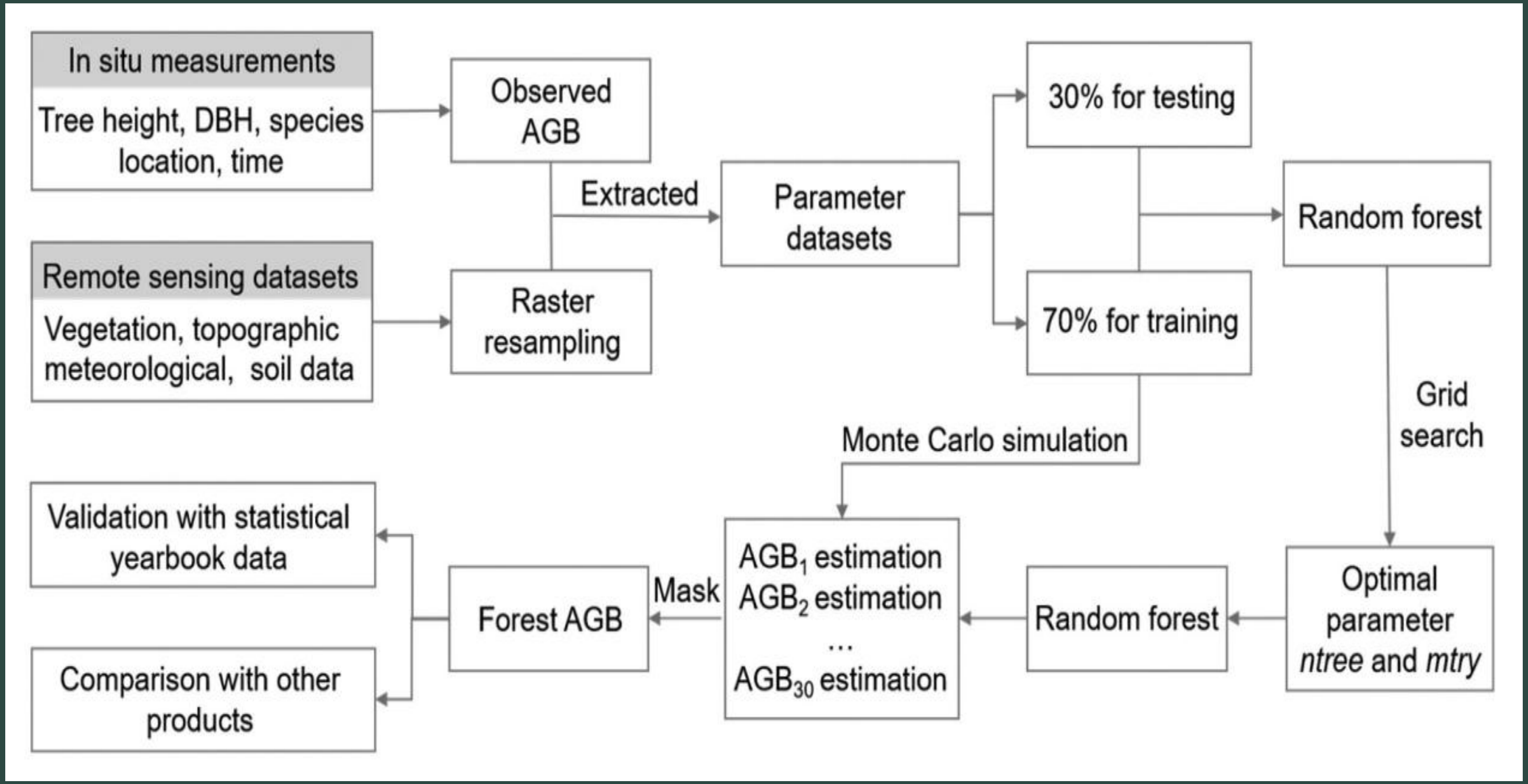
Source: *Liu et al., 2022; Yang et al., 2023.*

Example3: ICESat–2, GEDI, Sentinel–2

Table 2. Statistics for 17 variables used in this study.

Feature	Variables
Vegetation features	Forest canopy height (FCH, m)
	Normalized difference vegetation index (NDVI)
Terrain features	Elevation (m)
	Slope (°)
	Aspect (°)
Climate features	Average temperature (Tmean, °C)
	Temperature seasonality (Tseason, °C)
	Average precipitation (Pmean, mm)
	Precipitation seasonality (Pseason, mm)
	Average potential evapotranspiration (PET, mm)
	Maximum vapor pressure deficit (VPD, kPa)
Soil features	Average soil moisture (SM, mm)
	Soil organic carbon (SOC, kg m^{-2})
	Soil total nitrogen density (STN, kg m^{-2})
	Soil C: N mass ratios (C/N)
Ancillary data	Vegetation zone (Vegzone)
	Forest mask

- Second step: AGB mapping based on canopy height map

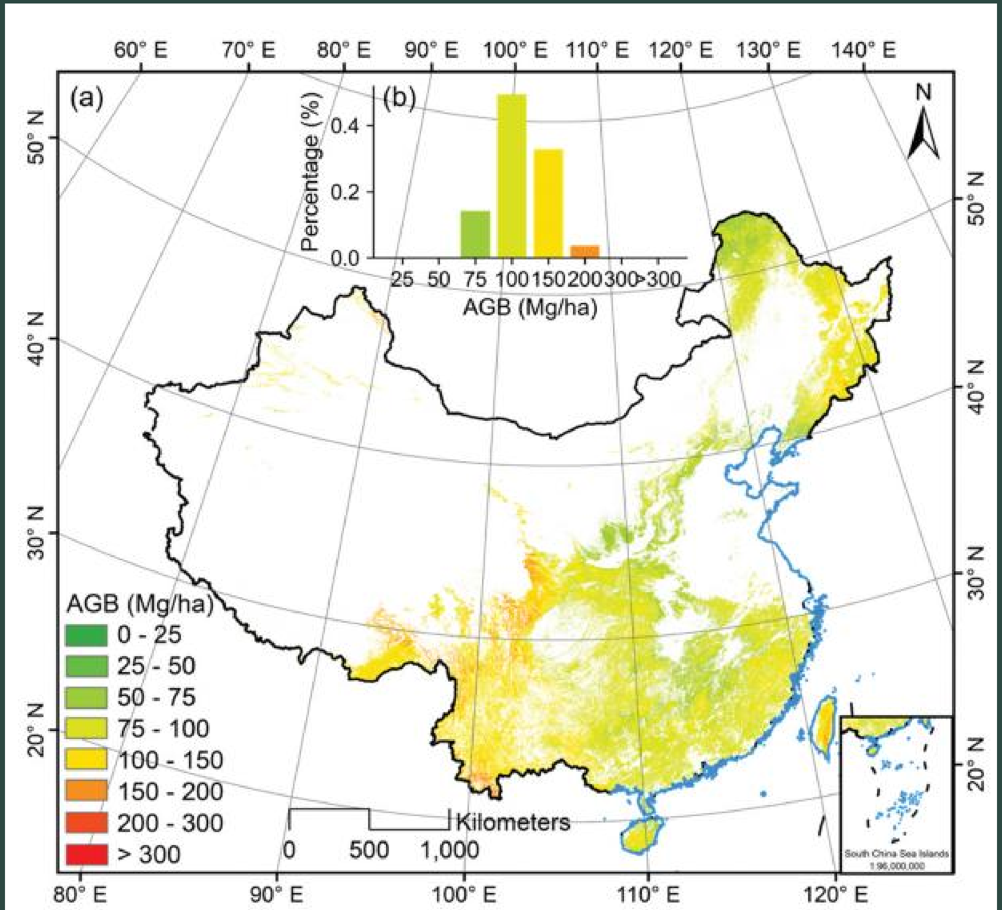


Overall workflow

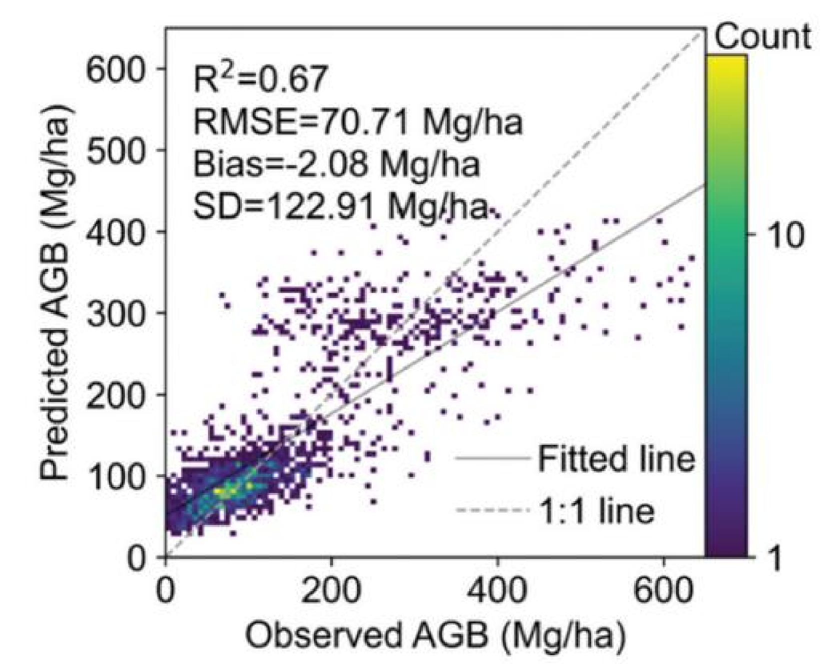
Source: *Liu et al., 2022; Yang et al., 2023.*

Example3: ICESat–2, GEDI, Sentinel–2

- Second step: AGB mapping based on canopy height map



30-m resolution AGB map of China



Model performance

Footprint tree height → Continuous tree height → Continuous AGB

Continuous images

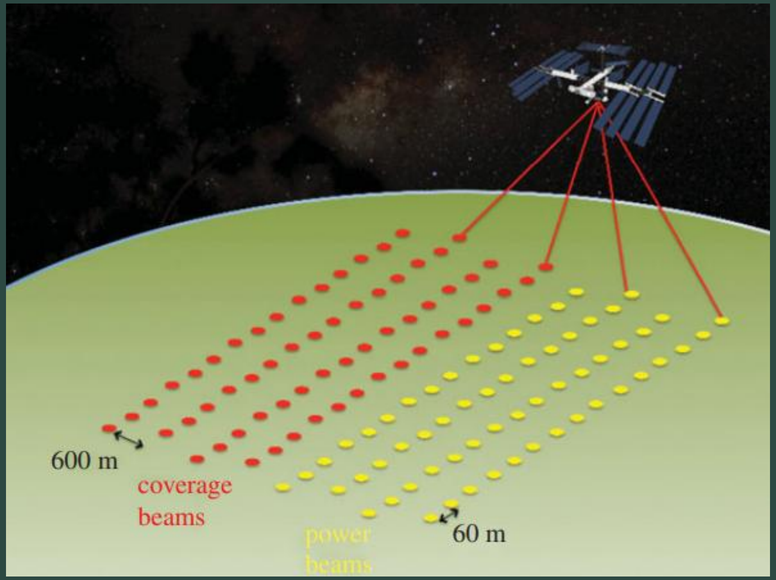
Other continuous variables

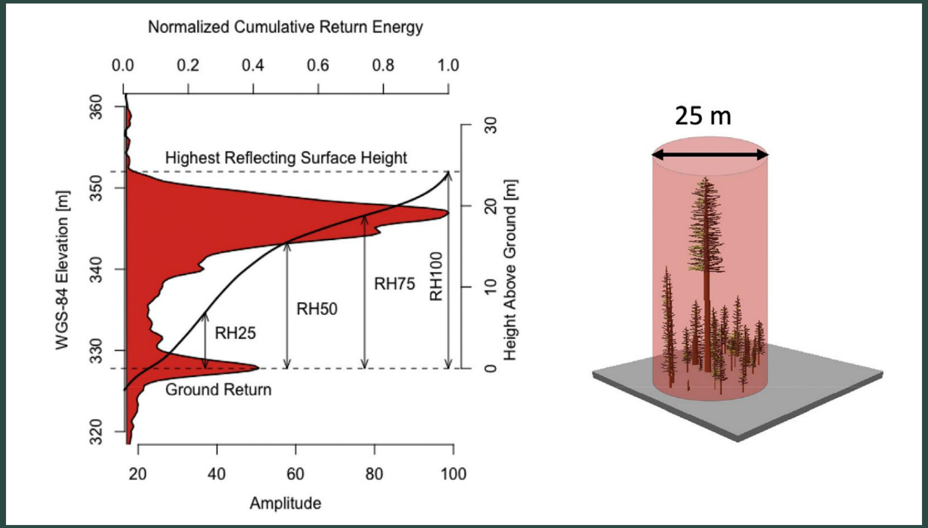
Source: *Liu et al., 2022; Yang et al., 2023.*

GEDI can do more...

Level	Product	Variables	Resolution
L1	01_A	Raw GEDI waveform	25m
	01_B	Geolocated GEDI waveform	25m
L2	02_A	Geolocated elevation and height metrics	25m
	02_B	Canopy cover and vertical profile metrics	25m
L3	03	Gridded land surface metrics	1km
L4	04_A	Footprint level aboveground biomass density	25m
	04_B	Gridded aboveground biomass density	1km

Ground elevation, canopy top height, relative return energy metrics (e.g., canopy vertical structure), canopy cover, plant area index, plant area volume density, foliage height diversity, **aboveground biomass density**...

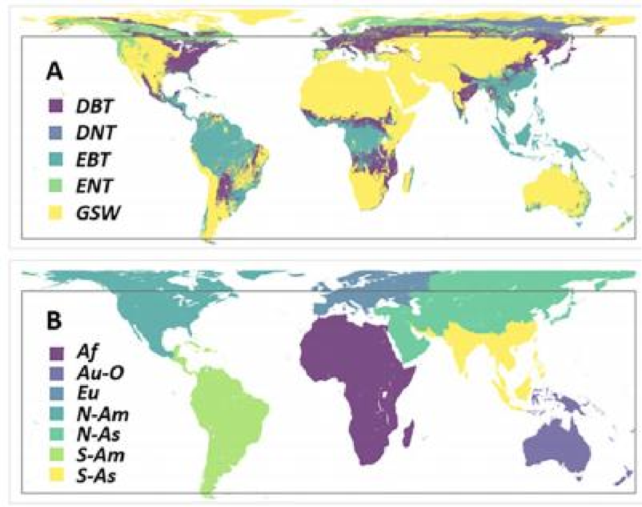


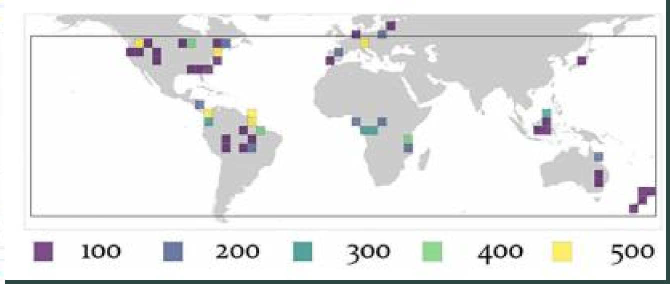


Source: *Hancock et al., 2021; Dubayah et al., 2020.*

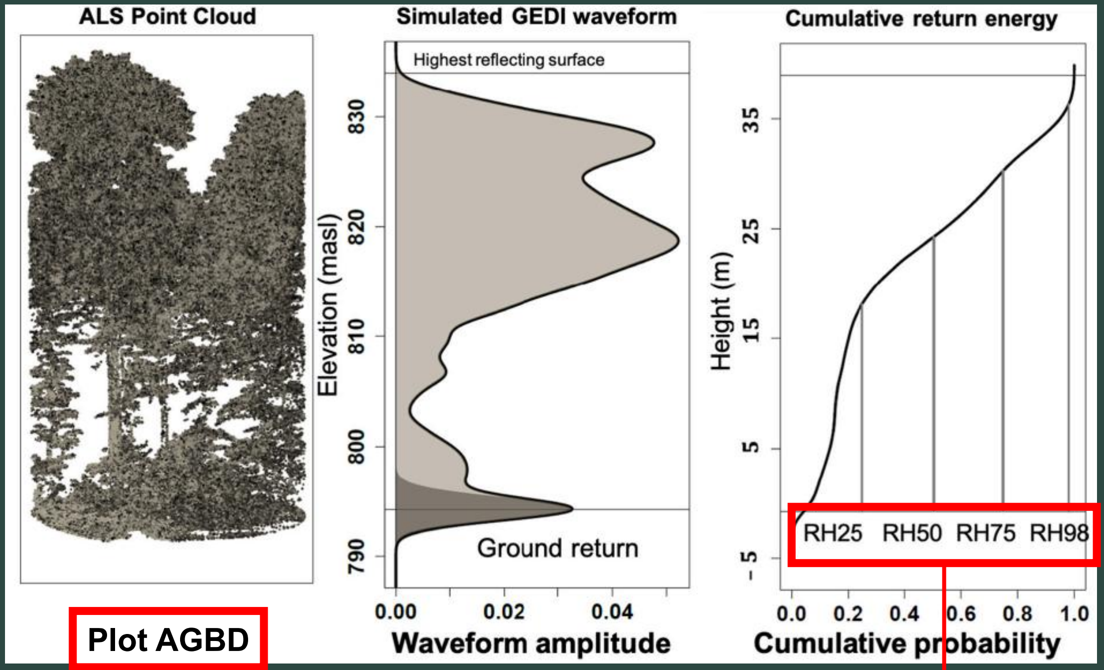
How does GEDI estimate footprint AGBD?

Pre-launch calibration *VS* Post-hoc calibration ?





Simulated waveform stratification



Parametric Ordinary Least Squares (OLS) model

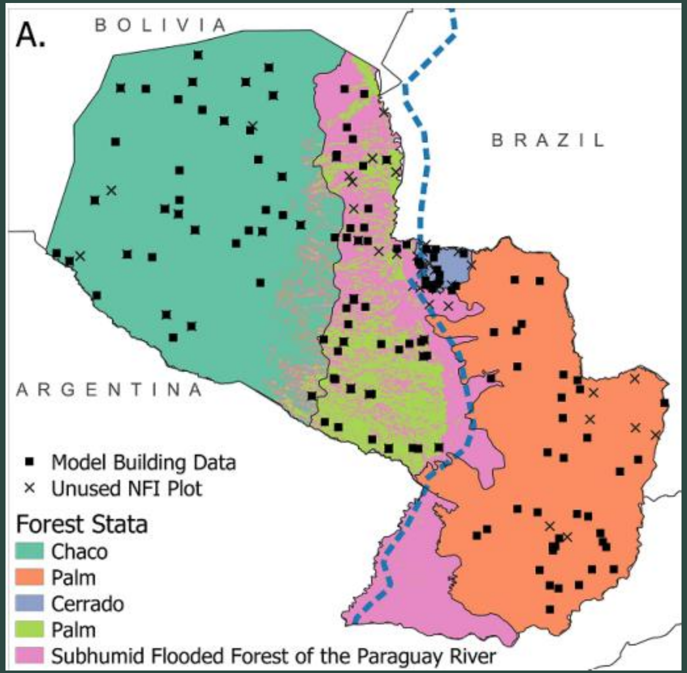
eg. DBT: $\text{AGBD} = 1.017 \times (-110.059 + 5.134 \times \sqrt{RH60 + 100} + 6.172 \times \sqrt{RH98 + 100})^2$

- More training samples especially in under-represented areas (eg. **Continental Asia**)
- Improved allometric models (eg. Using LiDAR)
- Alternative models (eg. Machine Learning)
- Auxiliary data (eg. L2B metrics)

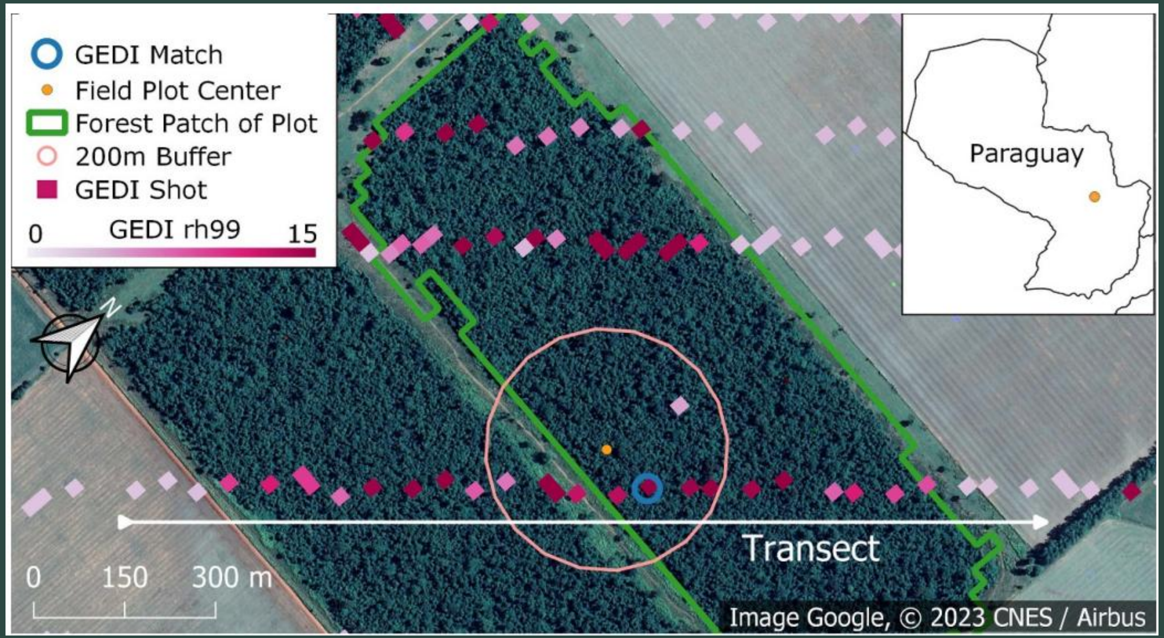
"The importance of these data cannot be stressed enough and their continued development should be encouraged and supported."

__Source:__ Kellner et al., 2023; Duncanson et al., 2022.

Example4: Calibrated GEDI footprint AGBD model



Field plots



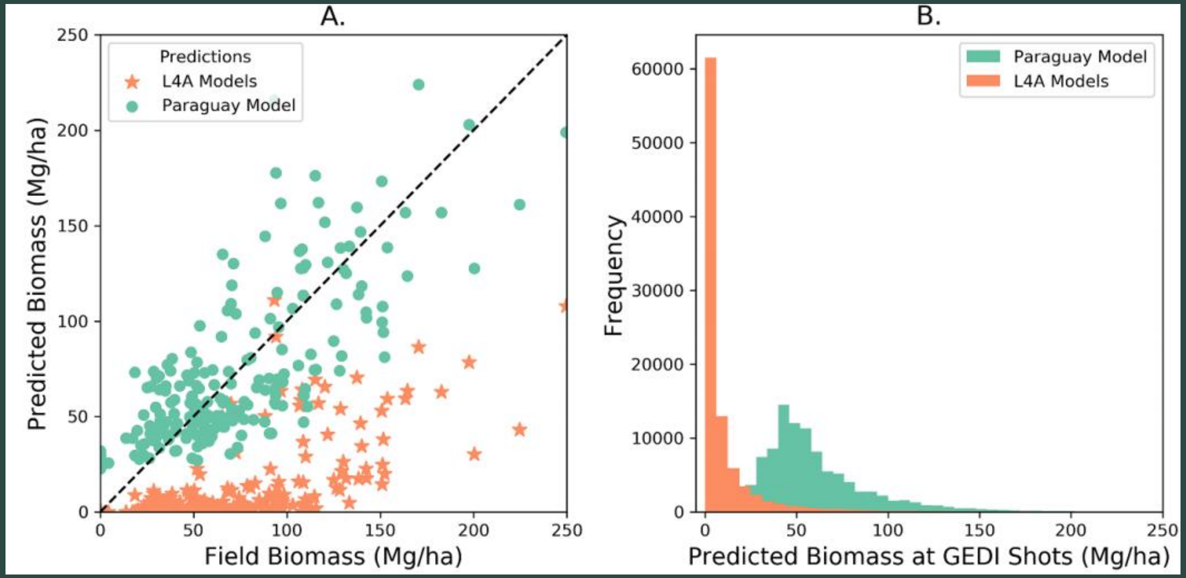
Matching field plots with GEDI footprints

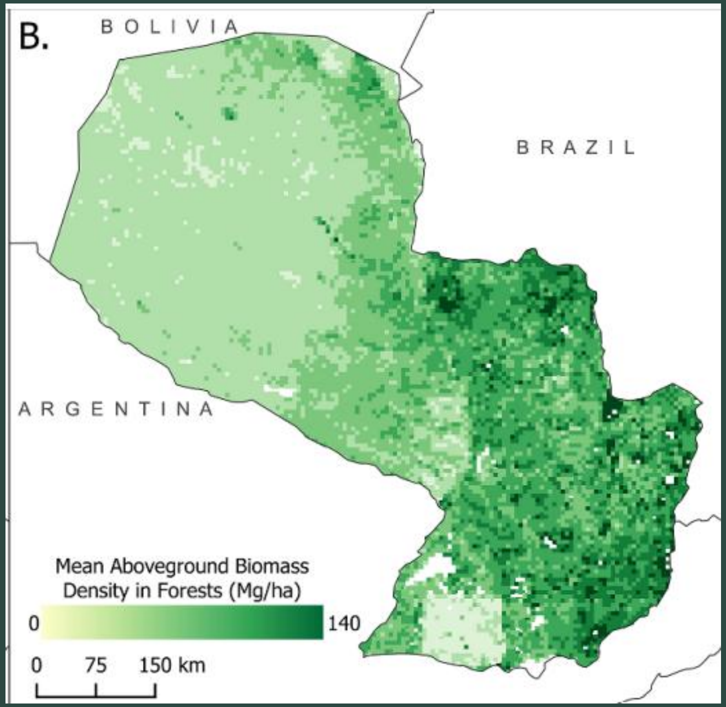
"Reconstruct" GEDI L4A model:

$$\sqrt{\text{AGBD}_i} = \beta_0 + \beta_1 \text{rh5}_i + \beta_2 \sqrt{\text{rh45}_i} + \beta_3 \sqrt{\text{rh91}_i} + \epsilon_i \; \epsilon_i \sim N\left(0, \sigma^2\right)$$

Source: *Bullock et al., 2023.*

Example4: Calibrated GEDI footprint AGBD model





L4A model VS Calibrated model

6km grid AGB map

Footprint RH → Footprint AGBD → Continuous AGB

Local plots calibrated model

Hybrid inference

Source: *Bullock et al., 2023.*

Example5: Other ways of calibration

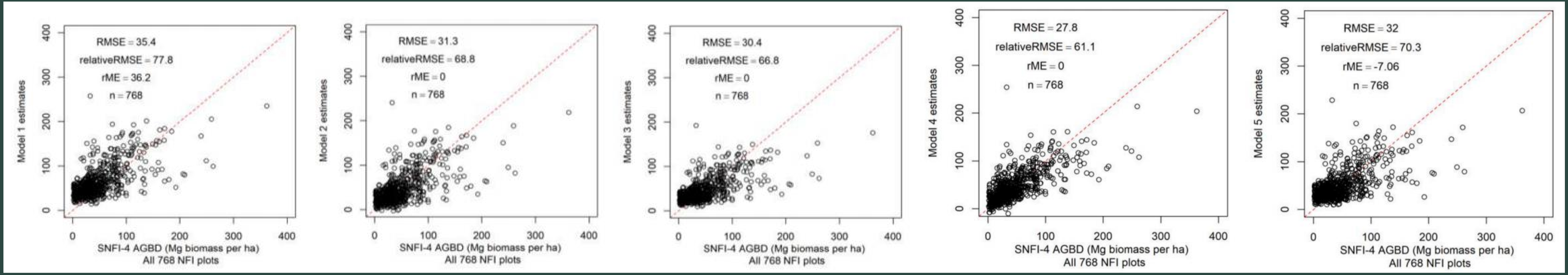
Plot AGBD ⟷ **L4A AGBD**

Calibration

ALS point cloud → Simulated GEDI waveform

Model	Description	Parameters		
1	$AGBD_{SNFI-4} = \beta_0 + \beta_1 AGBD_{L4A} + \varepsilon$	$\beta_0 = 0$	$\beta_1 = 1$	
2	$AGBD_{SNFI-4} = \beta_0 + \beta_1 AGBD_{L4A} + \varepsilon$	$\beta_0 = f$ (OLS)	$\beta_1 = 1$	
3	$AGBD_{SNFI-4} = \beta_0 + \beta_1 AGBD_{L4A} + \varepsilon$	$\beta_0 = f$ (OLS)	$\beta_1 = f$ (OLS)	
4	$AGBD_{SNFI-4} = \beta_0 + \beta_1 AGBD_{L4A} + \beta_2 FT + \varepsilon$	$\beta_0 = 0$	$\beta_1 = 1$	$\beta_2 = f$ (OLS)
5	$AGBD^{FT}_{SNFI-4} = \beta_0 + \beta_1 AGBD_{L4A} + \beta_2 FC + \varepsilon$	$\beta_0 = 0$	$\beta_1 = 1$	$\beta_2 = f$ (OLS)

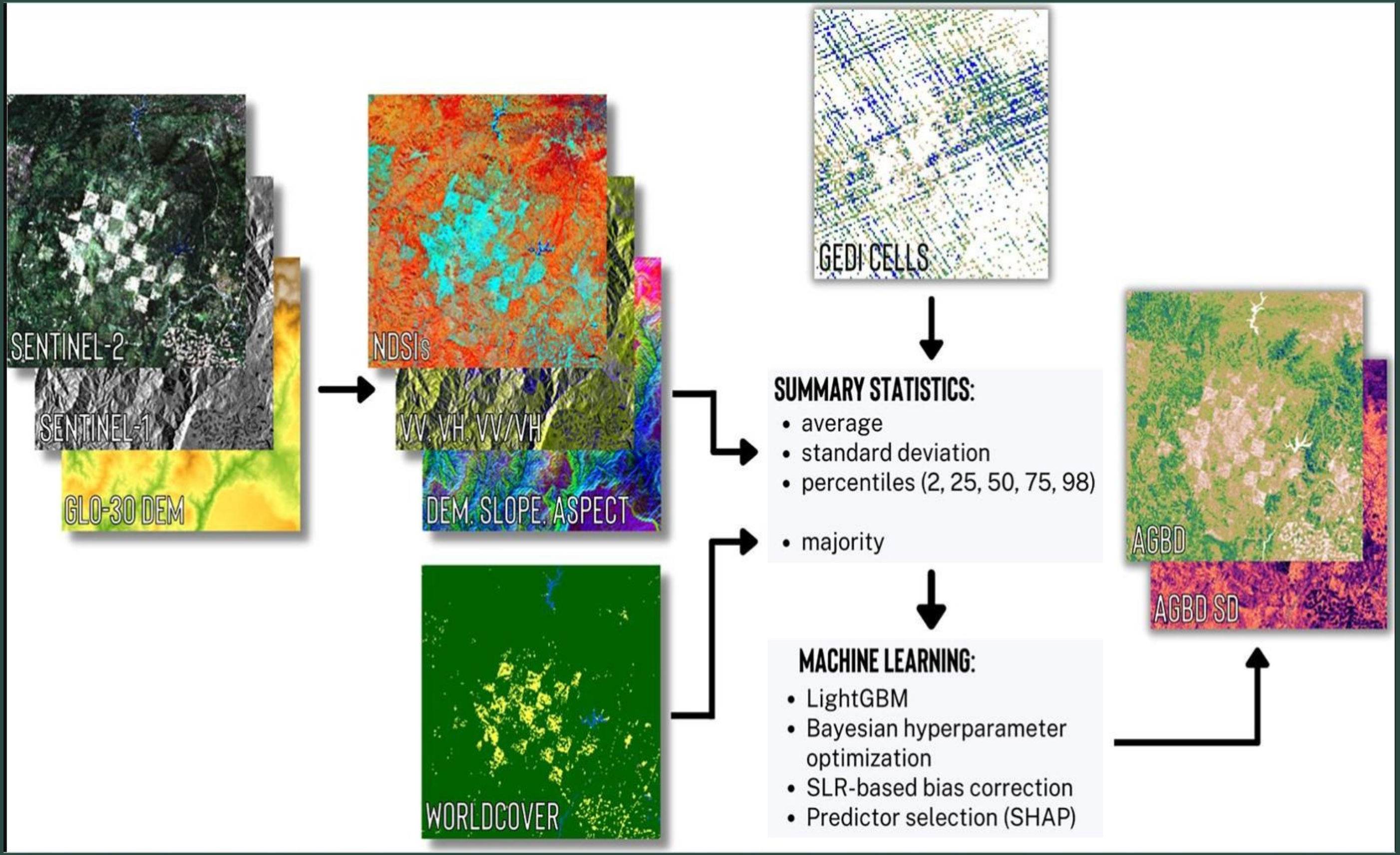
* $AGBD_{SNFI-4}$ is the measured AGBD in SNFI-4 samples, $AGBD_{L4A}$ is the AGBD prediction using L4A model from Duncason et al., (2022), *FT* is a regression factor defining the forest type of SNFI-4 plots, $AGBD^{FT}_{SNFI-4}$ is the measured $AGBD_{SNFI-4}$ for a given *FT* and *FC* expresses the ratio between the number of first ALS echoes over a height-break of 2 m.



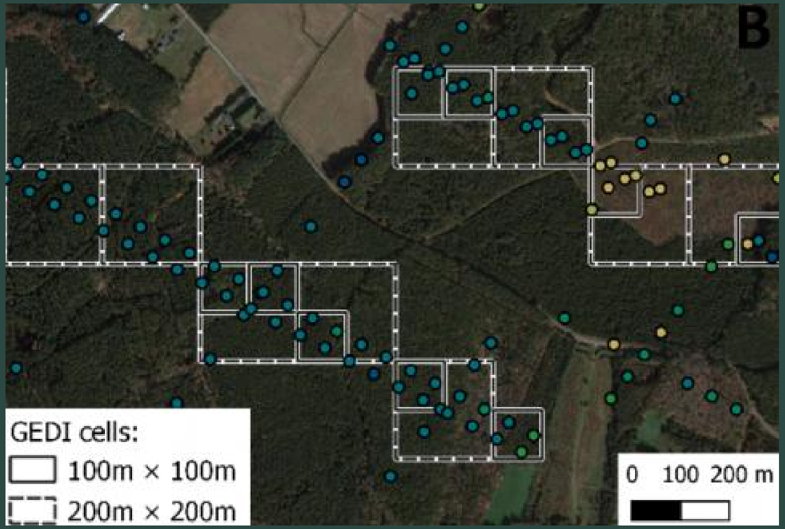
Model performance

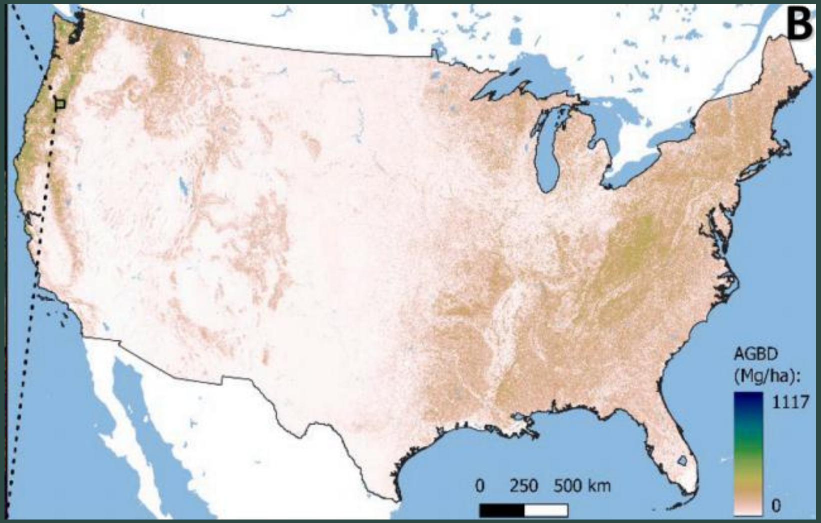
Source: *Pascual et al., 2023.*

Example6: GEDI_04A, Sentinel–1, Sentinel–2



Overall workflow



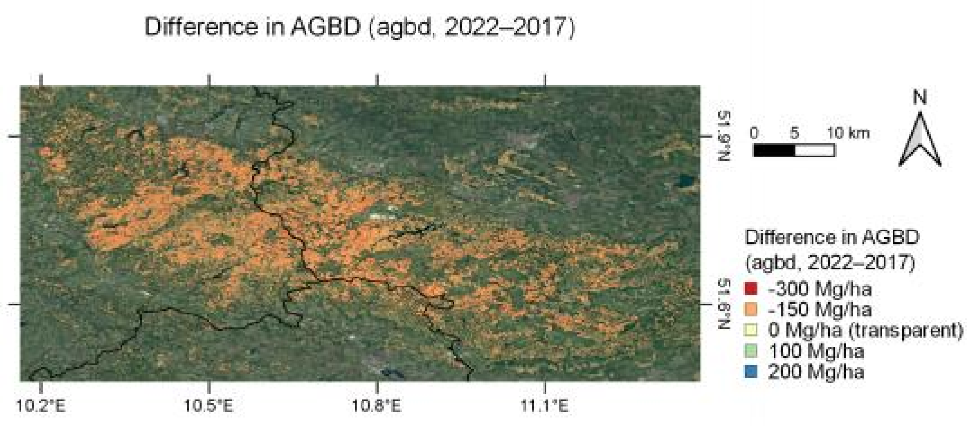


100m grid AGB map

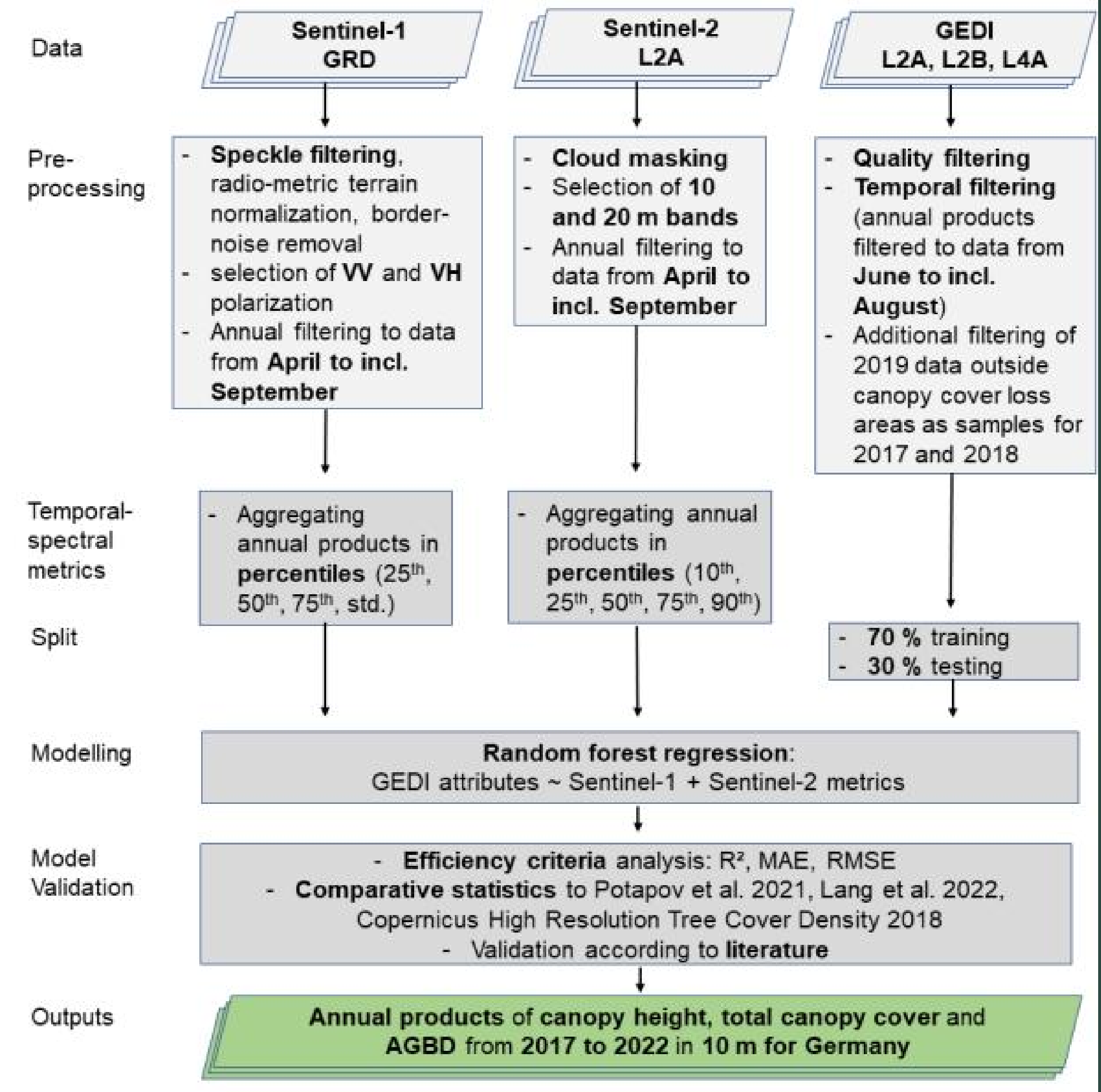
Source: *Shendryk, 2022.*

Example7: GEDI_04A time series AGB map

GEDI Attribute	Above-Round Biomass Density (Agbd)		
Efficiency Criteria	R^2 [%]	MAE [Mg/ha]	RMSE [Mg/ha]
2017	61.1	41.2	65.3
2018	62.8	38.8	61.3
2019	61.3	40.6	63.7
2020	61.9	38.3	60.2
2021	50.9	47.7	73.0
2022	54.7	39.5	62.4
Mean	58.8	41.0	64.3



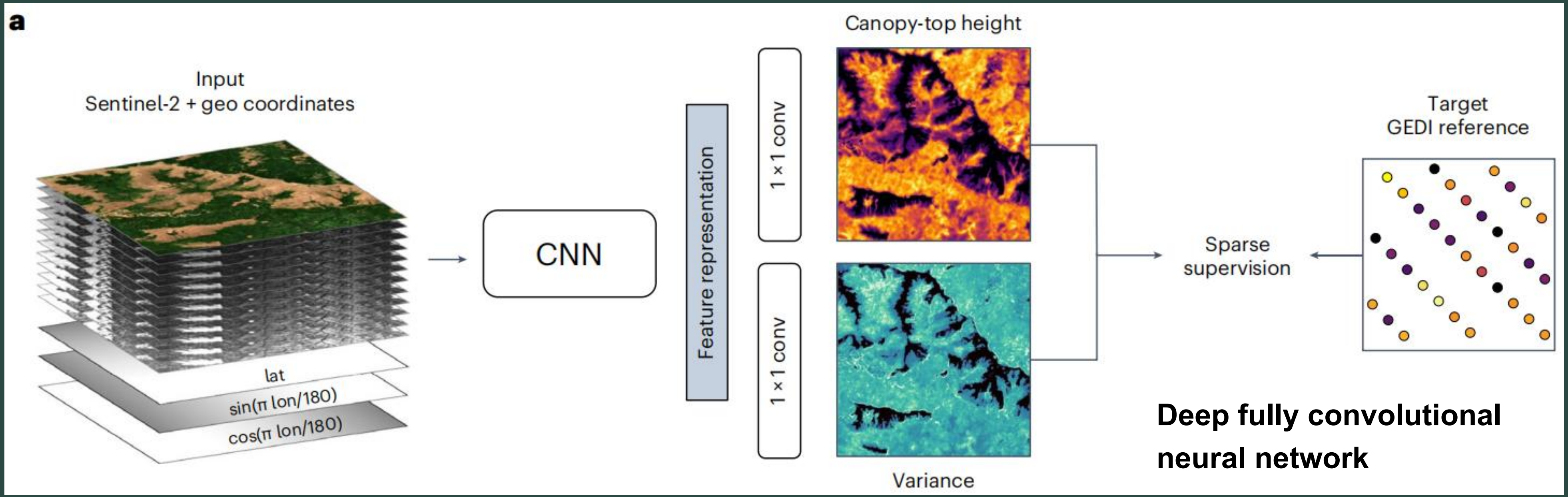
10m resolution AGB map



Overall workflow

Source: *Kacic et al., 2023.*

Example8: The state-of-the-art global canopy height map

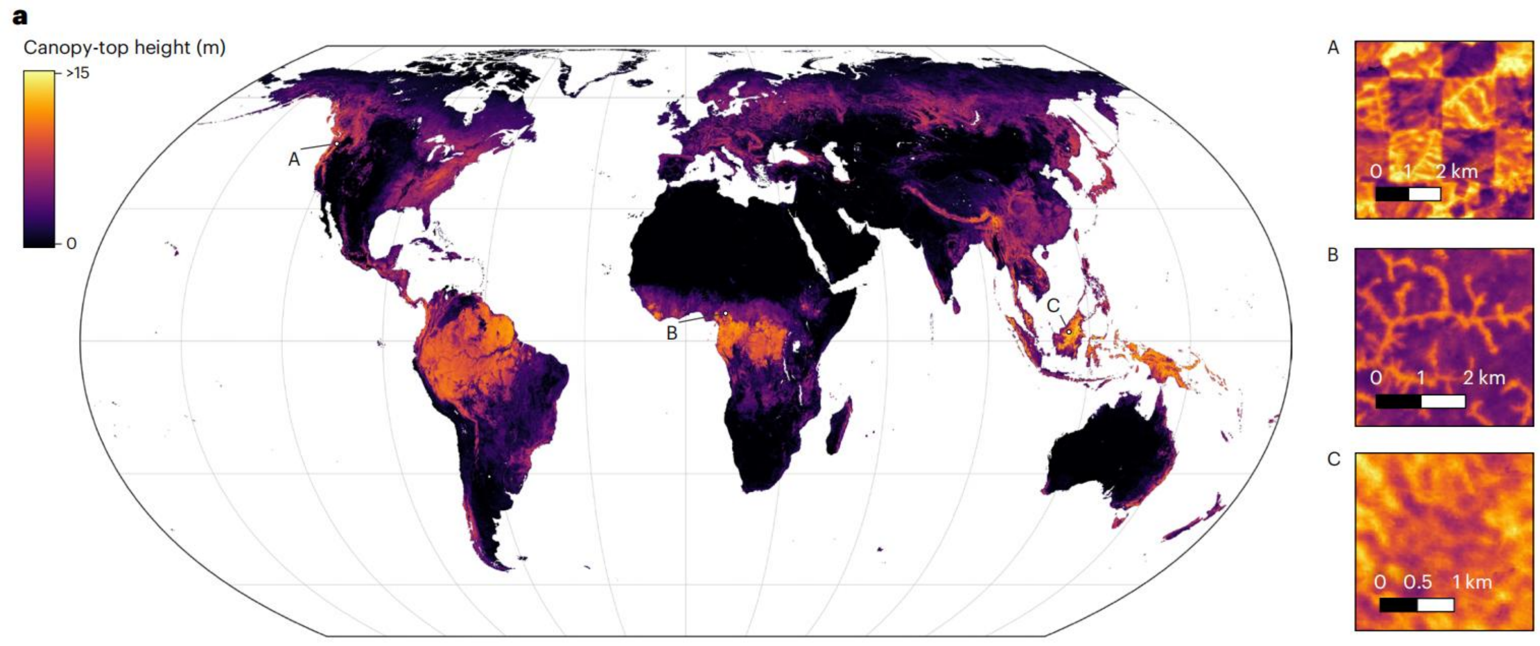


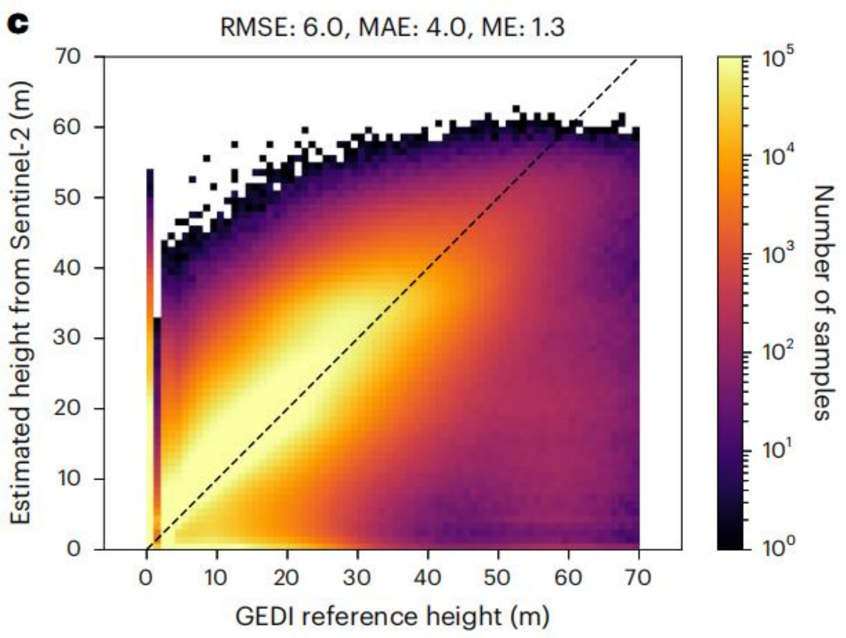
Dataset:

- **GEDI**: Calibrated footprint canopy height, from April to August in 2019 & 2020
- **Sentinel-2**: L2A 12 band values (+ lat & lon), the tile with the least cloud coverage between May and September in 2020
- 600 million samples by extracting Sentinel-2 image patches of 15 × 15 pixels around every GEDI point

Source: Lang et al., 2023.

Example8: The state–of–the–art global canopy height map





10m resolution global canopy height map
for the year 2020

Model performance

1. **Model**: GEDI L4A footprint AGBD + 10m Sentinel-2 = 10m AGBD map
2. **Data**: 10m canopy height map + other variables = 10m AGBD map

Source: *Lang et al., 2023.*

Mapping high–resolution, wall–to–wall forest aboveground biomass in Northeast China by integrating GEDI and Sentinel–2 *

Dataset

Time period: 2022 04-09

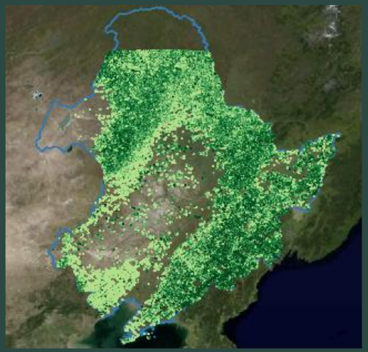
GEDI L4A raw data

↓ Quality flag

Good quality footprints

↓ LULC map

Forested footprints



Sparse points

AGBD, lat, lon

Overlay

+

Sentinel-2 L2A raw data

QA60 + Median composite ↓

Cloud-masked image

Resampling ↓

10-m resolution image

Continuous image

12 band values

=

	AGBD	Lat	Lon	Band1	Band2	···	Band12
0	63.10	45.37	123.09	0.025	0.029	···	0.083
1	75.70	46.33	125.61	0.028	0.026	···	0.060
2	71.38	46.33	125.61	0.028	0.026	···	0.069
···	···	···	···	···	···	···	···
5716319	93.83	44.85	130.71	0.039	0.033	···	0.108

Band values extraction

*This report is the outcome of the final project for the "Machine Learning for Earth and Environmental Sciences" course at the University of Lausanne, Autumn Semester 2023.

Methodology

12 bands values

Lat + Lon

Regression → AGBD values

Linear regression

Train/test split

StandardScaler

PCA

LinearRegression

MSE & R^2

Embedding learning

Train/test split

RandomForestRegressor

XGBRegressor

LGBMRegressor

MSE & R^2

Neural networks

Train/validation/test split

StandardScaler

ANN

Dense (128, ReLU)

Dense (64, ReLU)

Dense (32, ReLU)

Dense (16, ReLU)

Dense (1, Linear)

1-D CNN

Conv1D (64, 3, ReLU) ×2

Maxpooling (2)

Conv1D (128, 3, ReLU) ×2

Maxpooling (2)

Conv1D (256, 3, ReLU) ×2

Maxpooling (2)

Flatten

Dense (128, ReLU)

Dense (1, Linear)

MSE & R^2

Results

Linear regression

Train: MSE=7158, R^2=0.036
Test: MSE=7220, R^2=0.036

Random forest

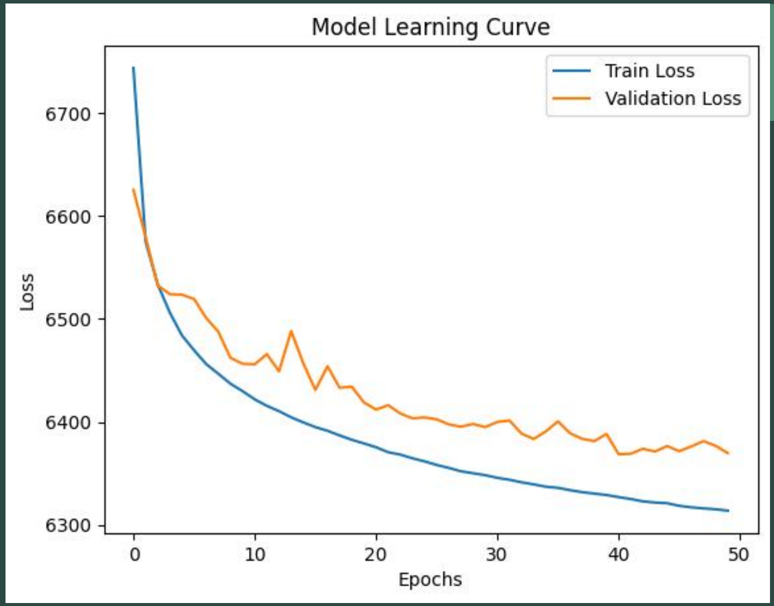
n_estimators=50:
Train: MSE=782, R^2=0.89
Test: MSE=5424, R^2=0.28

Overfitting?

n_estimators=50, max_depth=120,
min_samples_leaf=5, max_features=20:
Train: MSE=2740, R^2=0.63
Test: MSE=5393, R^2=0.28

ANN

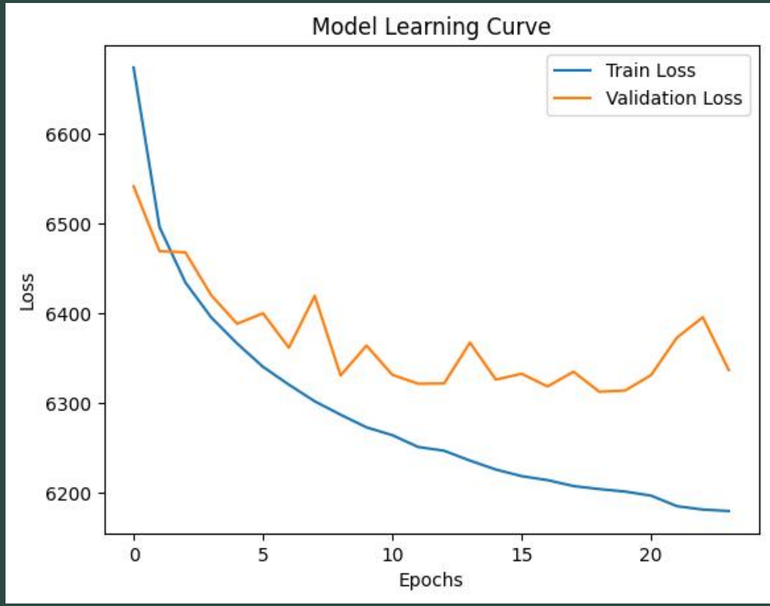
Test:
MSE=6301,
R^2=0.148



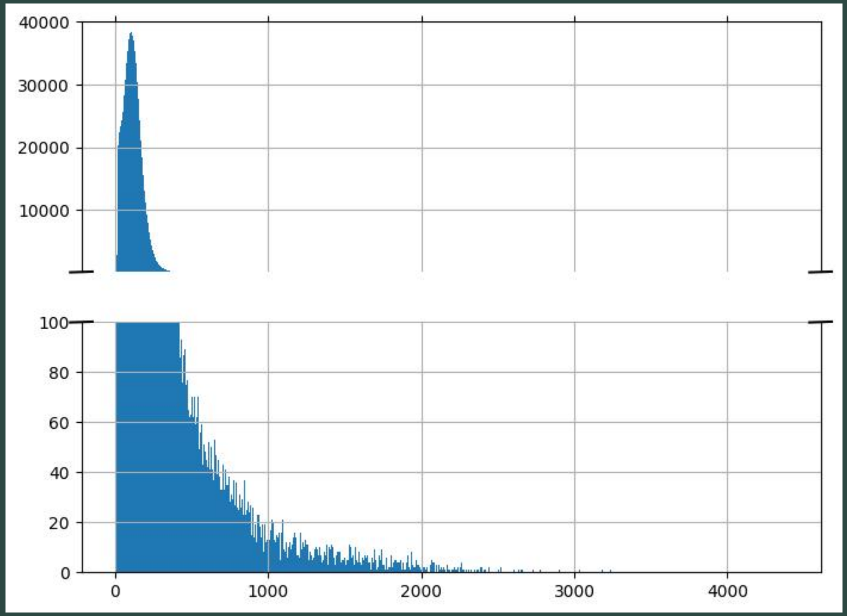
Optimizer: Adam
Loss: MSE
Epochs: 50
Batch size: 256
Callbacks: 5

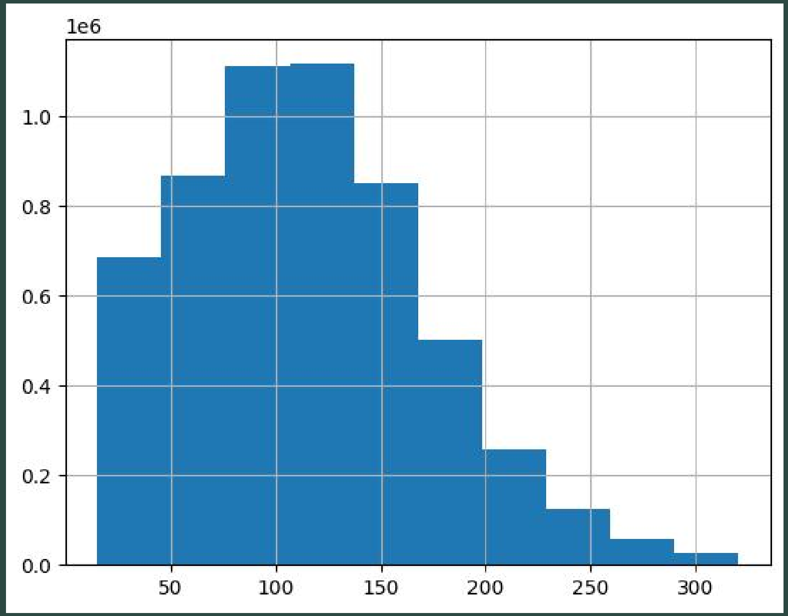
Underfitting?

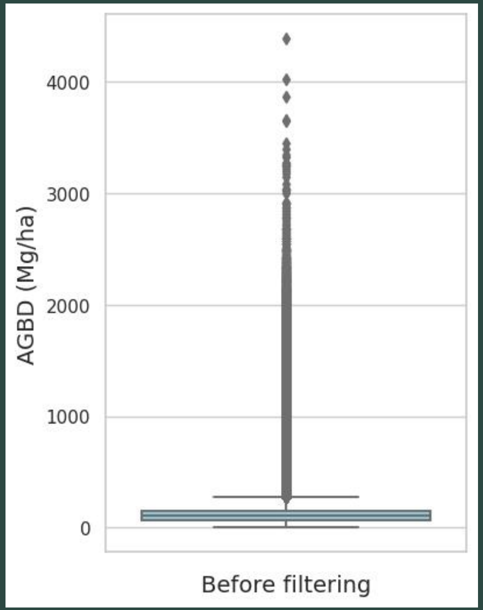
CNN



Test:
MSE=6266,
R^2=0.153



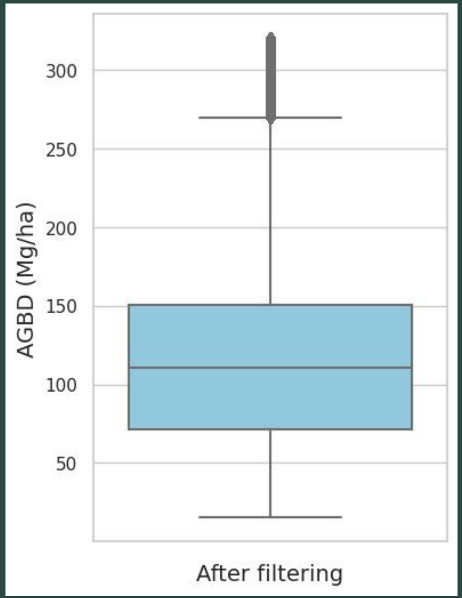




5,716,320 samples

Data filtration

$P_1 \sim P_{99}$



5,602,097 samples

Dataset

Methodology

Train: 80%, Test: 20%

Train: 70%, Validation: 15%, Test: 15%

Random forest

n_estimators = 50

Neural networks

Optimizer: Adam, Loss: MSE,
Epochs: 100, Batch size: 256, Callbacks: 10

Light GBM

Bayesian Optimization

init_points = 50,

n_iter = 100

Hyperparameter	Range	Best
n_estimators	50 - 2000	1261
max_depth	100 - 1000	756
num_leaves	1000 - 10000	6499
min_child_samples	10 - 200	52
colsample_bytree	0.5 - 1	0.64
subsample	0.5 - 1	0.88
learning_rate	0.01 - 0.5	0.04

MLP

Dense (128, ReLU)

Dense (64, ReLU)

Dense (32, ReLU)

Dense (16, ReLU)

Dense (1, Linear)

1-D CNN

Conv1D (64, 3, ReLU) ×2
Maxpooling (2)
Conv1D (128, 3, ReLU) ×2
Maxpooling (2)
Conv1D (256, 3, ReLU) ×2
Maxpooling (2)
Flatten
Dense (128, ReLU)
Dense (1, Linear)

Results

Evaluation metrics:

RMSE

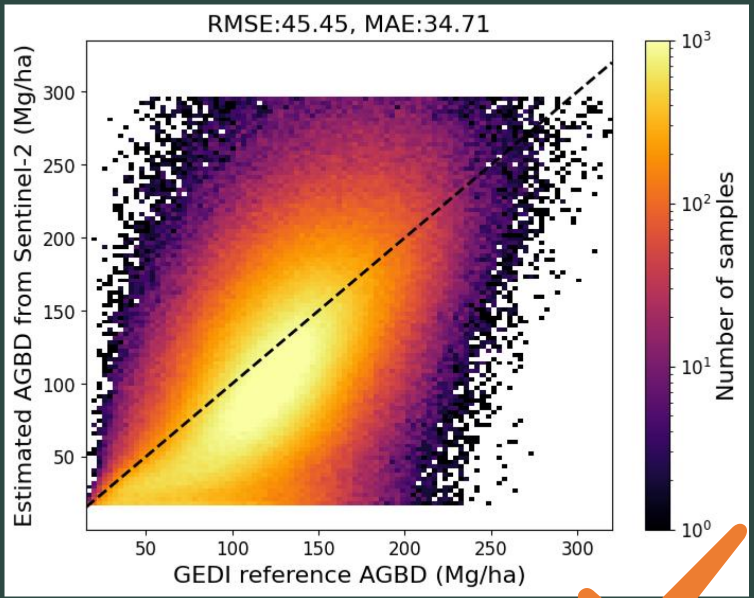
(Root Mean Square Error)

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

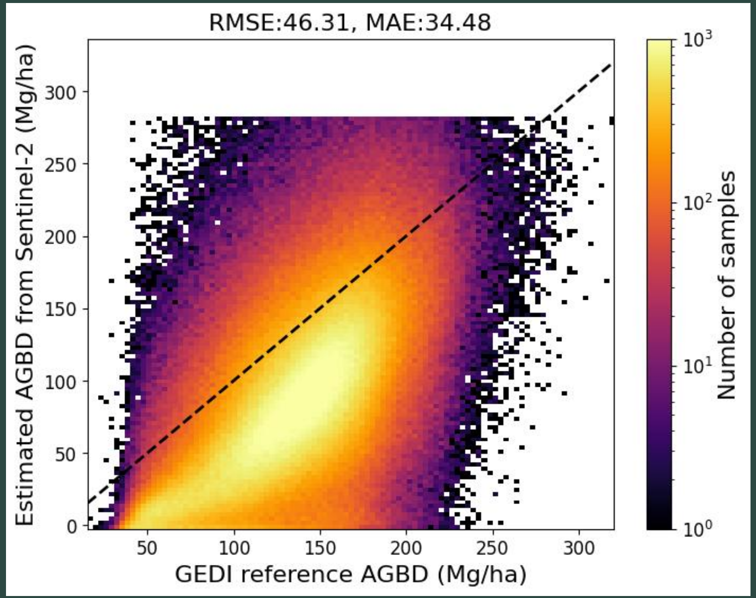
MAE

(Mean Absolute Error)

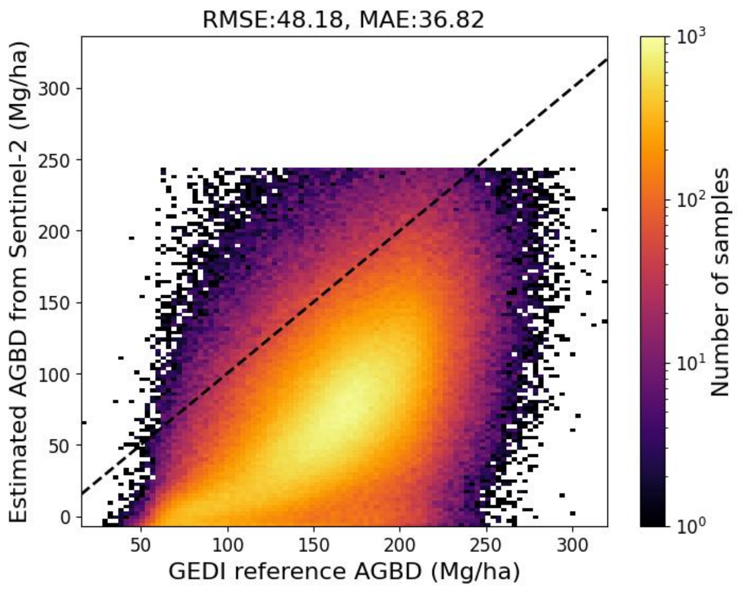
$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|$$



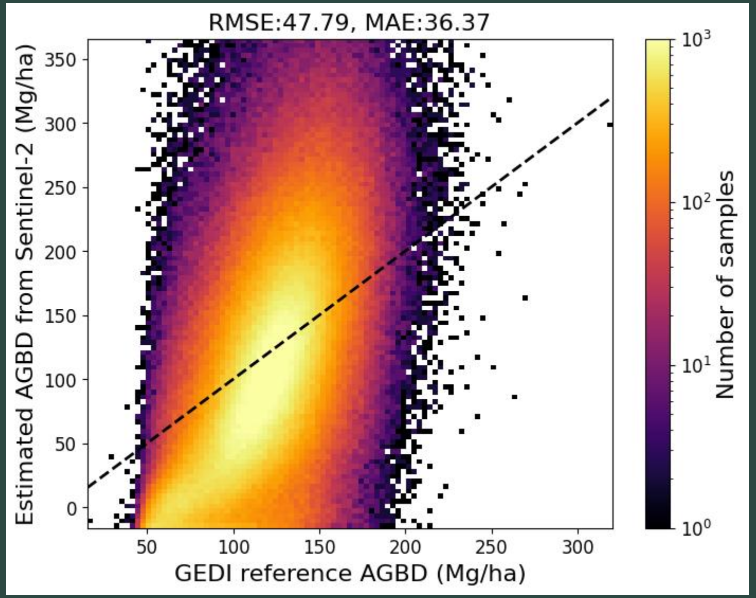
Random Forest



Light GBM



ANN

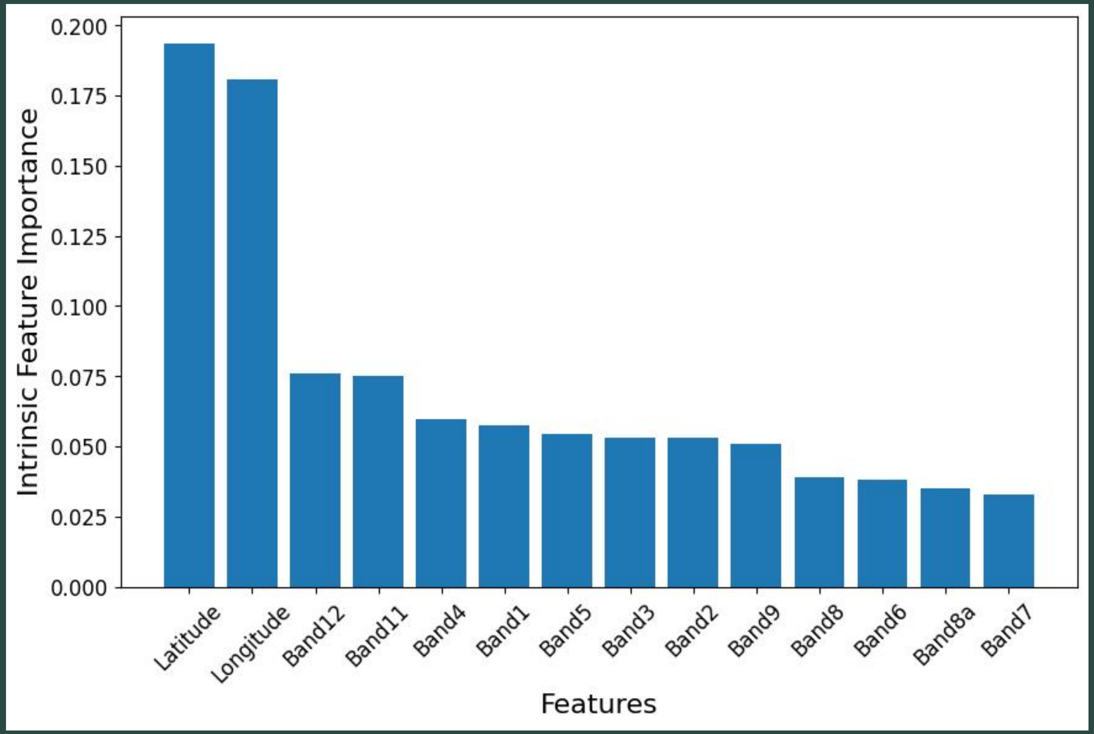


1-D CNN

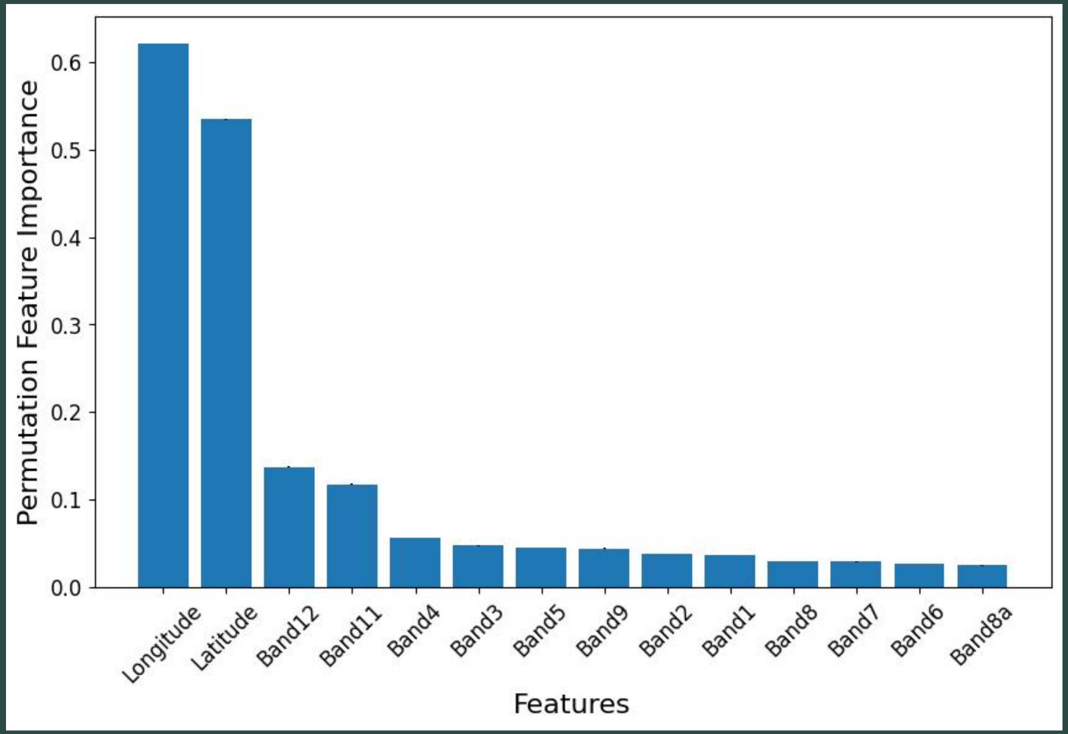
Results

More important features:

- ➢ Latitude, longitude (location)
- ➢ Band 12, 11 (short wave infrared)

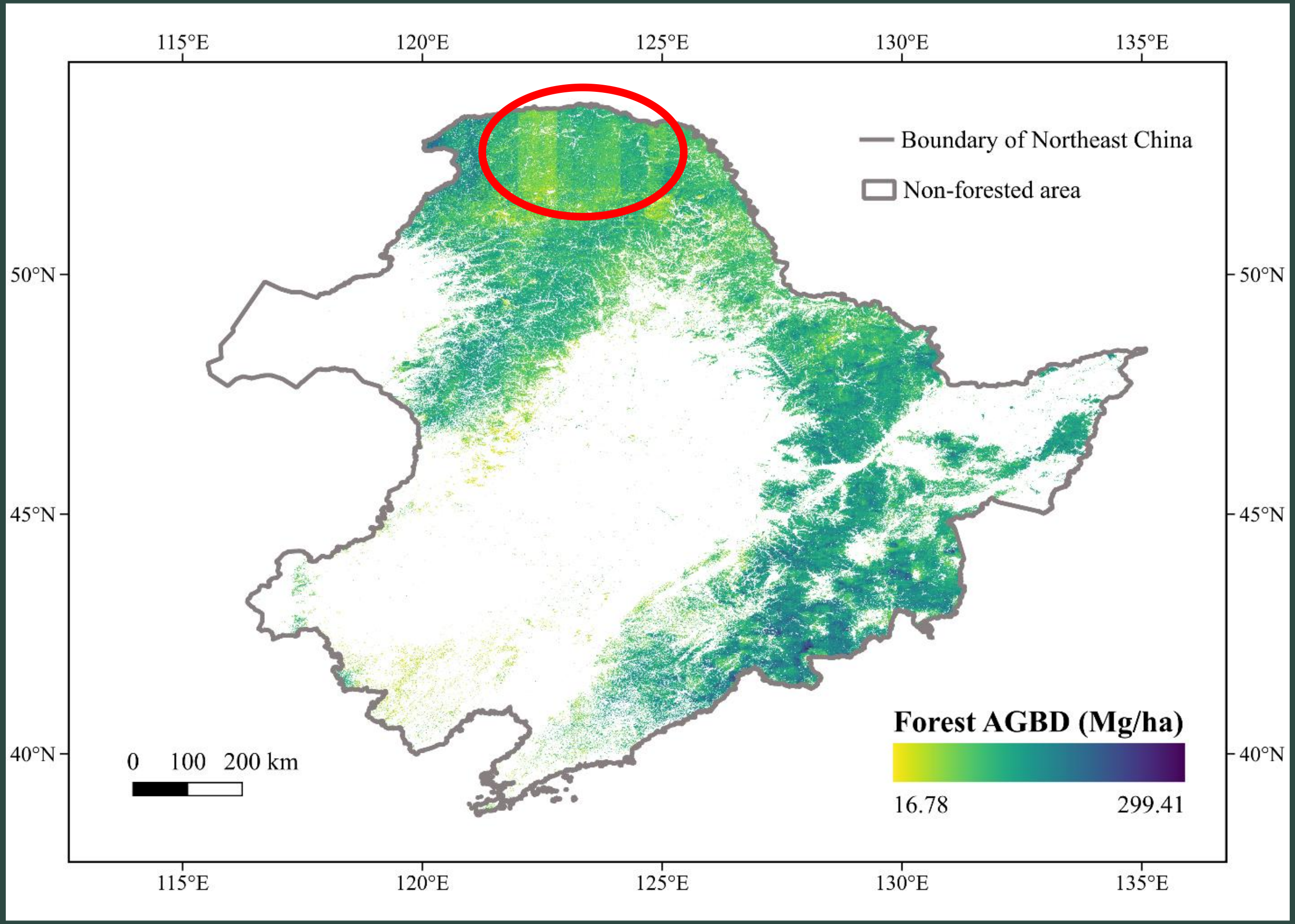


Intrinsic feature importance score



Permutation feature importance score

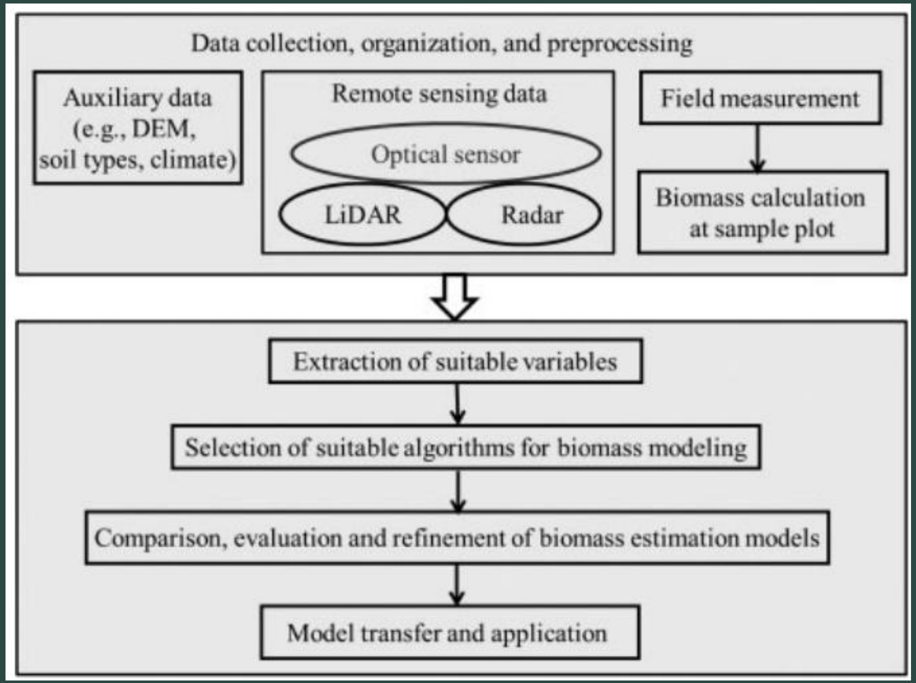
Final map



10 m resolution forest aboveground biomass map of Northeast China for the year 2022

- Strip effect
- Independent evaluation
- Uncertainty map
- Expand space and time dimension
- Other auxiliary data
- Quality of GEDI L4A

Take home message



Common structure of AGB mapping using remote sensing

Footprint tree height → Footprint AGB → Continuous AGB

ALS+plot AGB

Continuous images

Footprint tree height → Continuous tree height → Continuous AGB

Continuous images

Other continuous variables

Two ways of AGB mapping using GEDI

References & Resources

1. Ashton, P. M. S. (2012). Managing forest carbon in a changing climate. In Springer eBooks. https://doi.org/10.1007/978-94-007-2232-3

2. Roy, S., Mudi, S., Das, P., Ghosh, S. M., Shit, P. K., Bhunia, G. S., & Kim, J. (2020). Estimating Above Ground Biomass (AGB) and Tree Density using Sentinel-1 Data. In Environmental science and engineering (pp. 259–280). https://doi.org/10.1007/978-3-030-56542-8_11

3. Tian, L., Wu, X. B., Yu, T., Li, M., Qian, C., Liao, L., & Fu, W. (2023). Review of Remote Sensing-Based Methods for Forest Aboveground Biomass Estimation: Progress, Challenges, and Prospects. Forests, 14(6), 1086. https://doi.org/10.3390/f14061086

4. Liu, Y., Gong, W., Xing, Y., Hu, X., & Gong, J. (2019). Estimation of the forest stand mean height and aboveground biomass in Northeast China using SAR Sentinel-1B, multispectral Sentinel-2A, and DEM imagery. Isprs Journal of Photogrammetry and Remote Sensing, 151, 277–289. https://doi.org/10.1016/j.isprsjprs.2019.03.016

5. Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., & Moran, E. F. (2014). A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. International Journal of Digital Earth, 9(1), 63–105. https://doi.org/10.1080/17538947.2014.990526

6. Dubayah, R., Blair, J. B., Goetz, S. J., Fatoyinbo, L., Hansen, M. C., Healey, S. P., Hofton, M. A., Hurtt, G. C., Kellner, J. R., Luthcke, S. B., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W., & Silva, C. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of Remote Sensing, 1, 100002. https://doi.org/10.1016/j.srs.2020.100002

7. 王成,习晓环,杨学博,聂胜.激光雷达遥感导论[M].北京:高等教育出版社,2022:67-73.

8. Huang, H., Liu, C., Wang, X., Zhou, X., & Gong, P. (2019). Integration of multi-resource remotely sensed data and allometric models for forest aboveground biomass estimation in China. Remote Sensing of Environment, 221, 225–234. https://doi.org/10.1016/j.rse.2018.11.017

9. 邢艳秋等.星载激光雷达反演生物量方法及其应用[M].北京:科学出版社,2016:21.

10. Liu, X., Su, Y., Hu, T., Yang, Q., Liu, B., Deng, Y., Tang, H., Tang, Z., Fang, J., & Guo, Q. (2022). Neural network guided interpolation for mapping canopy height of China's forests by integrating GEDI and ICESat-2 data. Remote Sensing of Environment, 269, 112844. https://doi.org/10.1016/j.rse.2021.112844

11. Yang, Q., Niu, C., Liu, X., Feng, Y., Ma, Q., Wang, X., Tang, H., & Guo, Q. (2023). Mapping high-resolution forest aboveground biomass of China using multisource remote sensing data. Giscience & Remote Sensing, 60(1). https://doi.org/10.1080/15481603.2023.2203303

12. Hancock, S., McGrath, C., Lowe, C., Davenport, I. J., & Woodhouse, I. (2021). Requirements for a global lidar system: spaceborne lidar with wall-to-wall coverage. Royal Society Open Science, 8(12). https://doi.org/10.1098/rsos.211166

References & Resources

13. Kellner, J. R., Armston, J., & Duncanson, L. (2023). Algorithm Theoretical Basis Document for GEDI footprint aboveground biomass density. Earth and Space Science, 10(4). https://doi.org/10.1029/2022ea002516

14. Duncanson, L., Kellner, J. R., Armston, J., Dubayah, R., Minor, D., Hancock, S., Healey, S. P., Patterson, P. L., Saarela, S., Marselis, S., Silva, C. E., Bruening, J. M., Goetz, S. J., Tang, H., Hofton, M. A., Blair, B., Luthcke, S. B., Fatoyinbo, L., Abernethy, K., . . . Zgraggen, C. (2022). Aboveground biomass density models for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission. Remote Sensing of Environment, 270, 112845. https://doi.org/10.1016/j.rse.2021.112845

15. Bullock, E. L., Healey, S. P., Yang, Z., Acosta, R. P., Villalba, H., Insfrán, K. P., Melo, J., Wilson, S., Duncanson, L. I., Næsset, E., Armston, J., Saarela, S., Ståhl, G., Patterson, P. L., & Dubayah, R. (2023). Estimating aboveground biomass density using hybrid statistical inference with GEDI lidar data and Paraguay's national forest inventory. Environmental Research Letters, 18(8), 085001. https://doi.org/10.1088/1748-9326/acdf03

16. Pascual, A., Guerra-Hernández, J., Armston, J., Minor, D., Duncanson, L., May, P., Kellner, J. R., & Dubayah, R. (2023). Assessing the performance of NASA's GEDI L4A footprint aboveground biomass density models using National Forest Inventory and airborne laser scanning data in Mediterranean forest ecosystems. Forest Ecology and Management, 538, 120975. https://doi.org/10.1016/j.foreco.2023.120975

17. Shendryk, Y. (2022). Fusing GEDI with earth observation data for large area aboveground biomass mapping. International Journal of Applied Earth Observation and Geoinformation, 115, 103108. https://doi.org/10.1016/j.jag.2022.103108

18. Kacic, P., Thonfeld, F., Geßner, U., & Kuenzer, C. (2023). Forest Structure Characterization in Germany: novel products and analysis based on GEDI, Sentinel-1 and Sentinel-2 data. Remote Sensing, 15(8), 1969. https://doi.org/10.3390/rs15081969

19. Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2023). A high-resolution canopy height model of the Earth. Nature Ecology and Evolution, 7(11), 1778–1789. https://doi.org/10.1038/s41559-023-02206-6

Data & Codes:

https://github.com/ChguFan/2023_ML_EES/tree/main/FinalProject

谢谢大家！请批评指正

汇报人：范崇光　　日期：2023年12月15日