Machine learning for forest structure assessments

Jan Dirk Wegner
EcoVision Lab
Universität Zürich & ETH Zürich

Vegetation analysis at global scale

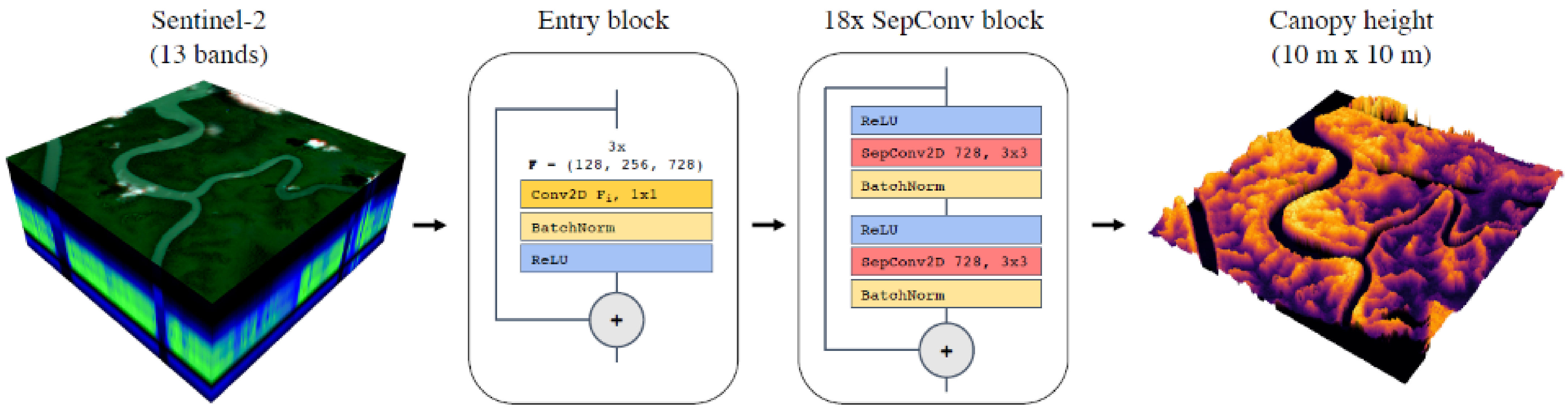
- **Goal**: Dense, near-realtime forest structure assessments at global scale with 10-20m ground sampling distance

- Monitor vegetation parameters to have accurate, up-to-date input for climate and biodiversity modelling

- **Idea**: use *single satellite images* to predict vegetation height (and later more variables like biomass)

Method: network architecture

- Avoid down- or up-sampling: stride 1, no max-pooling
- 18 identical, separable convolution (SepConv) blocks do not only learn spectral features that correlate with canopy height, but also spatial context and texture features.



Lang, N., Schindler, K., Wegner, J.D.: Country-wide high-resolution vegetation height mapping with Sentinel-2, Remote Sensing of Environment, 2019, vol. 233, article 111347.

Method: loss function

- Euclidean loss function for regression of continuous height values
- L2-penalty term on model parameters («weight decay») as regularizer

ground truth at pixel *i* **model parameters**

$$Loss = \frac{1}{N}\sum_{i=1}^{N}(f(x_i) - y_i)^2 + \lambda\frac{1}{W}\sum_{j=1}^{W}(w_j)^2$$

number of samples (pixels)

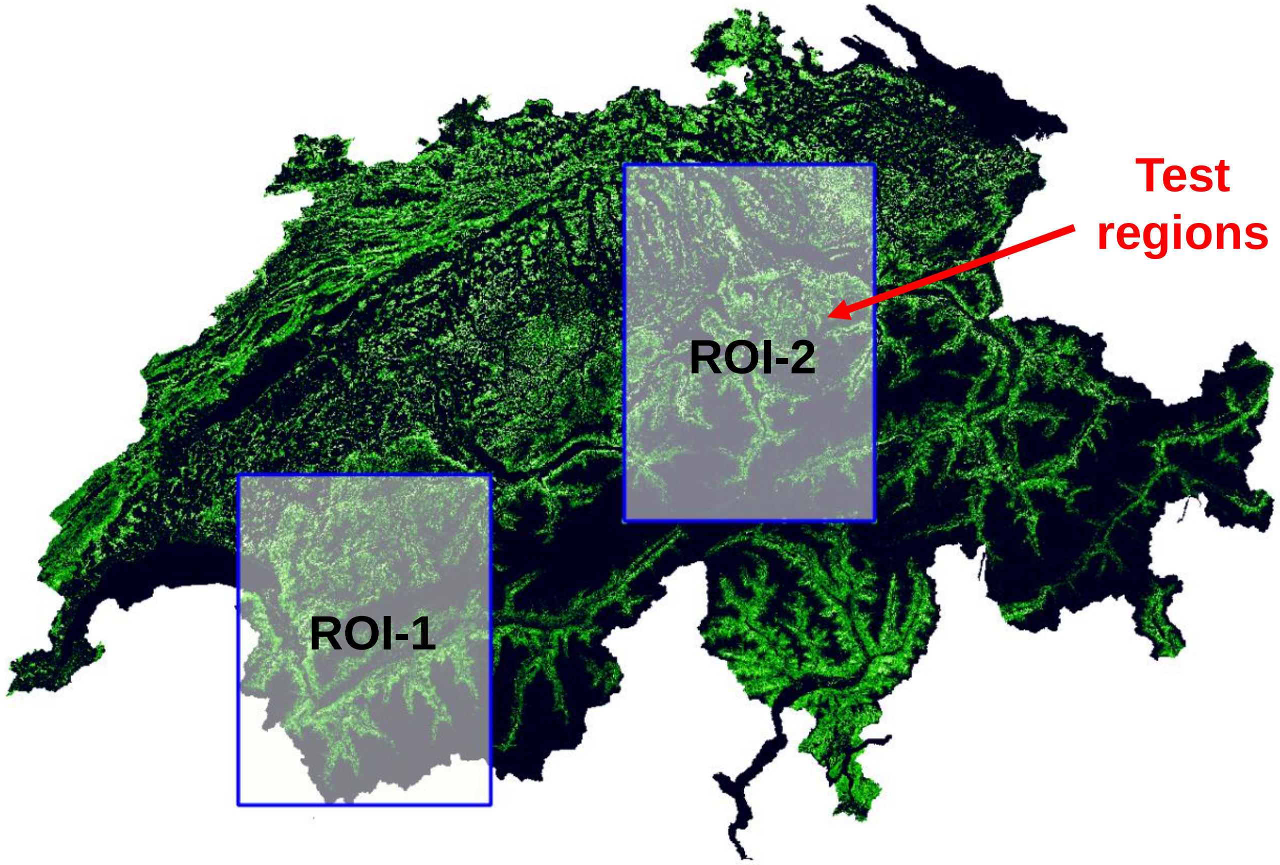
pixel intensities

Vegetation height predictions at pixel *i*

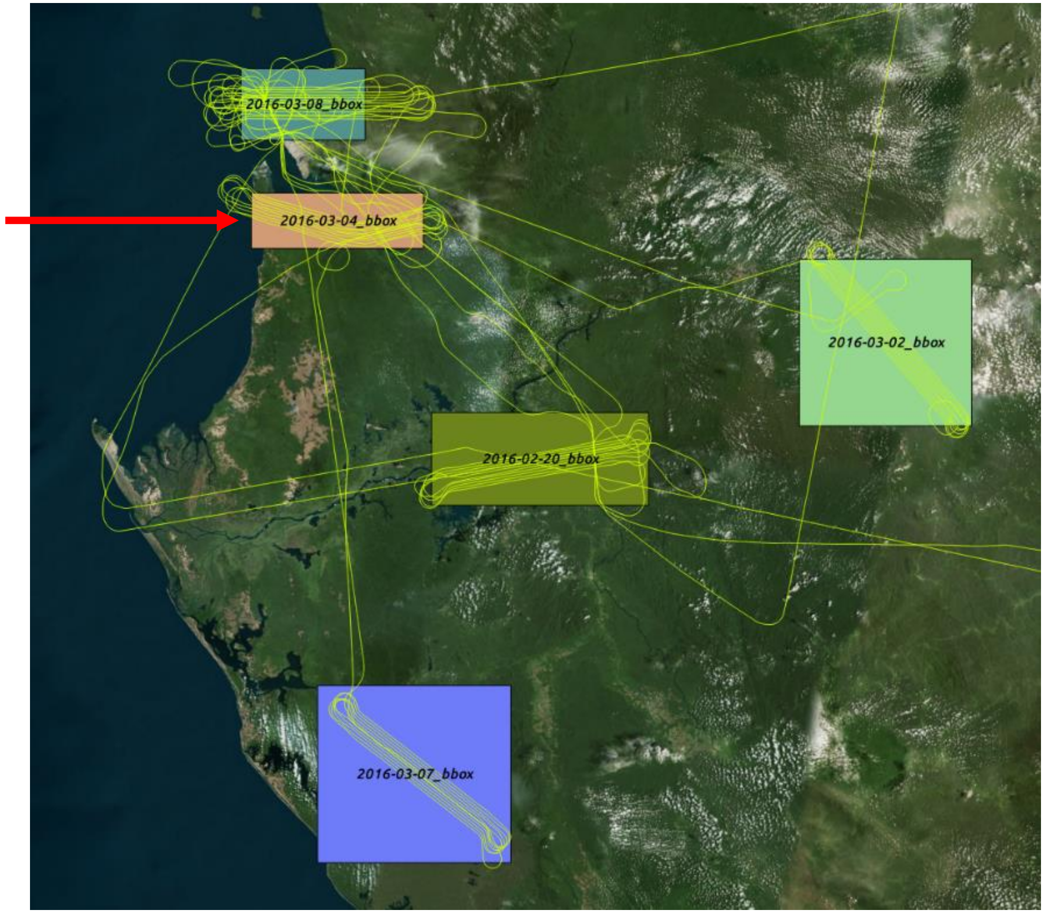
number of model parameters

Experiments

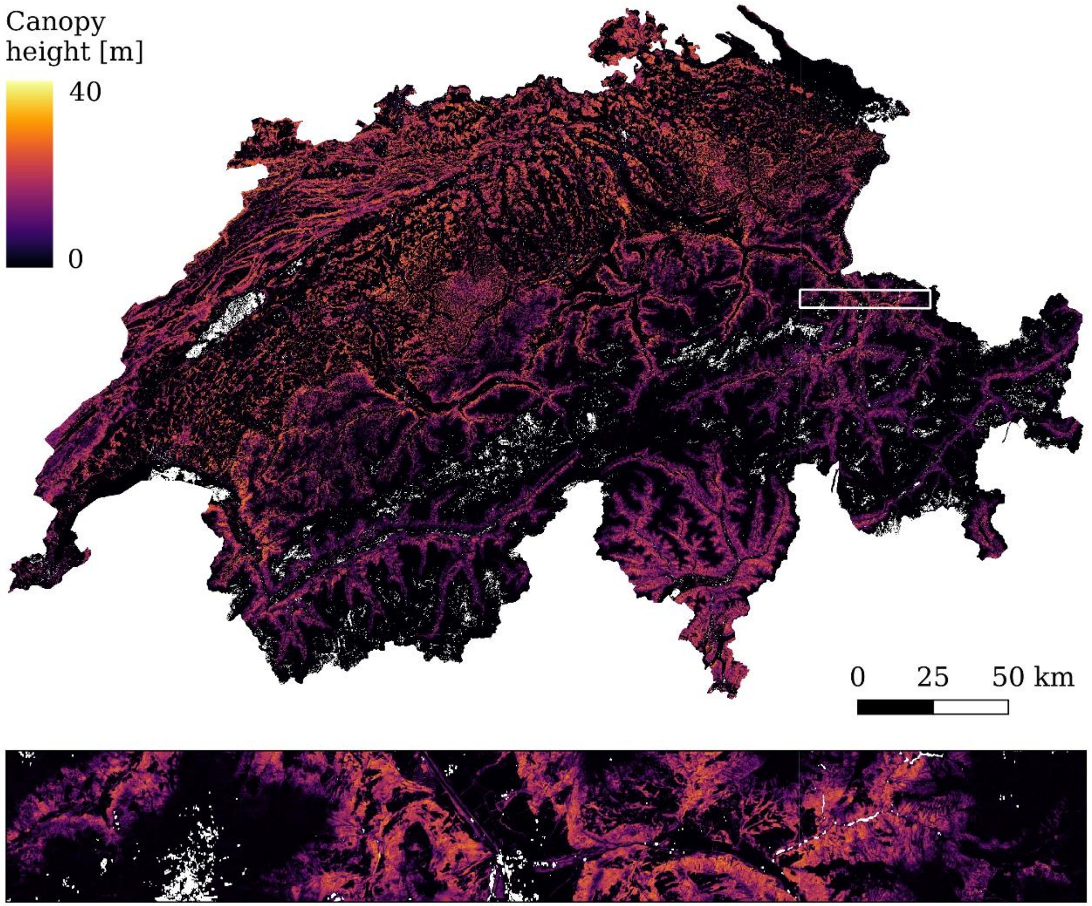
Switzerland



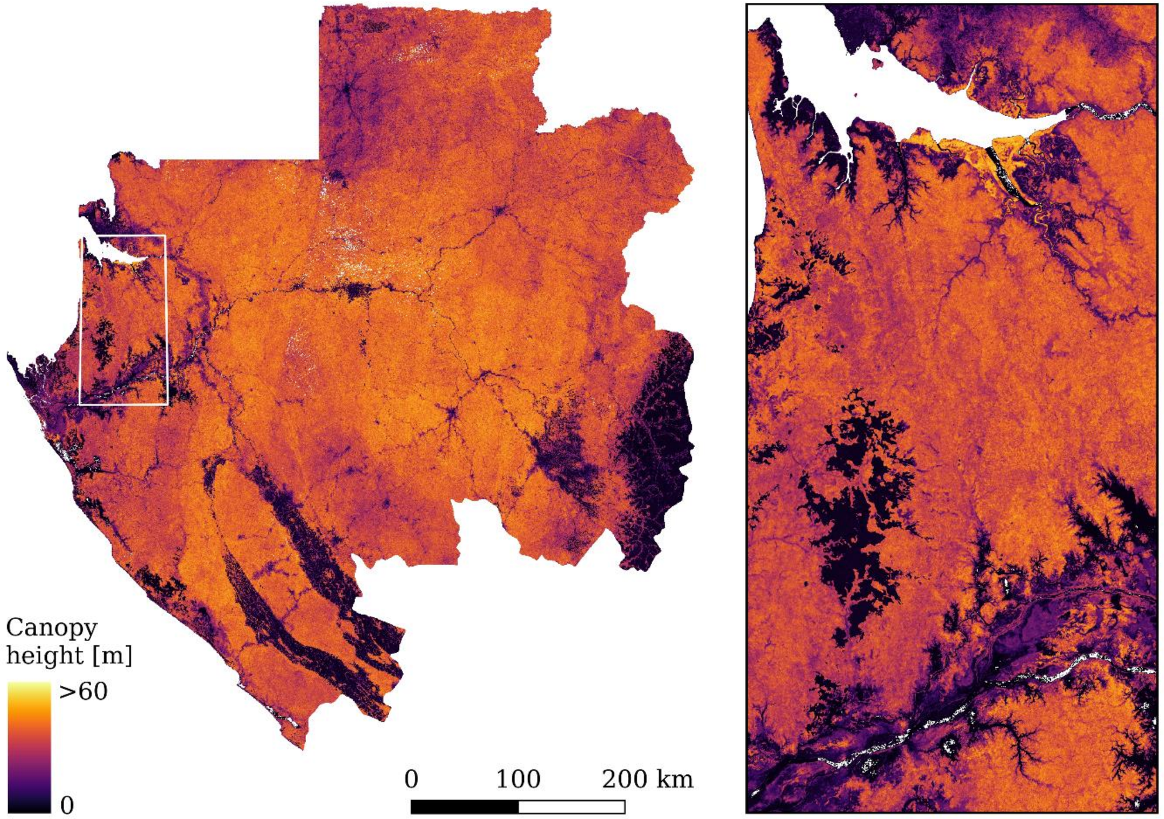
Gabon



Results at 10m ground sampling distance



*MAE **±1.7 m** (for vegetation heights 0 to 40m)*



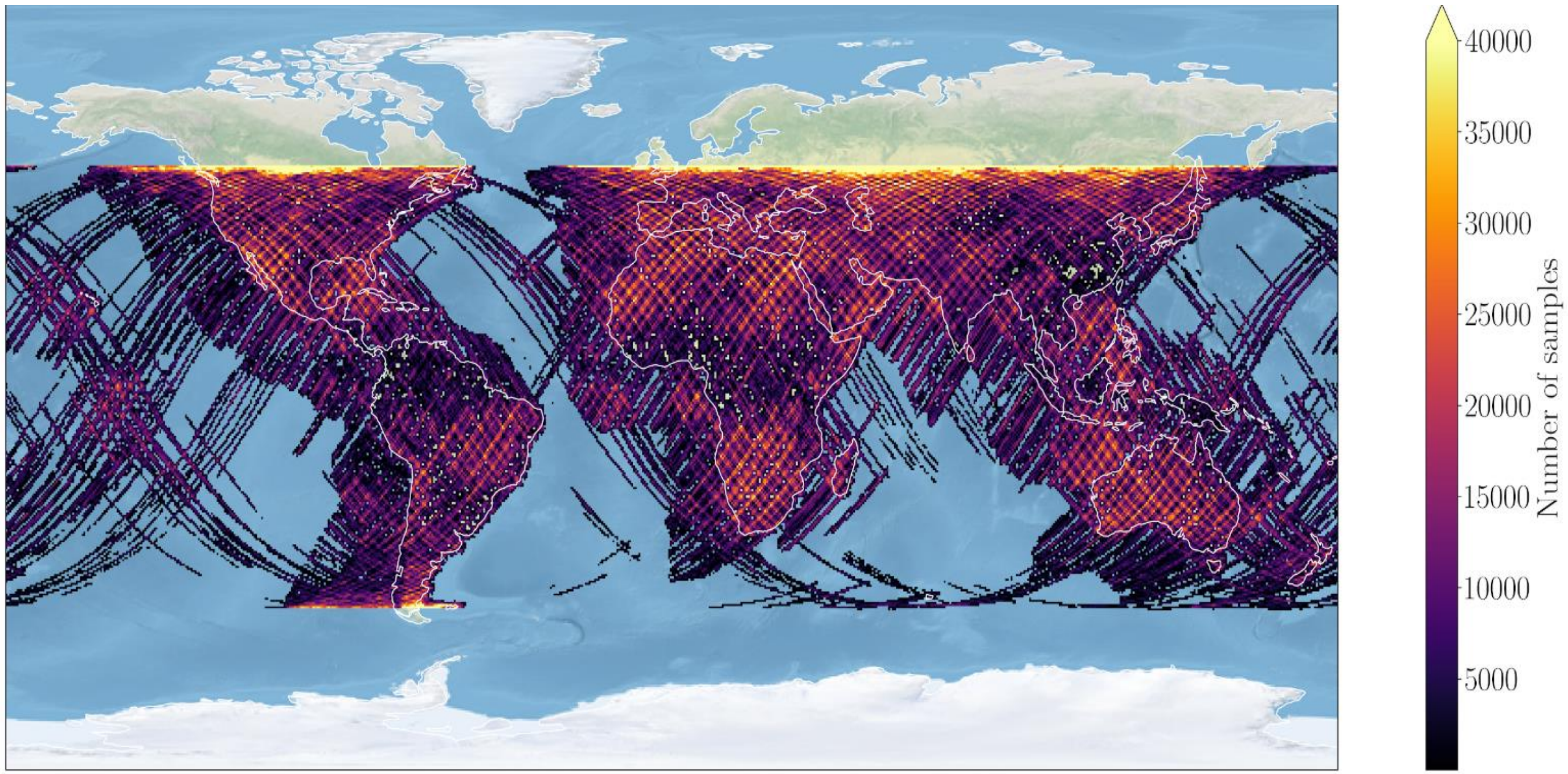
*MAE **±4.3 m** (for vegetation heights 0 to 60m)*

Universität Zürich[UZH] ETH zürich

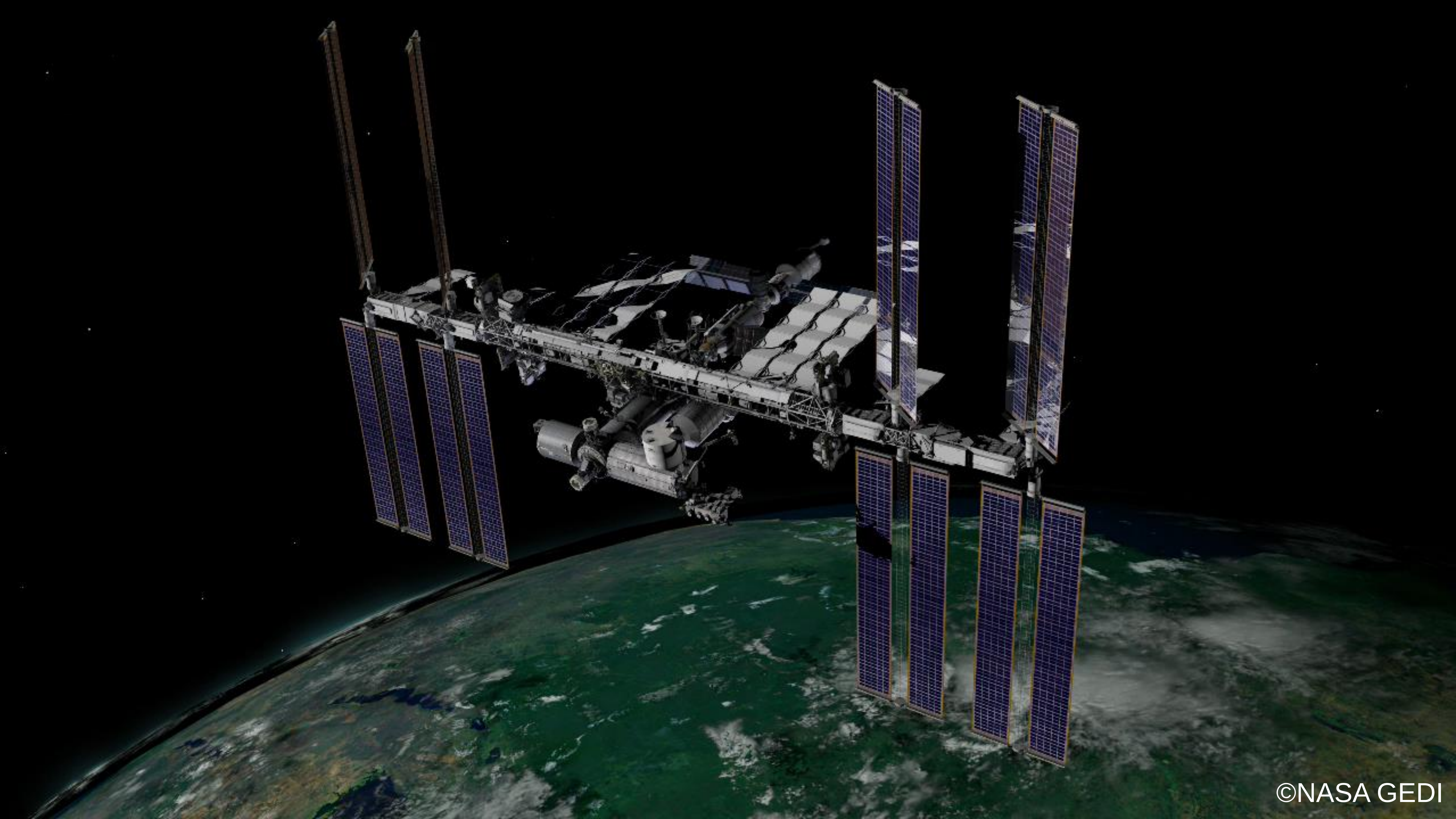
EcoVision

Ongoing work: scale globally

- **Goal**: global vegetation height, biomass and HCS map with 20 meter resolution and near-realtime updates
- Collaboration with **NASA GEDI** team und **Amazon Research**
- **Idea:** train deep learning model on *full-waveform spaceborne Laserscanning* points of NASA GEDI mission und interpolate with satellite data

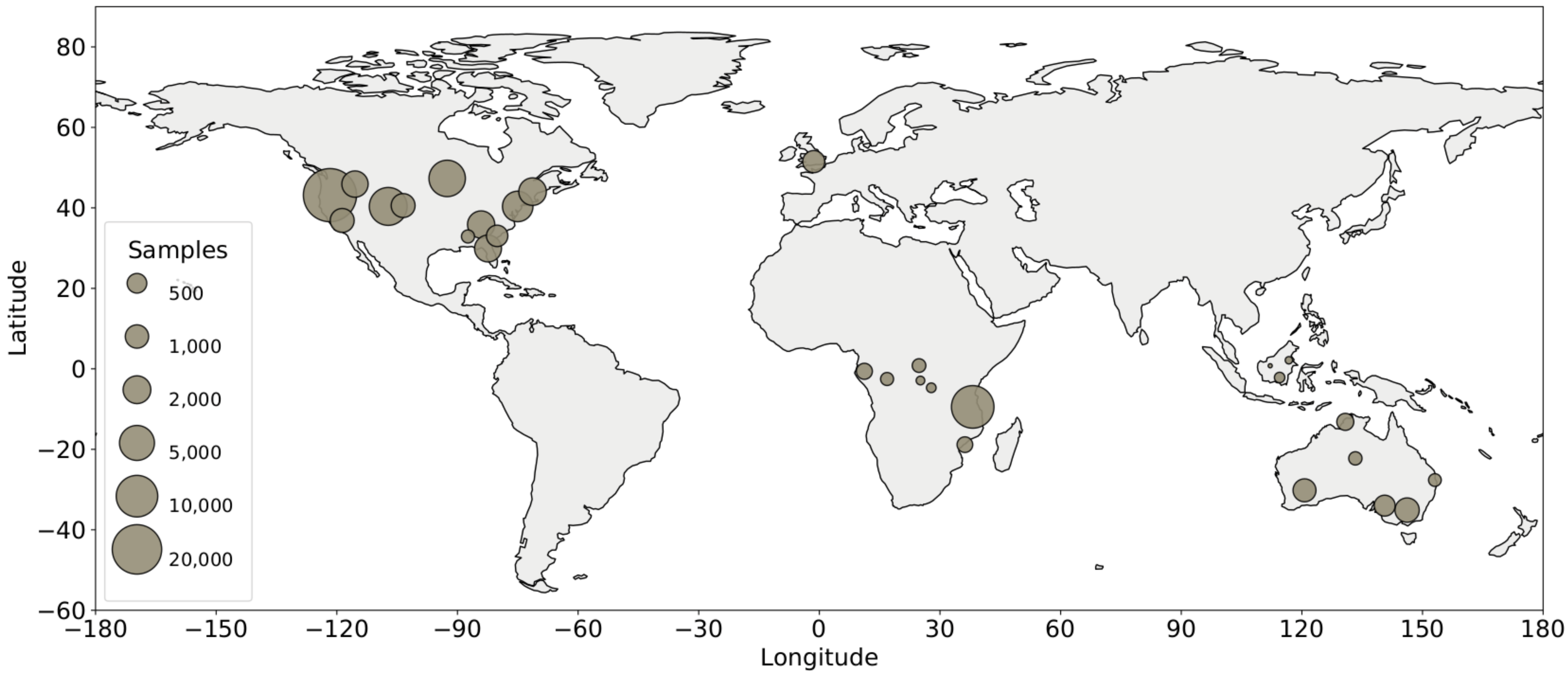


Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., Wegner, J.D.: Global canopy height estimation with GEDI LIDAR waveforms and Bayesian deep learning, 2021, under review.


©NASA GEDI

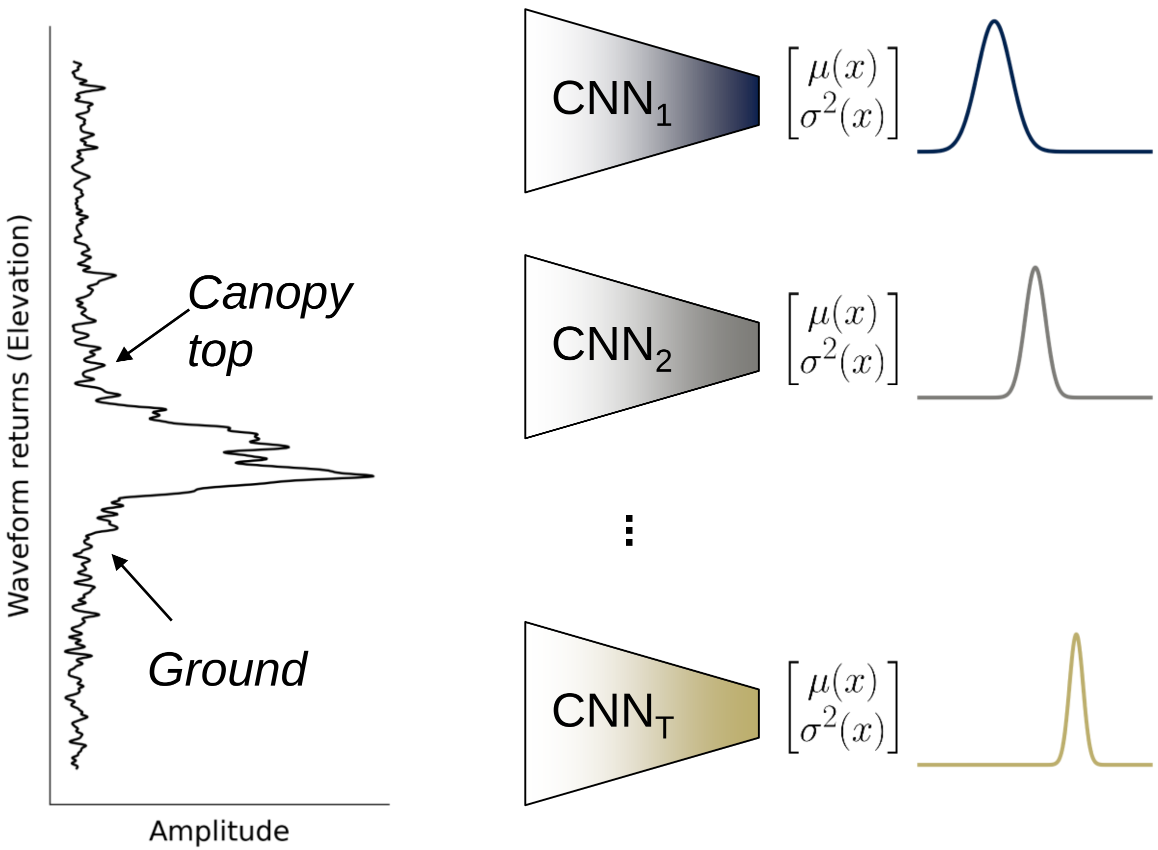
Reference sites

- Train, validate, test on reference sites collected by NASA GEDI team
- Full waveform airborne LiDAR used to simulate GEDI full waveform data
- Co-registration: waveform matching to align real GEDI waveforms with simulated waveforms



Calibration of GEDI data

Input waveform **Individual models**



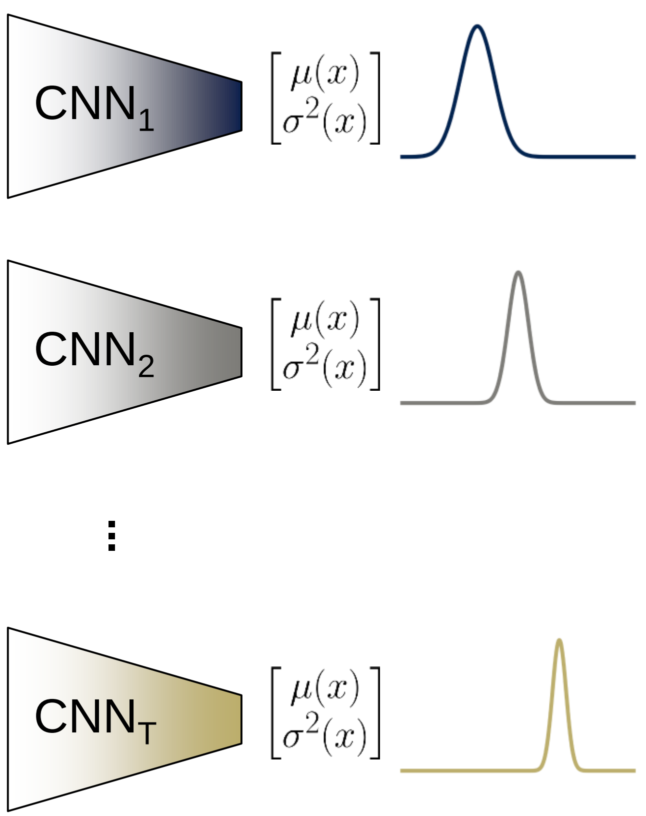
- Each CNN model is trained separately starting from random initializations
- Two outputs per model to approximate the conditional distribution p(y|x)
- Minimize the Gaussian negative log likelihood
- Optimize CNN parameters with stochastic gradient descent (SGD)

Training loss function

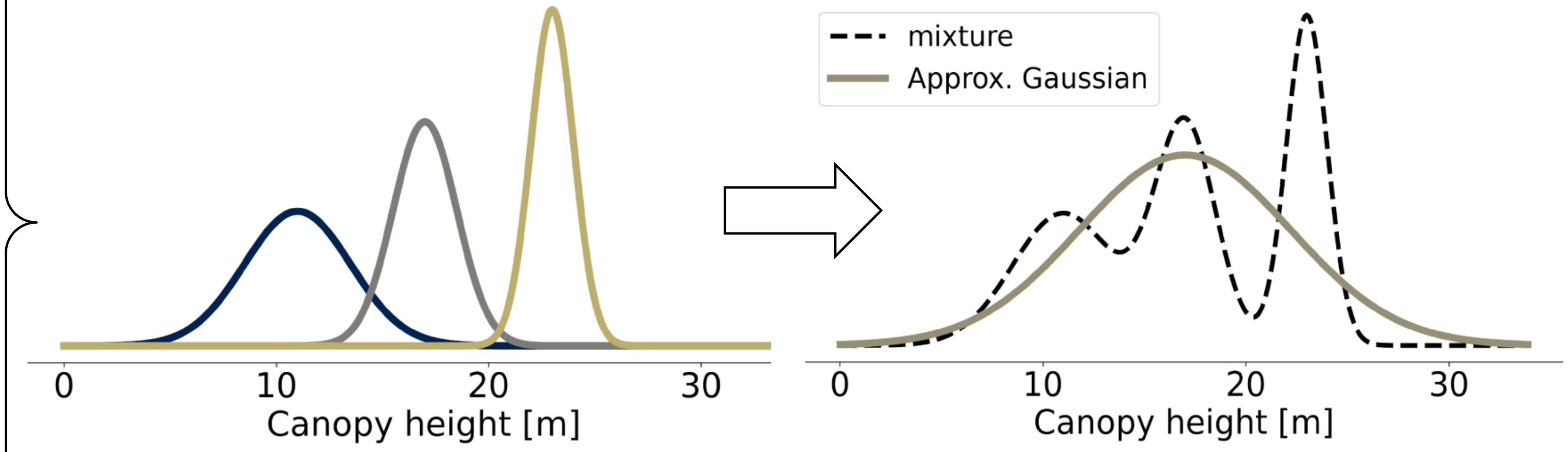
$$\mathcal{L}_{NLL} = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{2\sigma(x_i)^2}\left(\mu(x_i) - y_i\right)^2 + \frac{1}{2}\log \sigma(x_i)^2$$

Calibration of GEDI data

Individual models



Ensemble averaging



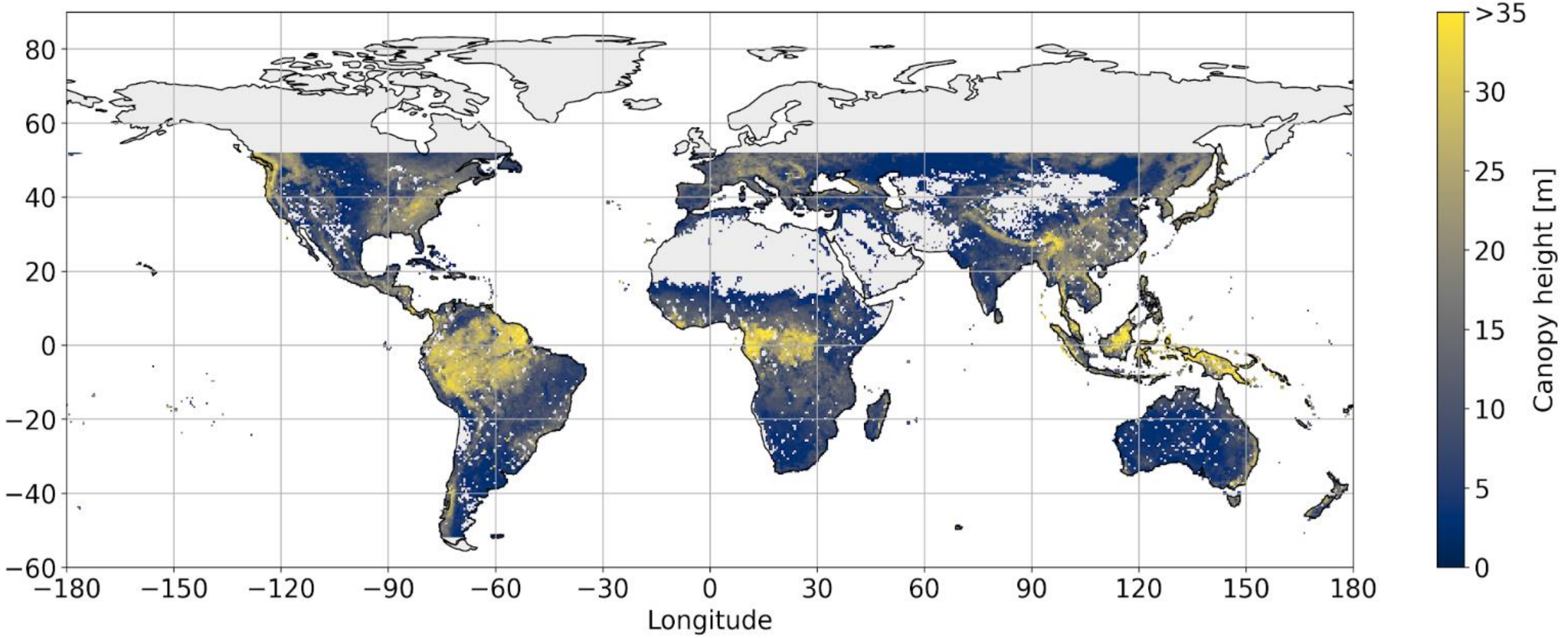
$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T}\hat{\mu}_t$$

$$Var(\hat{y}) = \underbrace{\frac{1}{T}\sum_{t=1}^{T}\hat{\mu}_t^{\,2} - \left(\frac{1}{T}\sum_{t=1}^{T}\hat{\mu}_t\right)^2}_{\text{epistemic}} + \underbrace{\frac{1}{T}\sum_{t=1}^{T}\hat{\sigma}_t^{\,2}}_{\text{aleatoric}}$$

epistemic
(model uncertainty)
"knowledge"

aleatoric
(data uncertainty)
"rolling a dice"

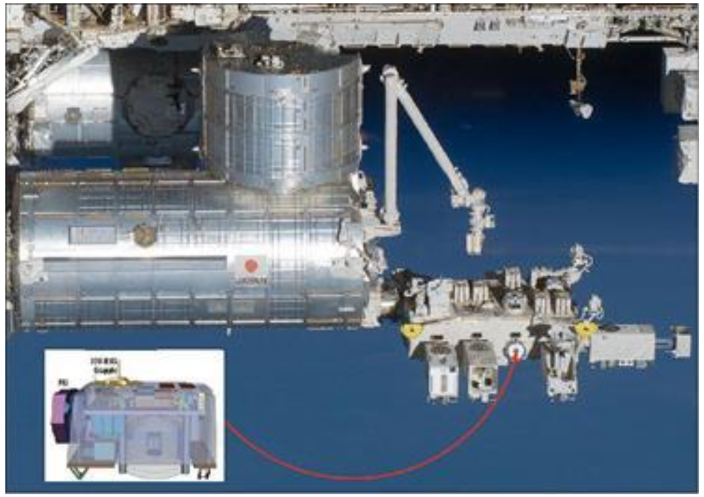
Results: global canopy height with 2.7 m RMSE

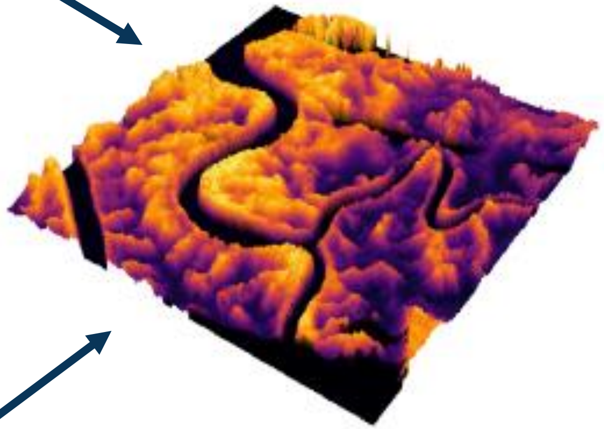


waveforms in non-vegetated areas are filtered out based on MODIS Vegetation Continuous Fields (MOD44B), waveforms filtered based on predictive uncertainty according to the 70% recall setting (i.e., 30% with highest uncertainty filtered out) and values below 0m height suppressed

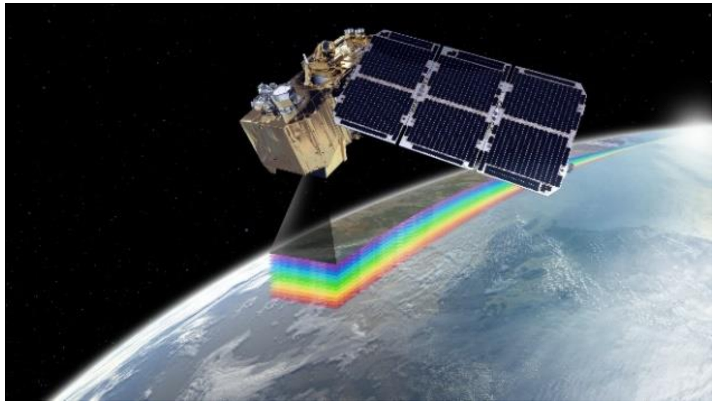
Combination of sparse LiDAR footprints and dense Sentinel-2 predictions

GEDI





Sentinel-2



Exciting future directions

Nico Lang

- ✓ *Add IceSat-2 for polar regions*
- ✓ *Investigate use of SAR data*
- ✓ *geometric deep learning for non-grid structured GEDI data*

Thanks to a great team @ EcoVision Lab !