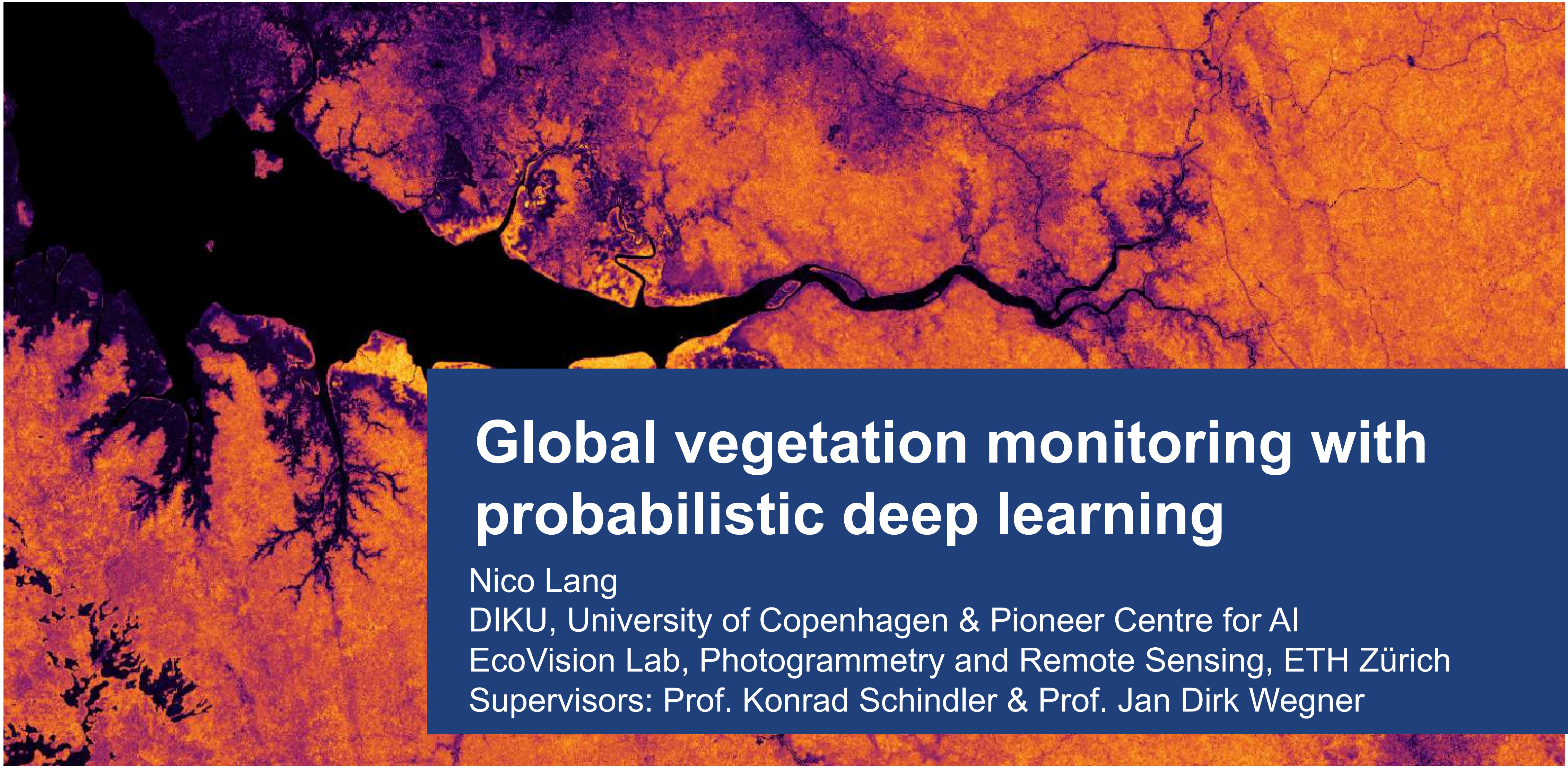
ETH zürich

EcoVision Lab,
Photogrammetry and Remote Sensing



Global vegetation monitoring with probabilistic deep learning

Nico Lang
DIKU, University of Copenhagen & Pioneer Centre for AI
EcoVision Lab, Photogrammetry and Remote Sensing, ETH Zürich
Supervisors: Prof. Konrad Schindler & Prof. Jan Dirk Wegner

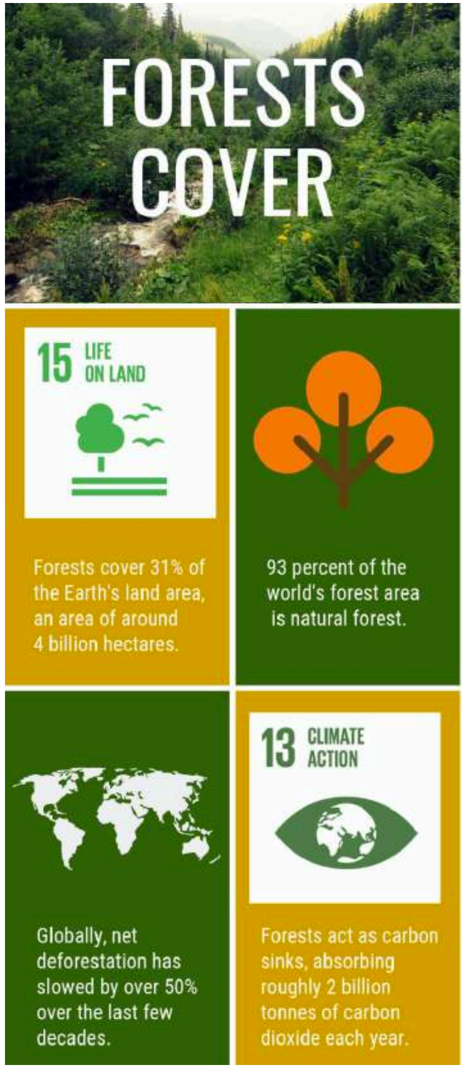
UNIVERSITY OF COPENHAGEN

PIONEER CENTRE FOR ARTIFICIAL INTELLIGENCE

PRS Photogrammetry Remote Sensing

EcoVision

United Nations Strategic Plan for Forests 2017-2030



un.org

Global Forest Goal 1

- Increase forest area by 3 percent worldwide
- Maintain or enhance forest carbon stocks

Global challenges

→ Climate action

→ Biodiversity conservation

→ *We need **global, high-resolution**, and **fine-grained data of vegetation properties** to monitor these goals.*

→ *Vegetation canopy height is a key indicator.*

Huge 750 years old sequoia tree, California. Photo by: Michael Nichols

How can we ~~measure~~ estimate vegetation height?

Deep Learning is a promising approach to:

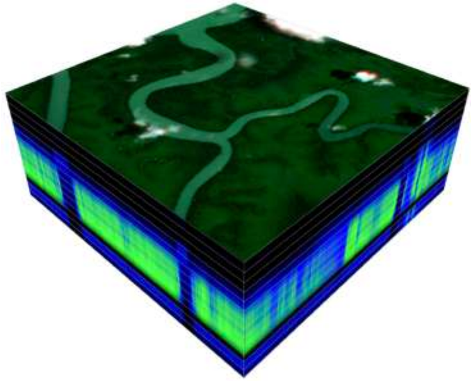
1. Exploit the large amount of data
2. Overcome lack of understanding

high

Scalability

low

Optical satellite images



+ global coverage
+ high resolution
- not designed for height

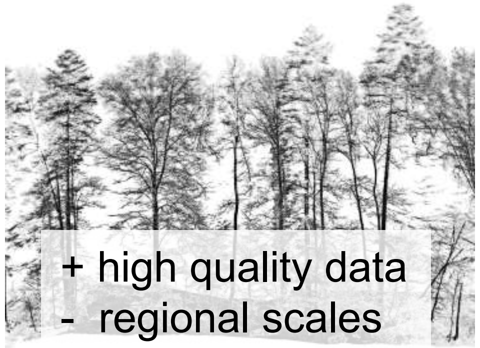
?

Spaceborne LIDAR

+ global distribution
- noise in the data

?

Airborne LIDAR / Photogrammetry



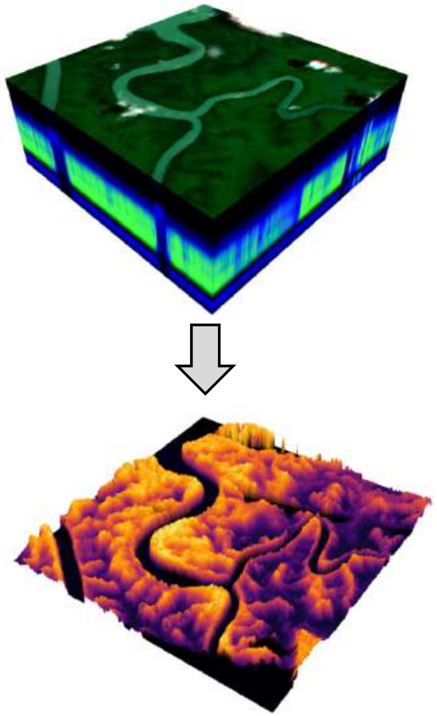
+ high quality data
- regional scales

low *high*

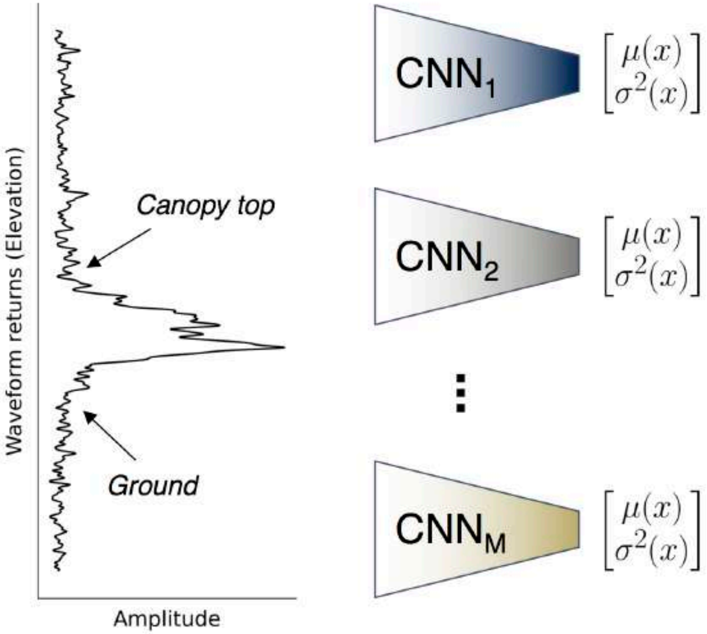
Suitability to directly measure vegetation structure

Overview

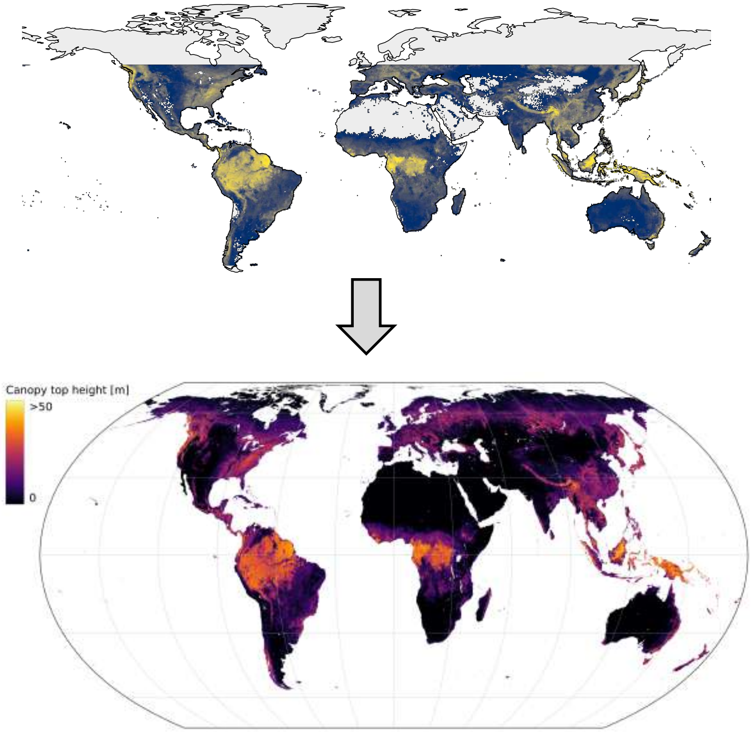
(1) Canopy height estimation from Sentinel-2 optical images



(2) Global canopy height regression and uncertainty estimation from GEDI spaceborne LIDAR

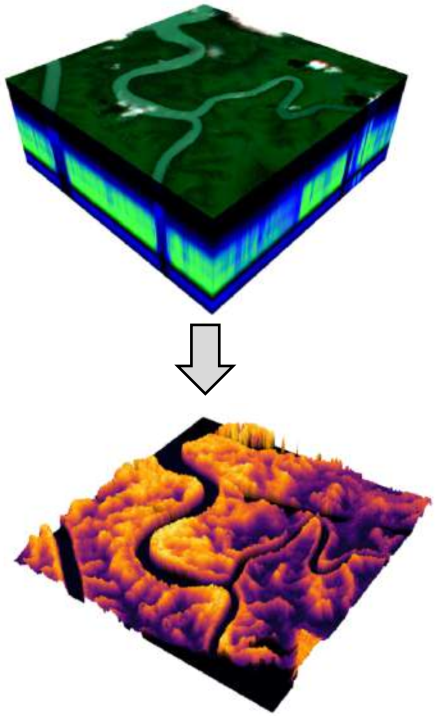


(3) A high-resolution canopy height model of the Earth

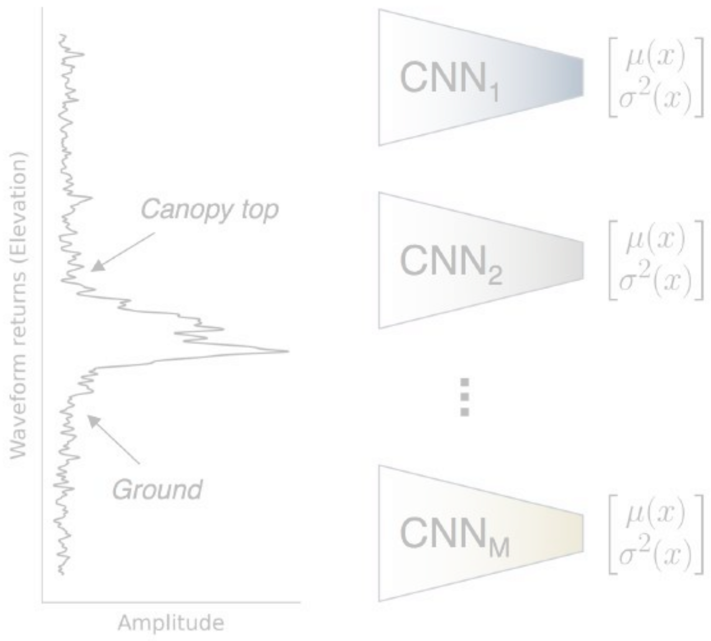


Overview

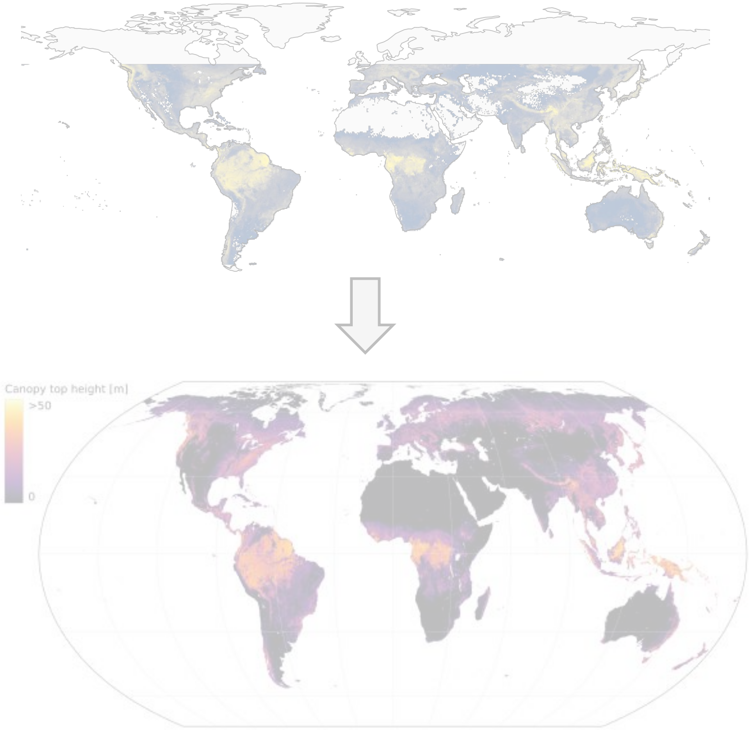
(1) Canopy height estimation from Sentinel-2 optical images



(2) Global canopy height regression and uncertainty estimation from GEDI spaceborne LIDAR

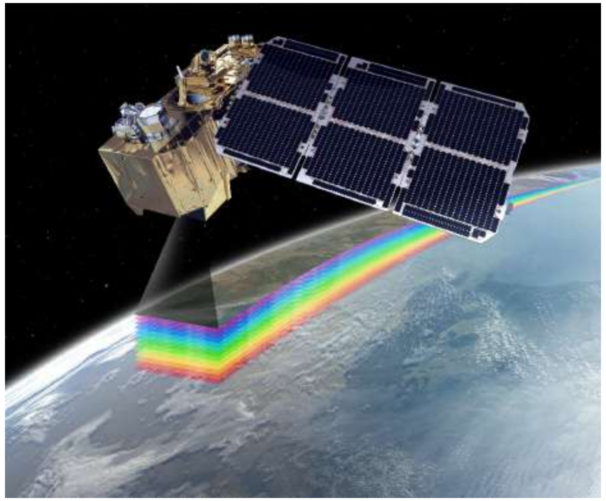


(3) A high-resolution canopy height model of the Earth



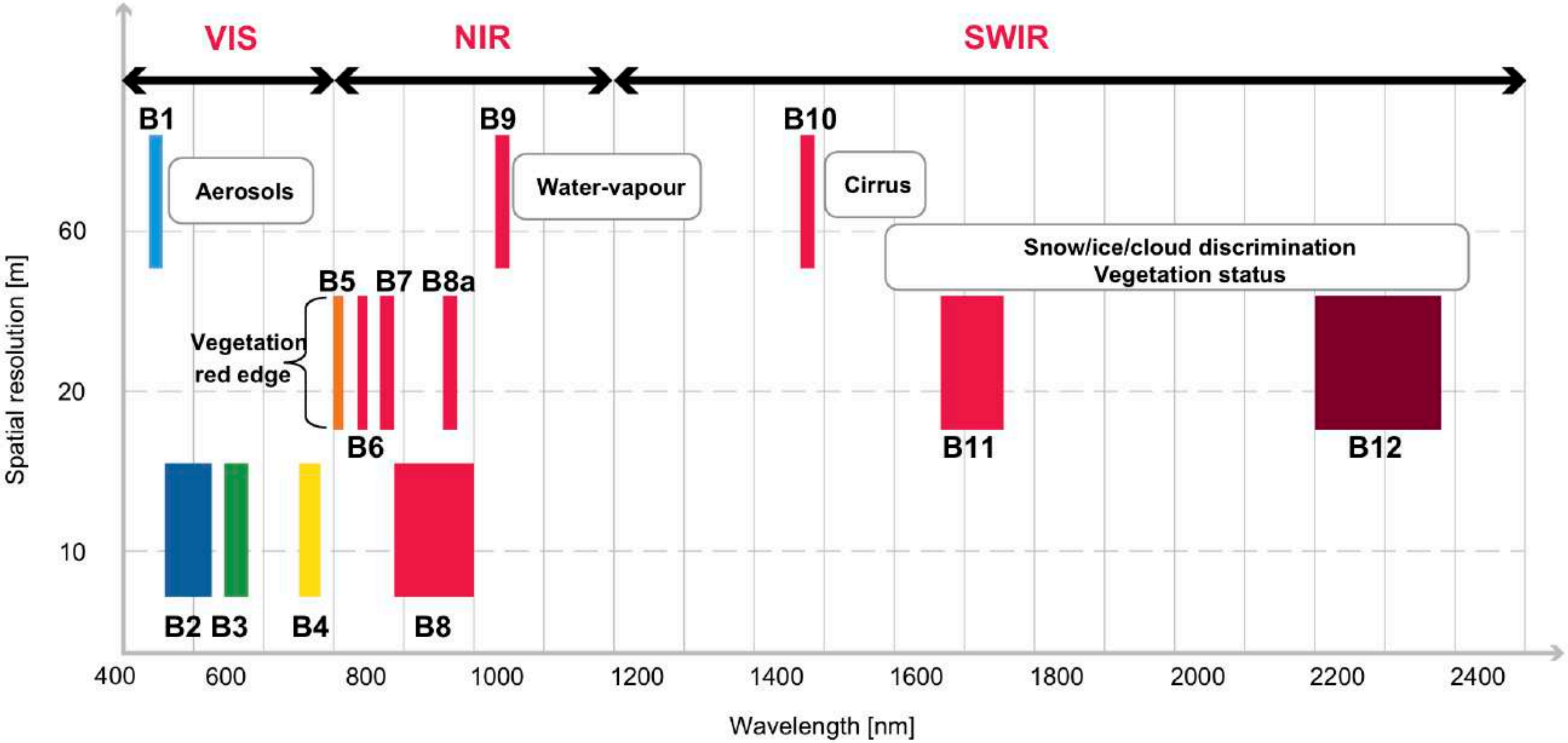
Lang, N., Schindler, K., & Wegner, J. D. (2019). Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sensing of Environment*, *233*, 111347.

(1) Copernicus Sentinel-2 optical satellite images



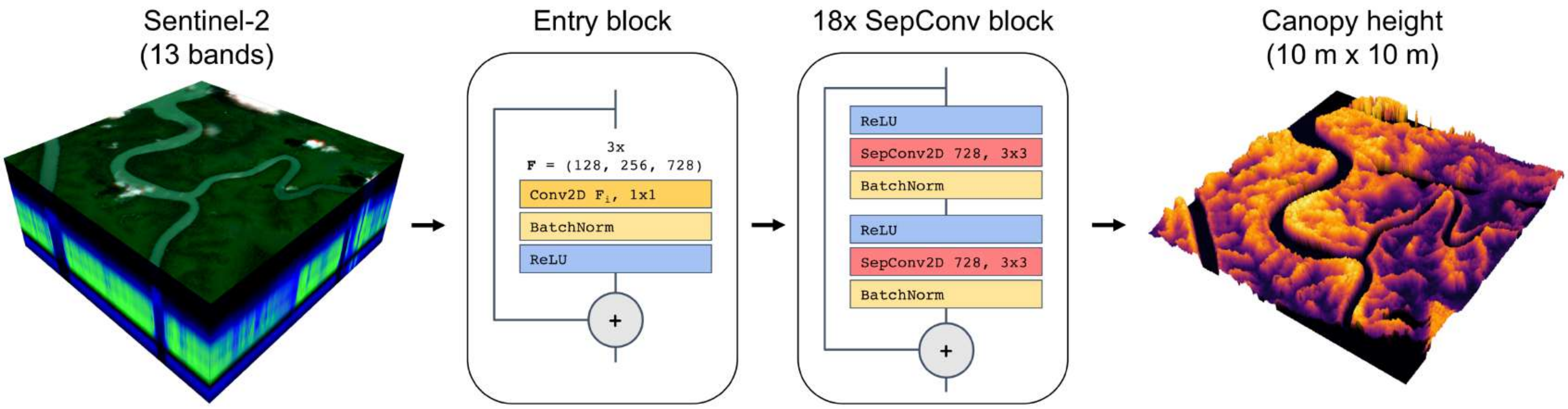
airbus.com

- Multispectral sensor
- 4 bands with 10 m GSD
- Global coverage
- New image every 3-5 days



European Space Agency bulletin (2015)

(1) Data-driven approach to estimate canopy height from Sentinel-2



Learn this mapping from reference data in a supervised fashion

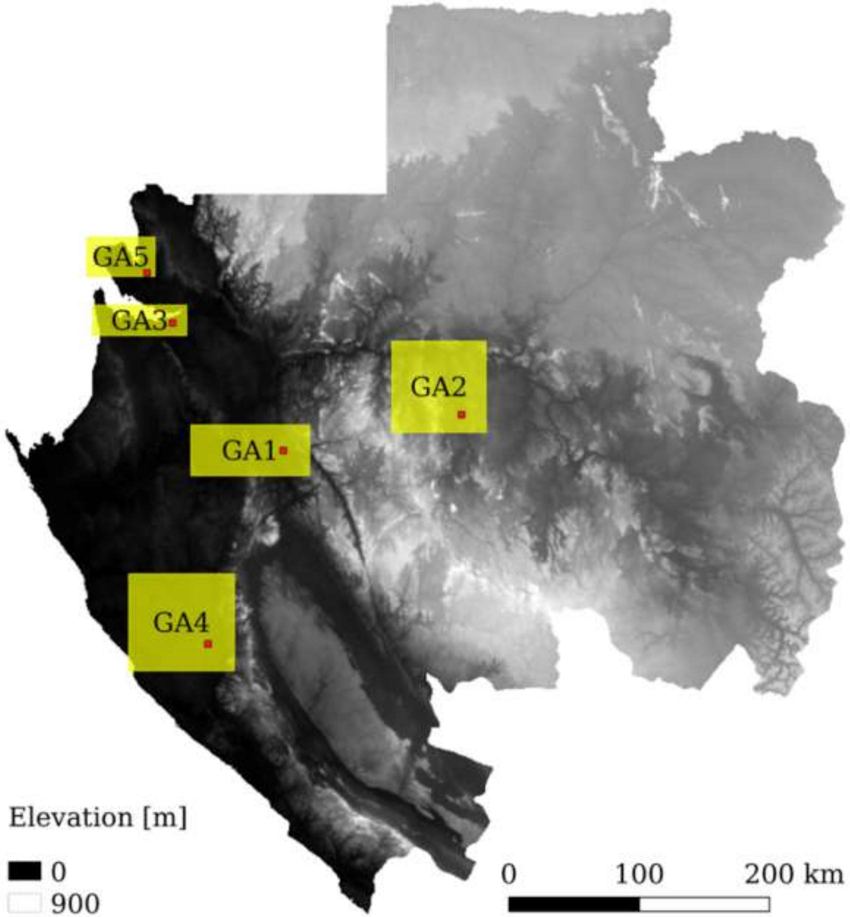
$$Loss = \frac{1}{N}\sum_{i=1}^{N}(\underbrace{f(x_i)}_{\text{Model output at pixel } i} - \underbrace{y_i}_{\text{Reference canopy height}})^2$$

Model output at pixel i *Reference canopy height*

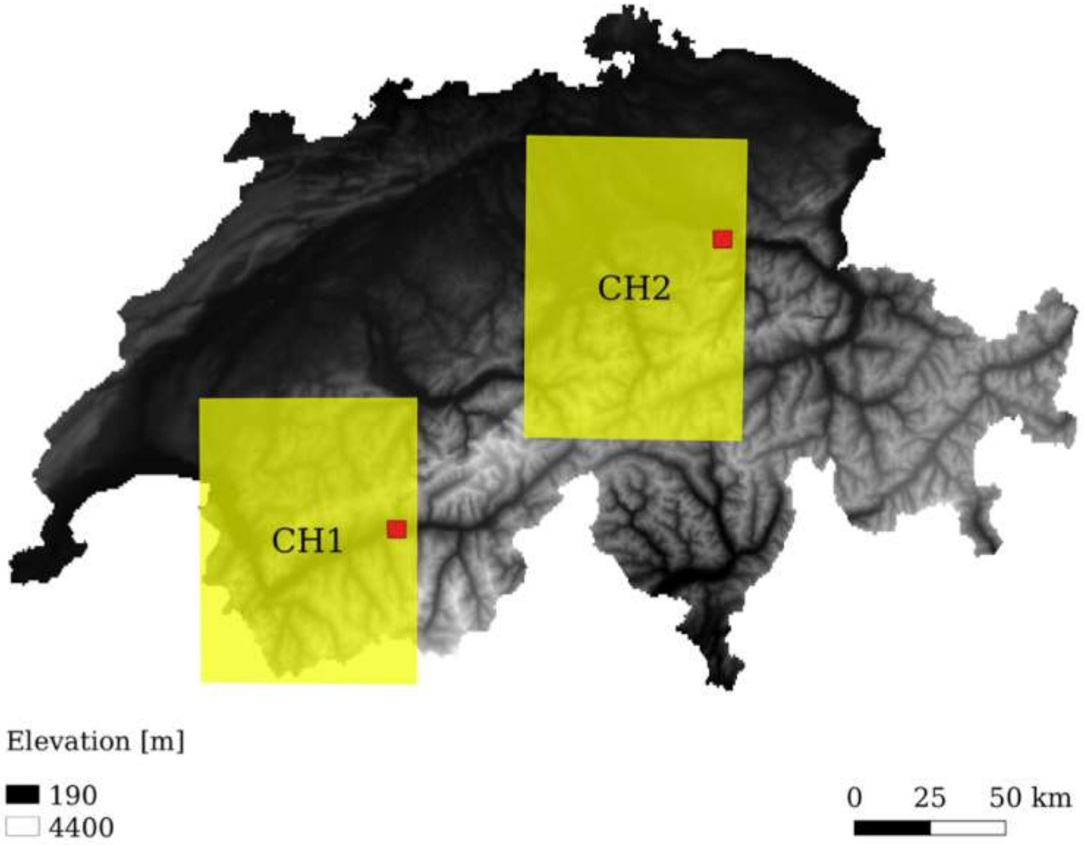
Lang, N., Schindler, K., & Wegner, J. D. (2019). Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sensing of Environment*, *233*, 111347.

(1) Reference data from airborne LIDAR / Photogrammetry

Gabon



Switzerland



Lang, N., Schindler, K., & Wegner, J. D. (2019). Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sensing of Environment*, *233*, 111347.

(1) Ablation studies using airborne LIDAR reference data

Mean absolute error (MAE) [m]

Name	overall	0-10 [m]	10-20 [m]	20-30 [m]	30-40 [m]	40-50 [m]	50-60 [m]	60-70 [m]
ALL	**3.7**	2.6	3.7	**4.4**	**3.8**	**4.5**	5.9	7.5
RGB	5.0	2.7	5.4	6.9	5.9	5.1	7.2	10.3
N	6.0	2.7	5.4	8.1	7.6	8.7	8.3	8.6
RGBN	3.8	**1.8**	**3.4**	5.1	4.8	5.3	**5.8**	**6.9**
woRGBN	4.8	2.0	4.7	5.8	5.2	7.6	11.2	14.5
ALL 1×1	6.0	1.8	5.1	7.7	6.6	11.3	17.1	22.0

Different input band combinations

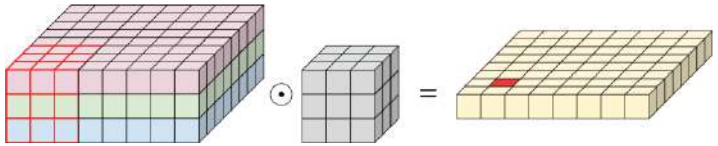
No texture features

Table 4: Ablation study Gabon (GA3): MAE for various band selections, and for strictly pixel-wise spectral features (ALL 1×1).

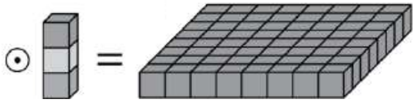
→ High spatial resolution bands RGBN (10 m) are most important

→ Texture features improve overall performance and reduce the error for tall canopies

Texture features (3x3 kernels)



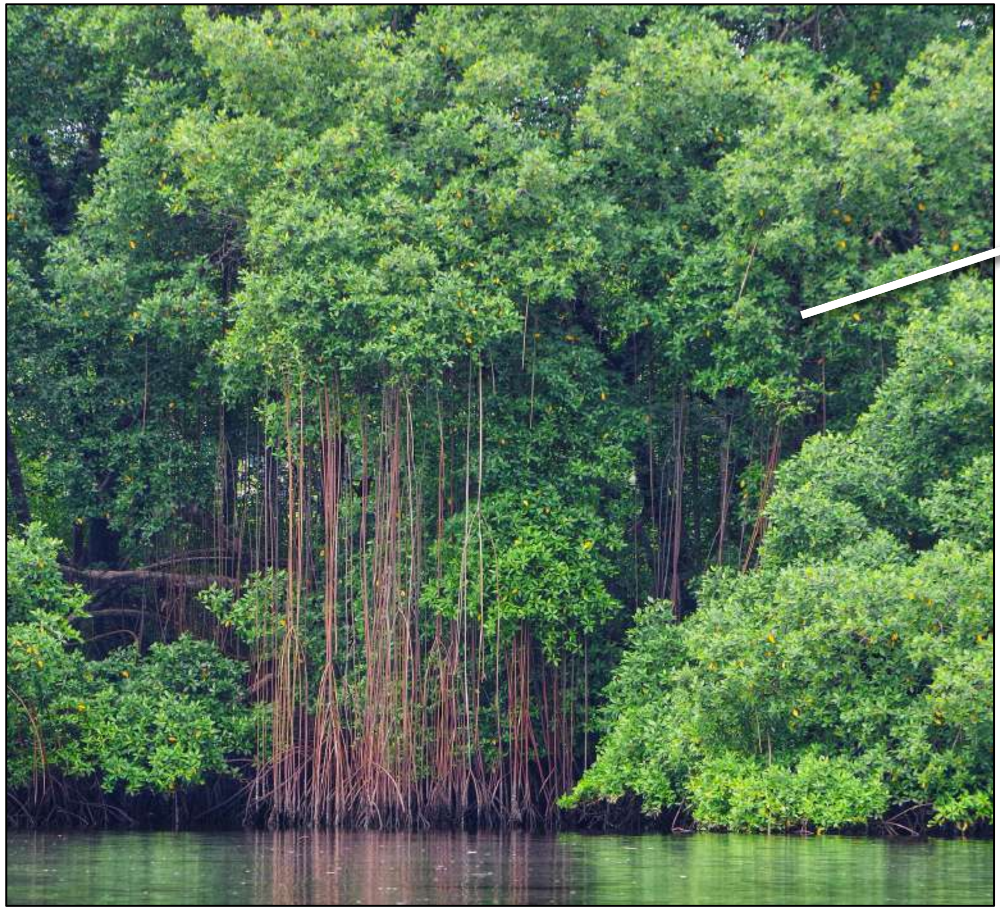
No texture features (1x1 kernels)



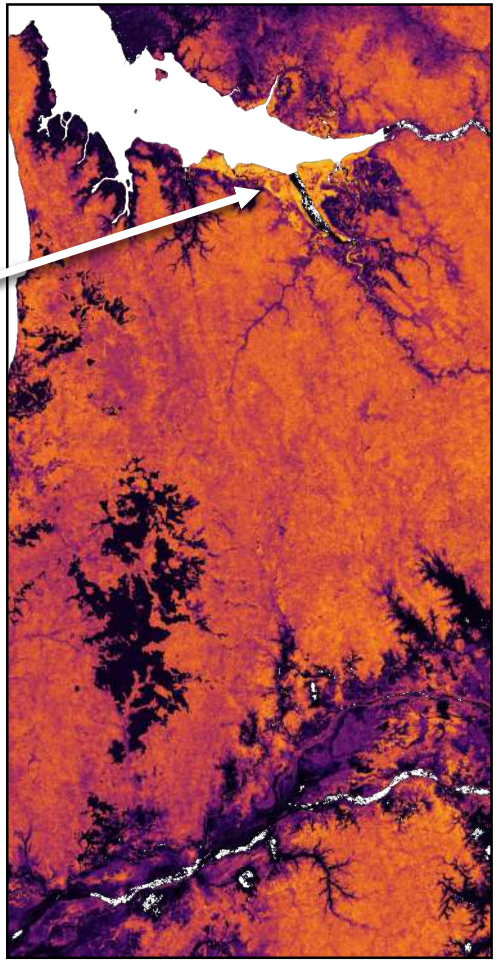
Lang, N., Schindler, K., & Wegner, J. D. (2019). Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sensing of Environment*, *233*, 111347.

(1) Country-wide map for Gabon

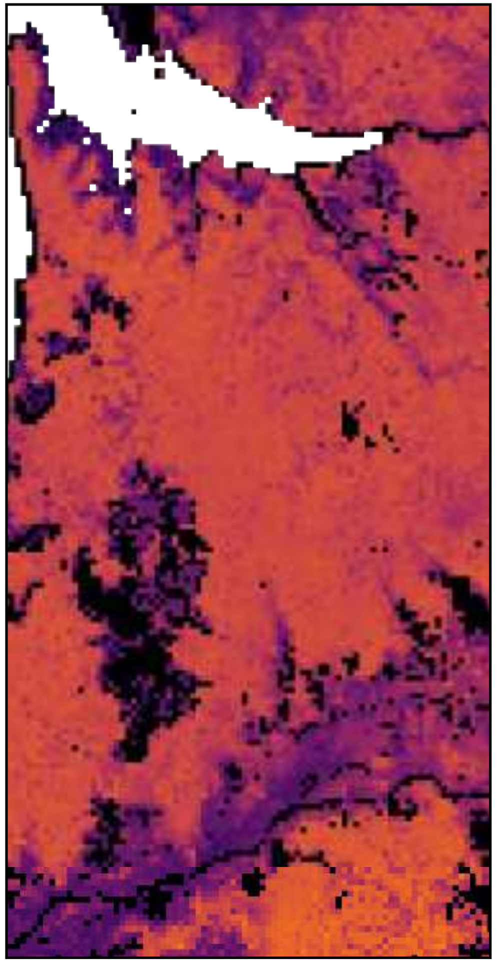
Pongara National Park: tallest mangroves in the world



lifeblink.com



Lang et al. (2019), 10 m GSD

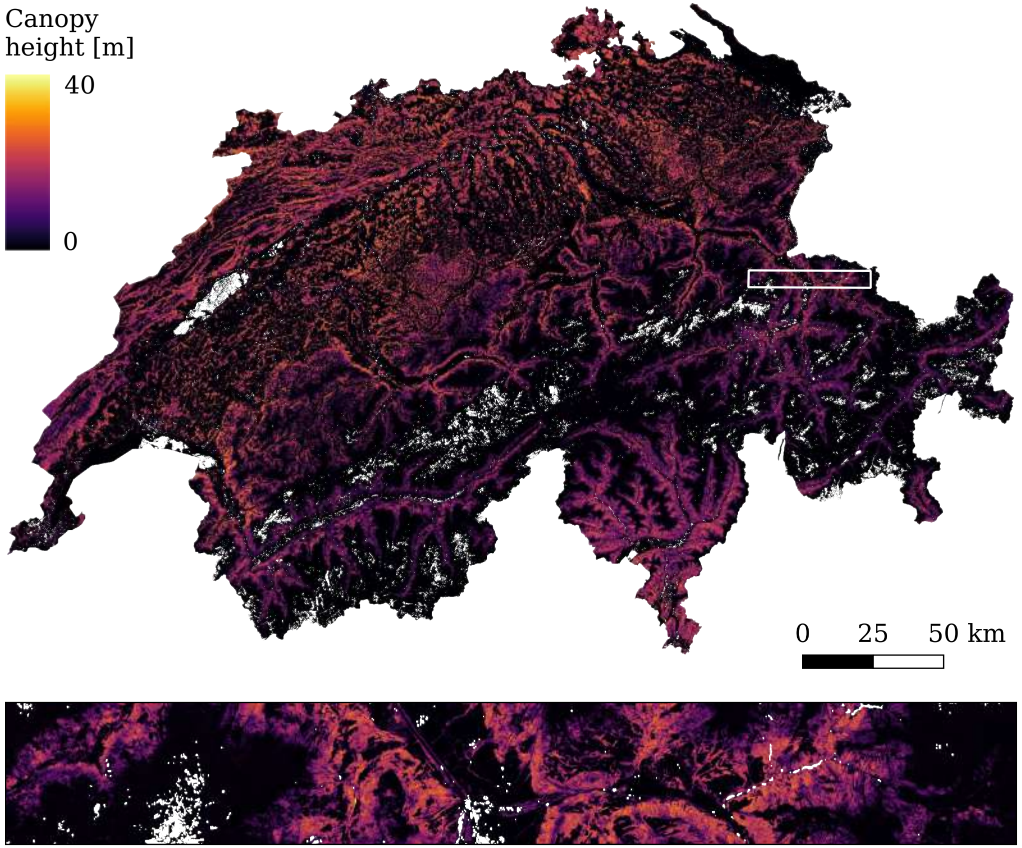


Simard et al. (2011), 1 km GSD

Lang, N., Schindler, K., & Wegner, J. D. (2019). Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sensing of Environment*, *233*, 111347.

(1) Country-wide map for Switzerland

MAE: 1.7 m



Lang et al. (2019)

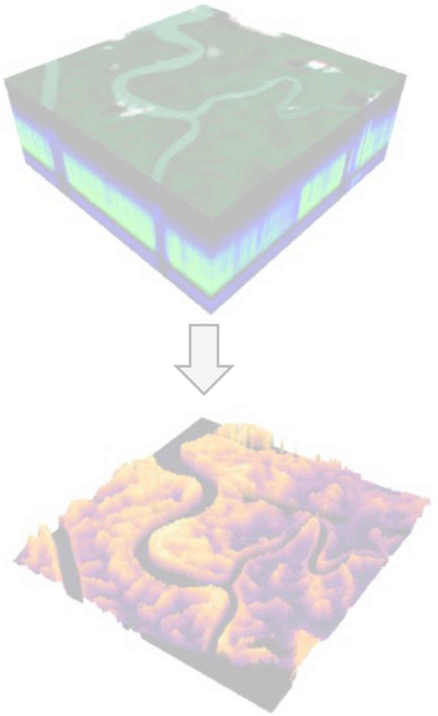
1. Canopy height can be regressed from Sentinel-2
2. Estimate canopy heights >50 m at 10 m GSD
3. Textural features and 10-meter bands are important

→ **This approach may scale globally, but we need globally distributed reference data for training.**

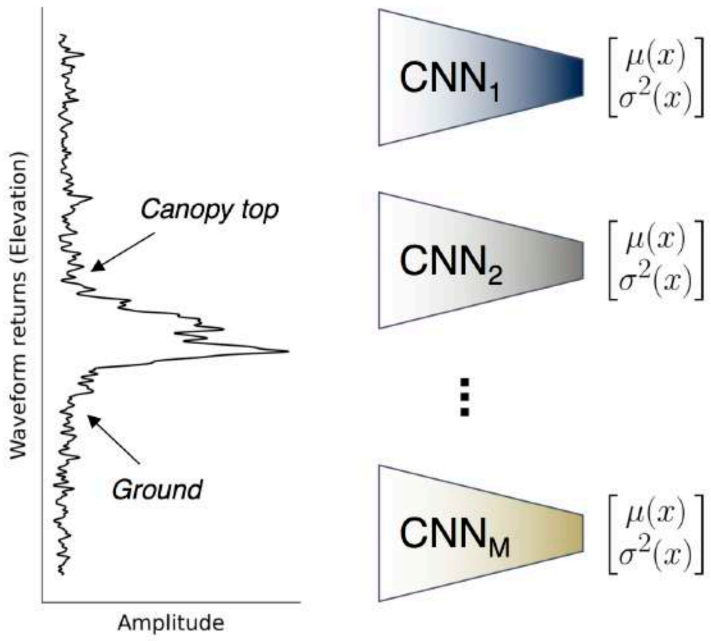
Lang, N., Schindler, K., & Wegner, J. D. (2019). Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sensing of Environment*, *233*, 111347.

Overview

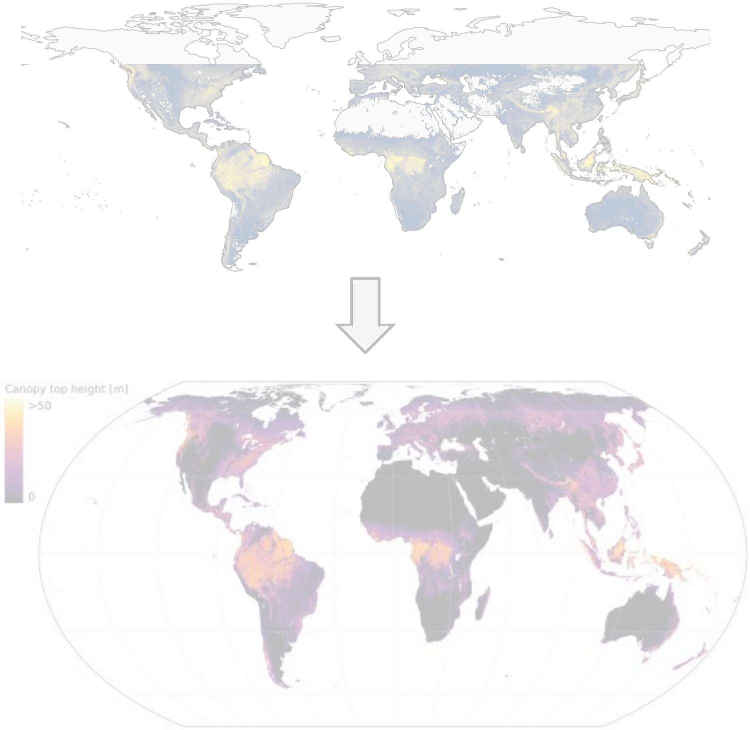
(1) Canopy height estimation from Sentinel-2 optical images



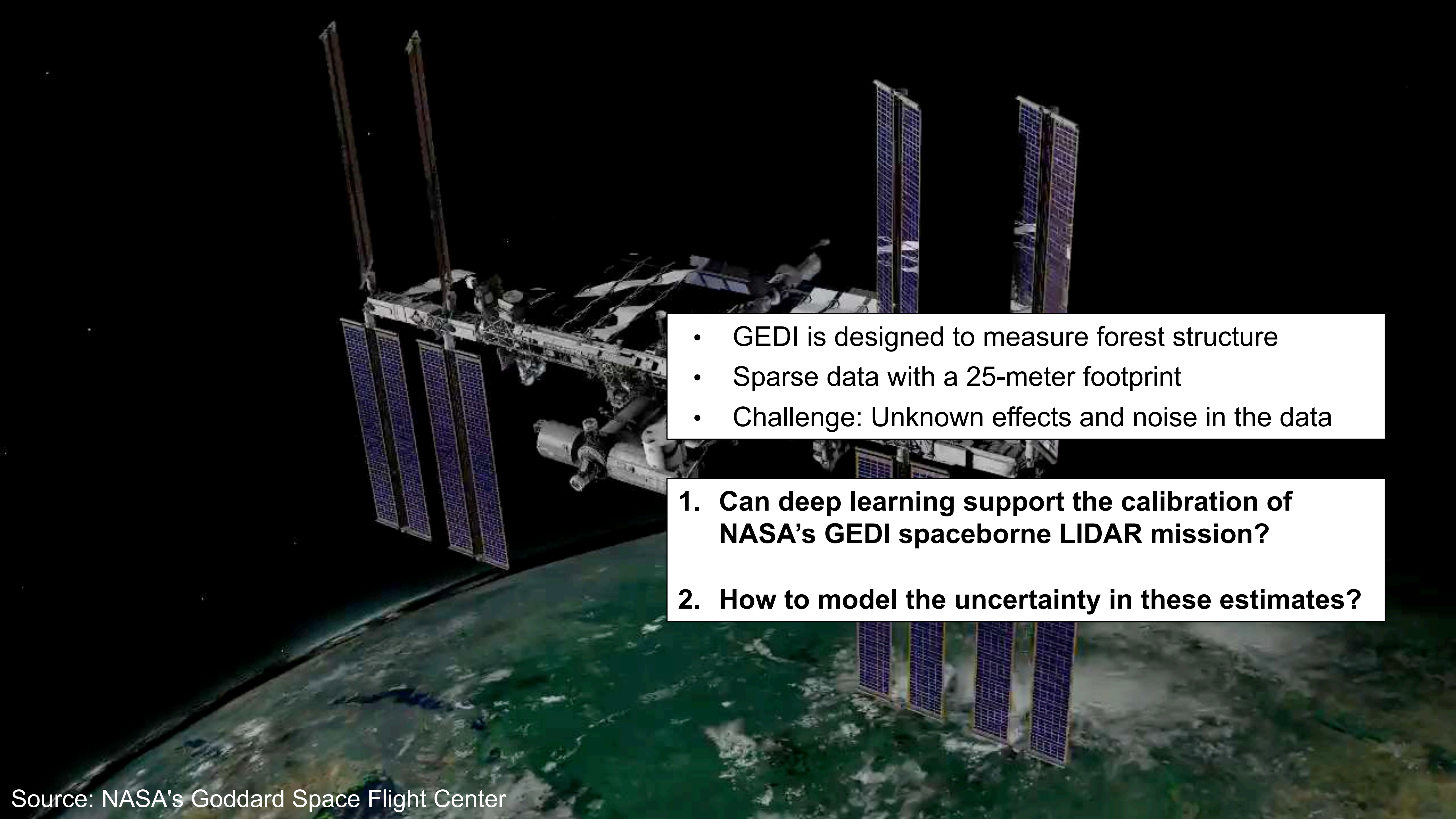
(2) Global canopy height regression and uncertainty estimation from GEDI spaceborne LIDAR



(3) A high-resolution canopy height model of the Earth



Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, *268*, 112760.



- GEDI is designed to measure forest structure
- Sparse data with a 25-meter footprint
- Challenge: Unknown effects and noise in the data

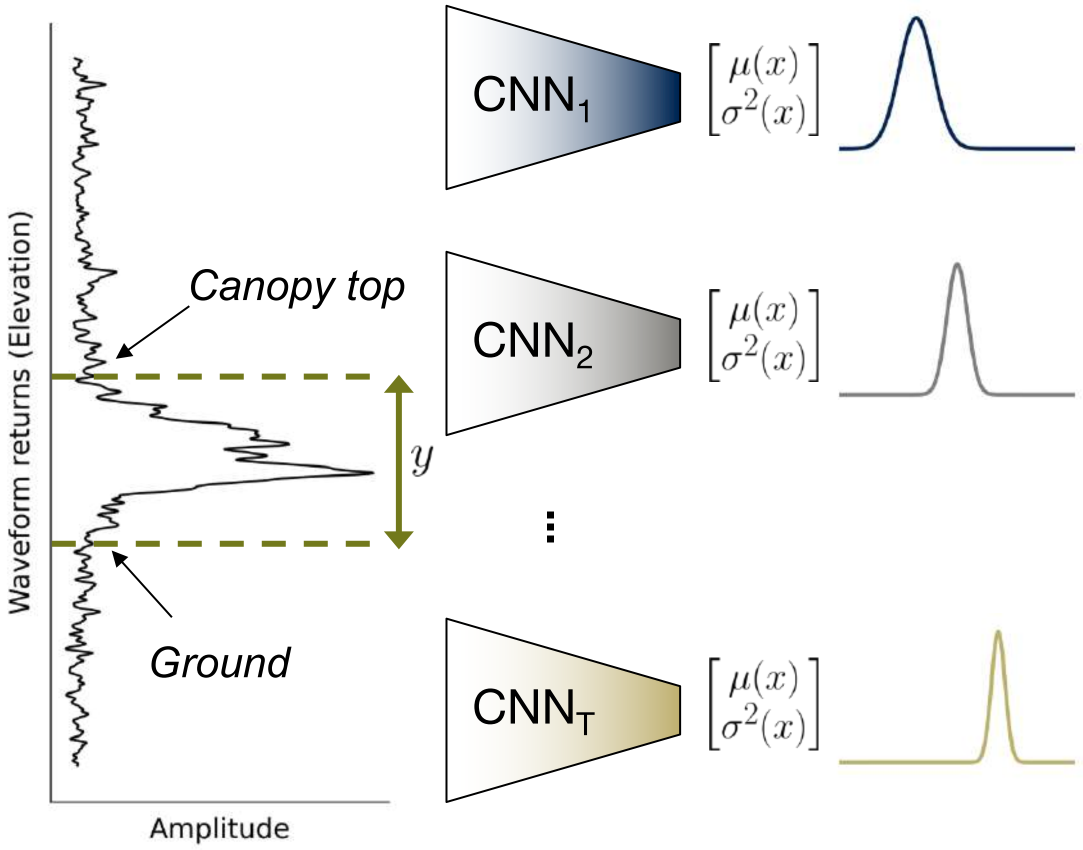
1. Can deep learning support the calibration of NASA’s GEDI spaceborne LIDAR mission?

2. How to model the uncertainty in these estimates?

Source: NASA's Goddard Space Flight Center

(2) Waveform interpretation with an ensemble of CNNs

Input waveform **Individual models**



Point estimate → **Distribution**

Gast and Roth (2018)

- Two outputs per model (mean & variance) to approximate the conditional distribution p(y|x)
- Minimize the Gaussian negative log likelihood

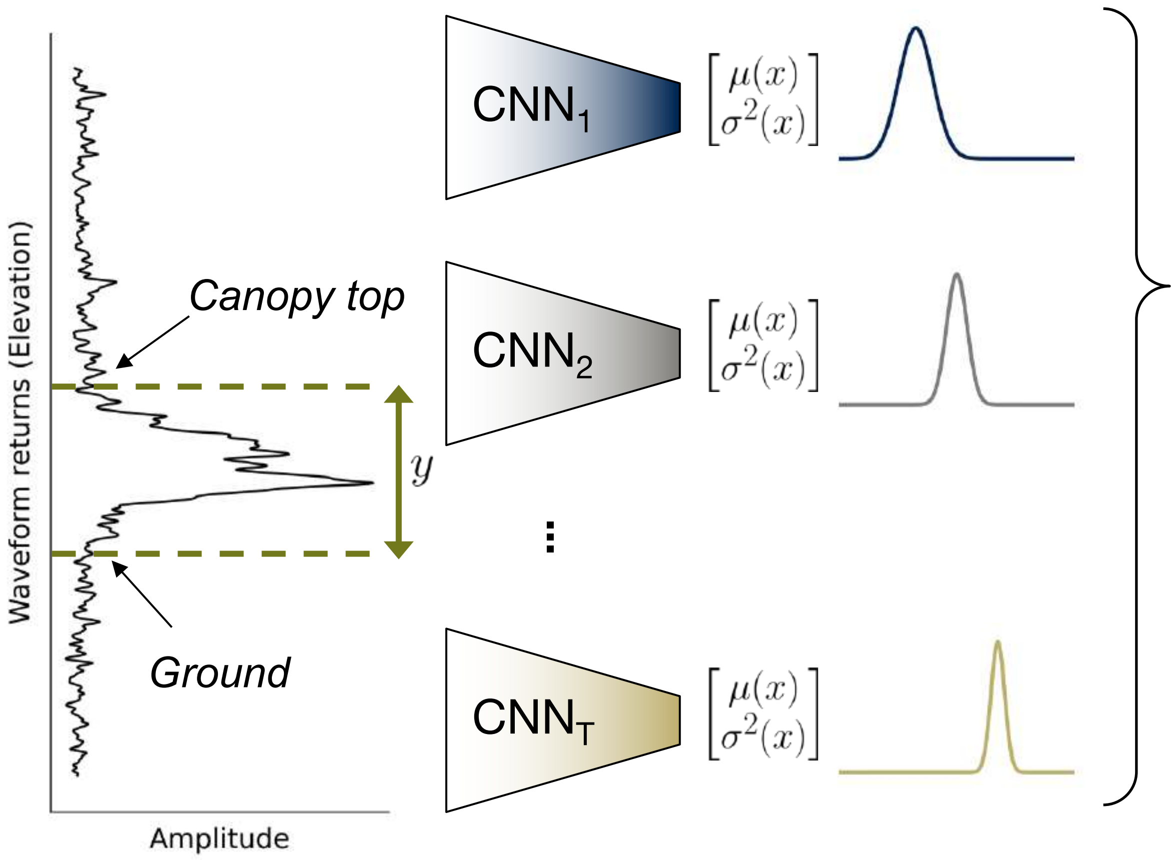
$$\mathcal{L}_{NLL} = \frac{1}{N}\sum_{i=1}^{N} \frac{(\hat{\mu}(x_i) - y_i)^2}{2\hat{\sigma}^2(x_i)} + \frac{1}{2}\log \hat{\sigma}^2(x_i)$$

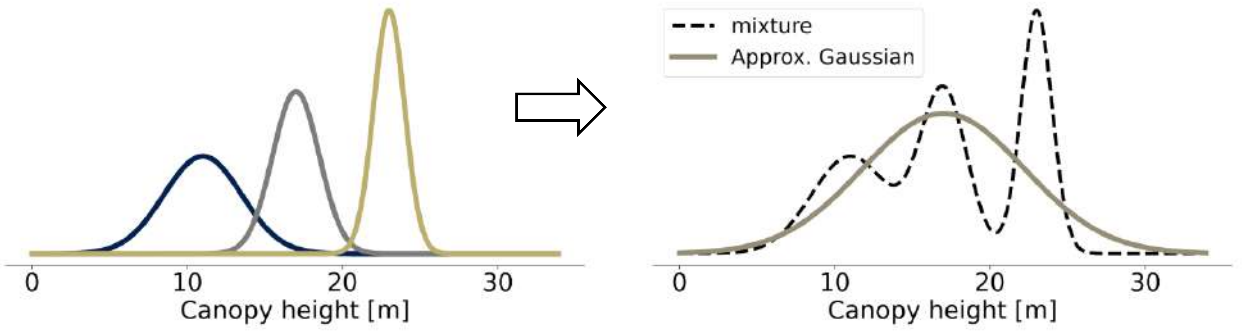
- Ensemble of CNNs, each trained separately, starting from different random initializations

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, *268*, 112760.

(2) Waveform interpretation with an ensemble of CNNs

Input waveform **Individual models** **Ensemble averaging**





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

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

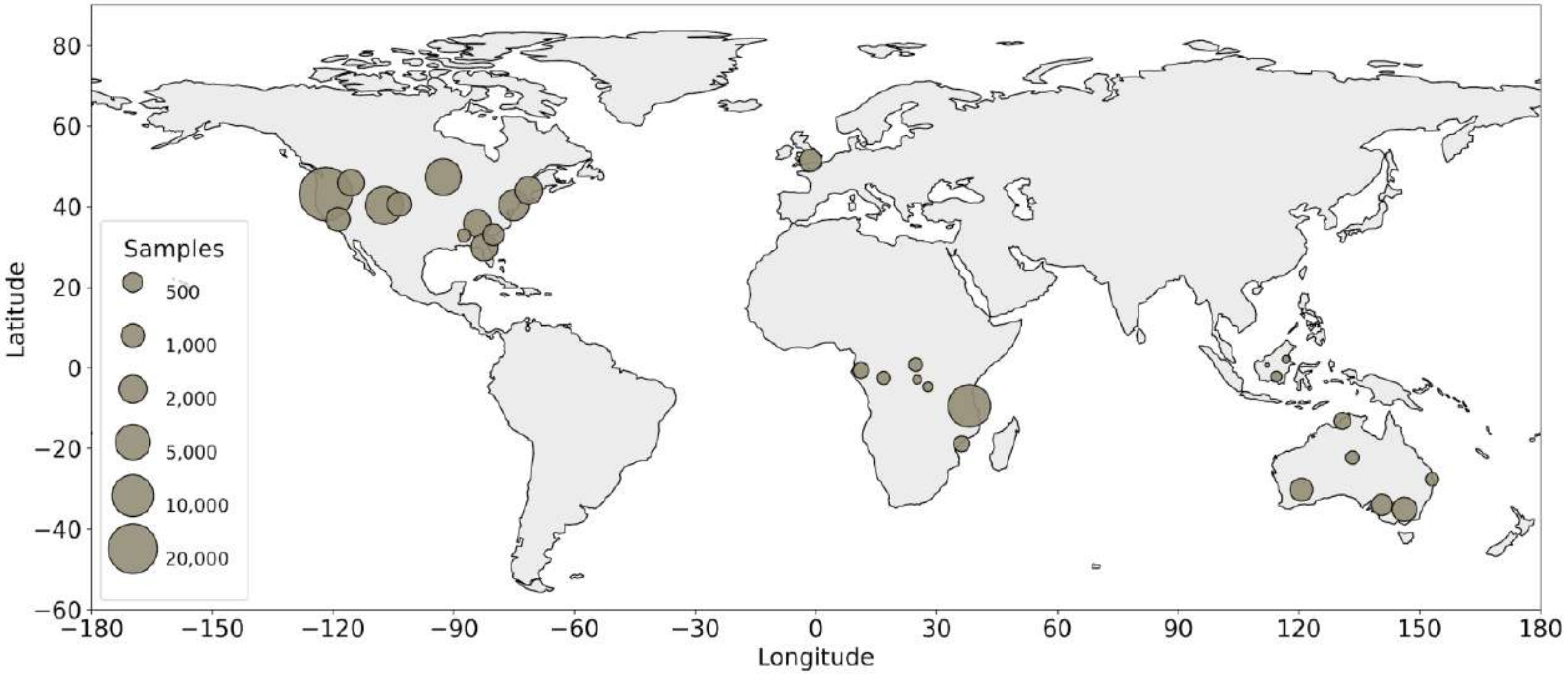
epistemic
(model uncertainty)
"knowledge"

aleatoric
(data uncertainty)
"rolling a dice"

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, *268*, 112760.

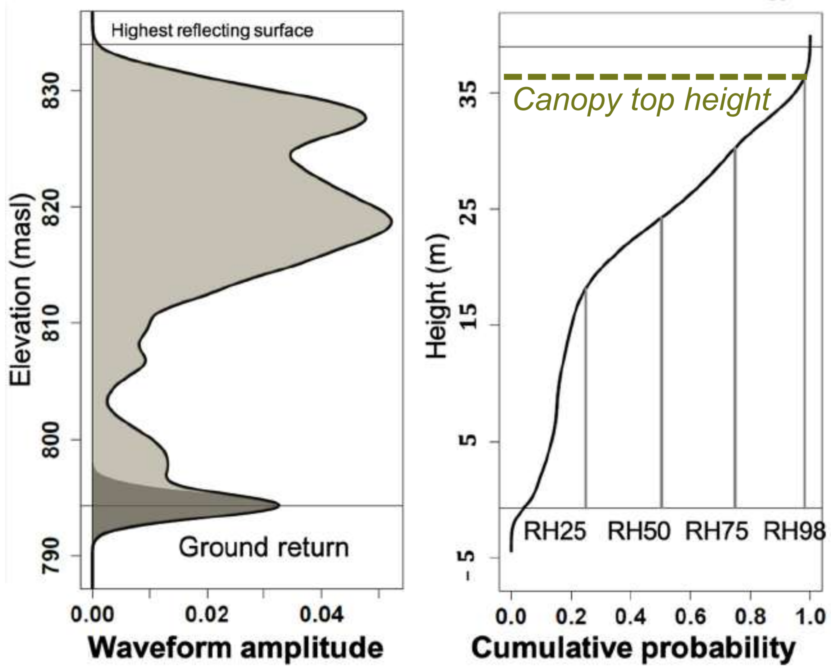
(2) Reference dataset: GEDI waveforms matched with ALS data

Total: 70,000 on-orbit waveforms matched with ALS reference data



Simulated waveform

Cumulative return energy

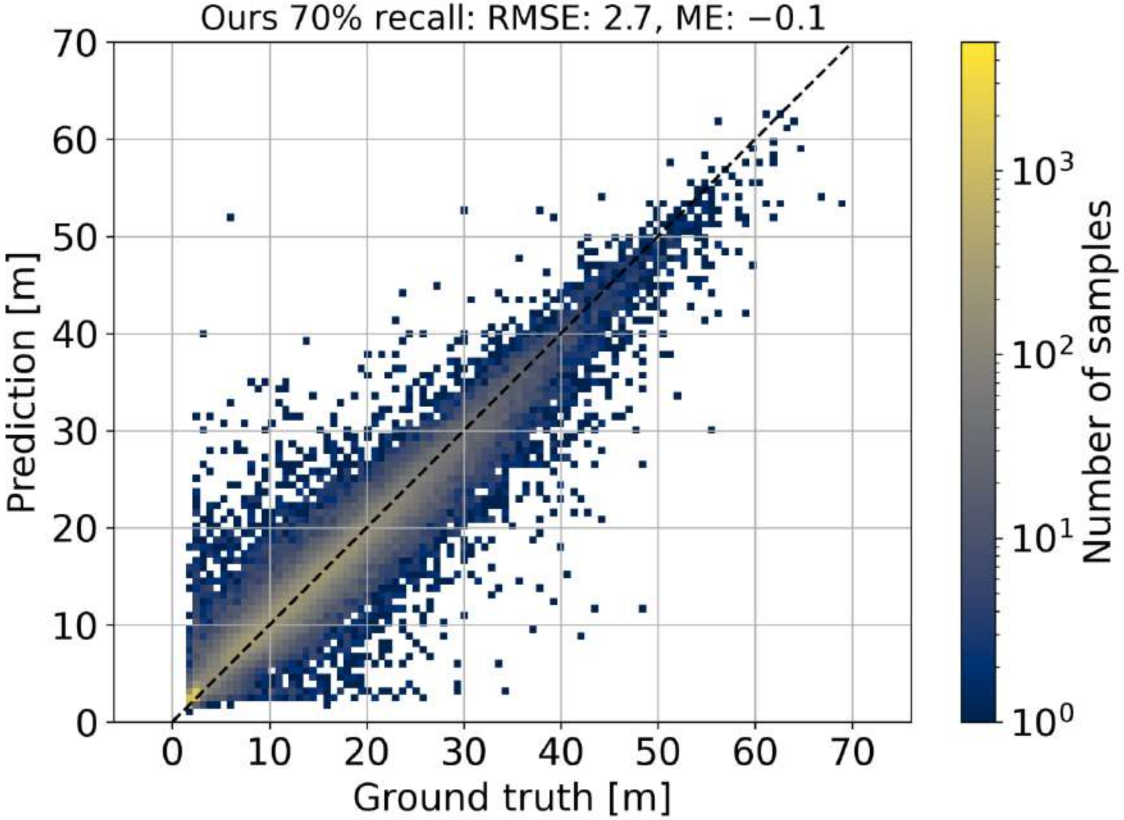


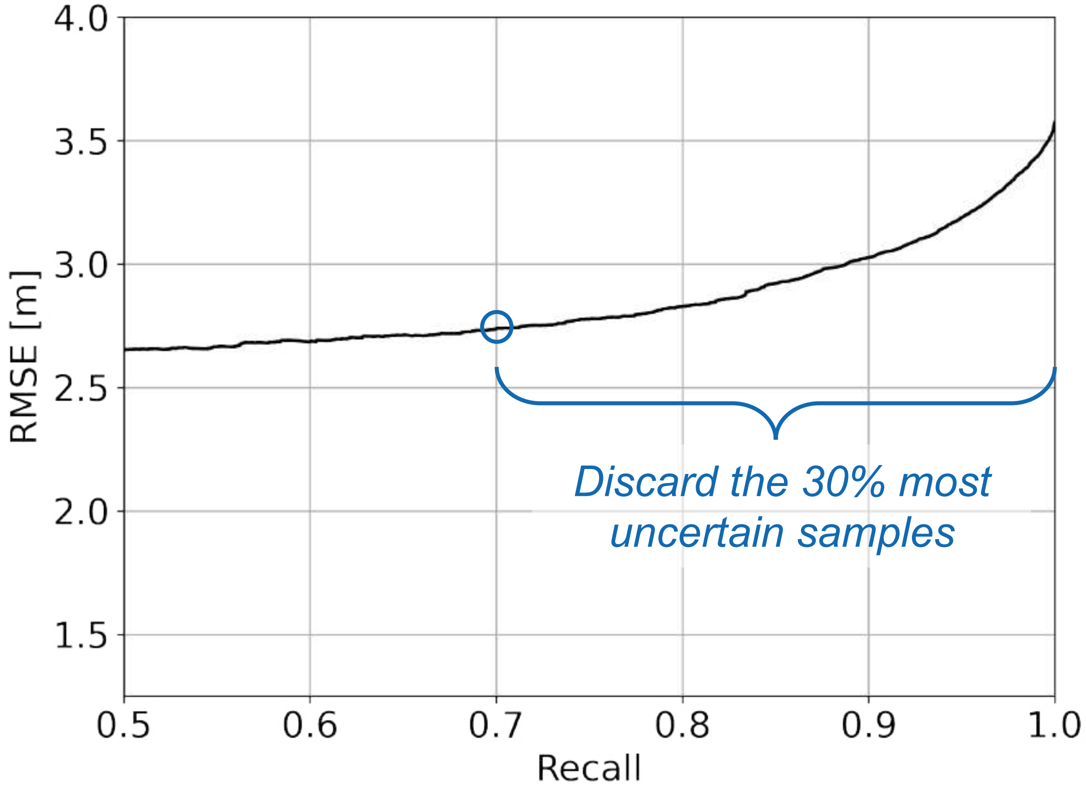
Duncanson et al. (2022)

- Reference canopy top height from airborne laser scanning (ALS)
- RH98 as a proxy for canopy top height

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, *268*, 112760.

(2) Canopy top height regression from GEDI waveforms

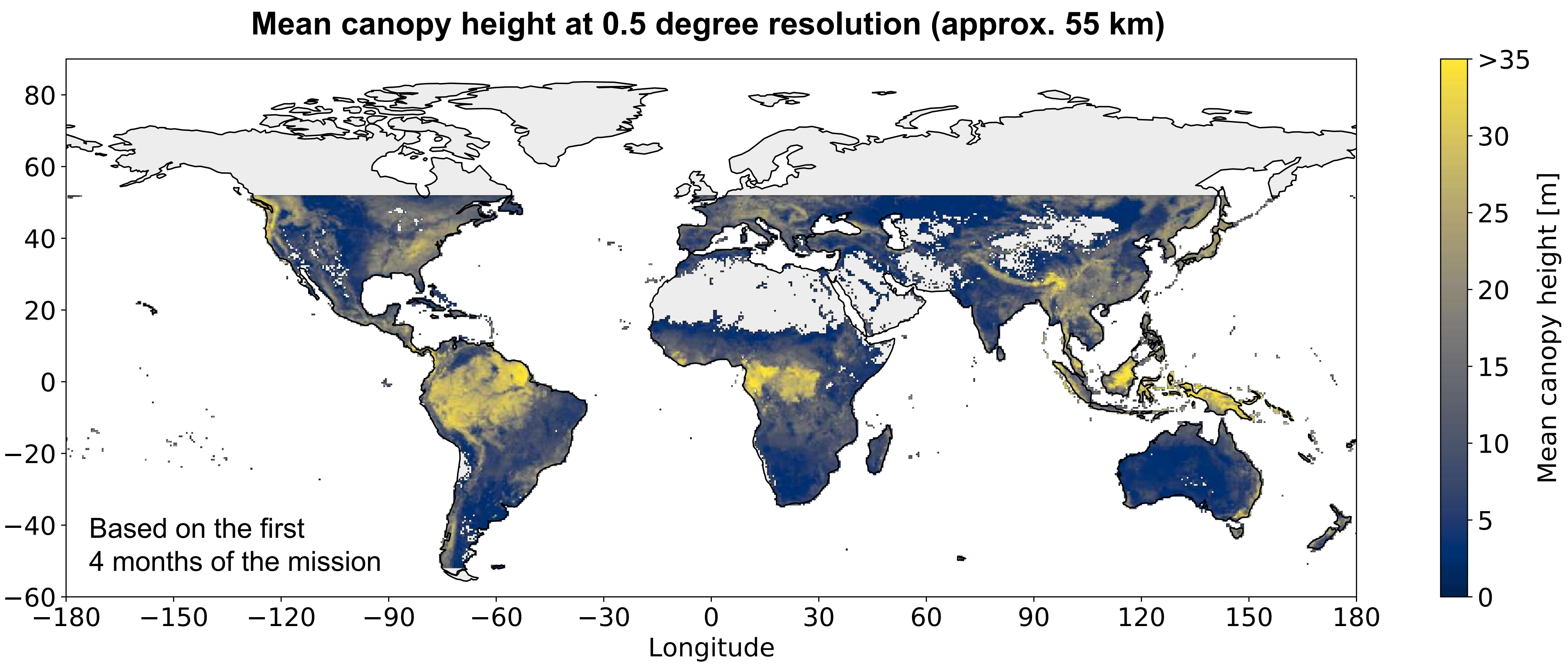




→ Predictive uncertainty allows to filter erroneous predictions and reduce overall error

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, *268*, 112760.

(2) Sparse canopy height estimates from GEDI at global scale

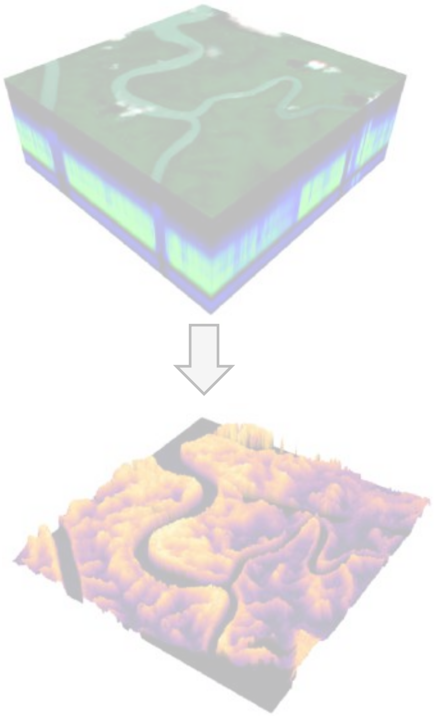


→ Now we have access to globally distributed reference data!

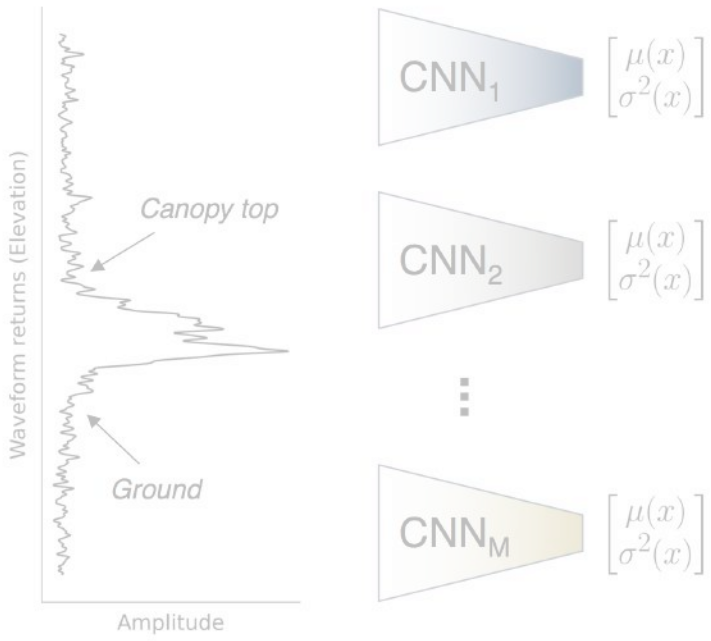
Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, *268*, 112760.

Overview

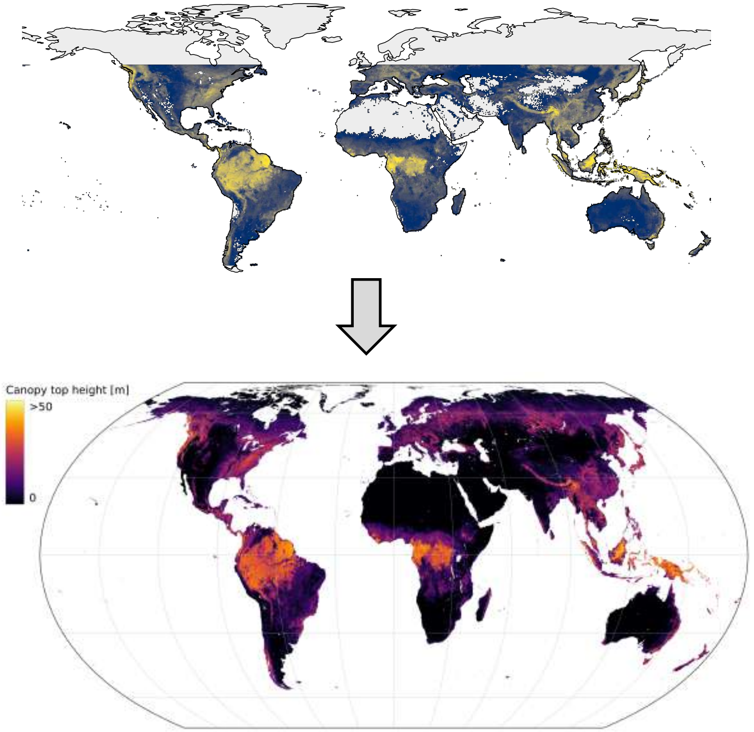
(1) Canopy height estimation from Sentinel-2 optical images



(2) Global canopy height regression and uncertainty estimation from GEDI spaceborne LIDAR



(3) A high-resolution canopy height model of the Earth

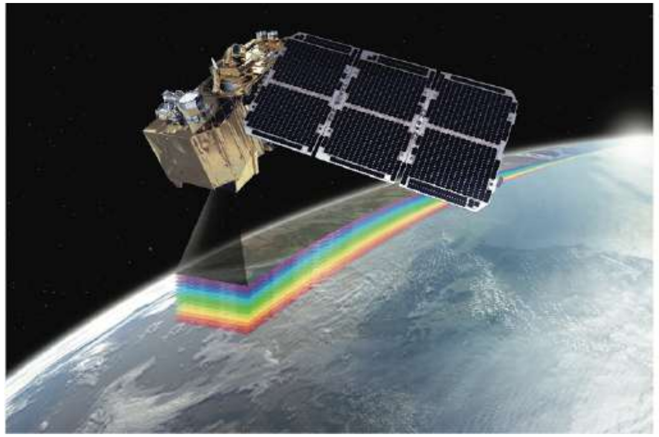


Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2022). A high-resolution canopy height model of the Earth. *arXiv preprint arXiv:2204.08322*. (Under review)

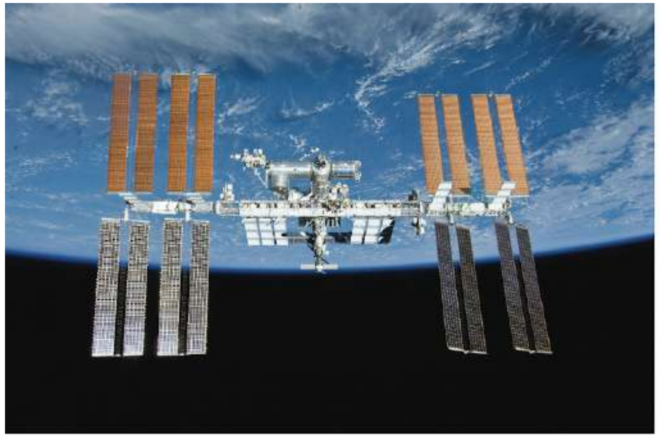
(3) Idea

Combine **Sentinel-2 optical satellite images** with sparse space-based **GEDI LIDAR** data to map **canopy height globally at high-resolution.**

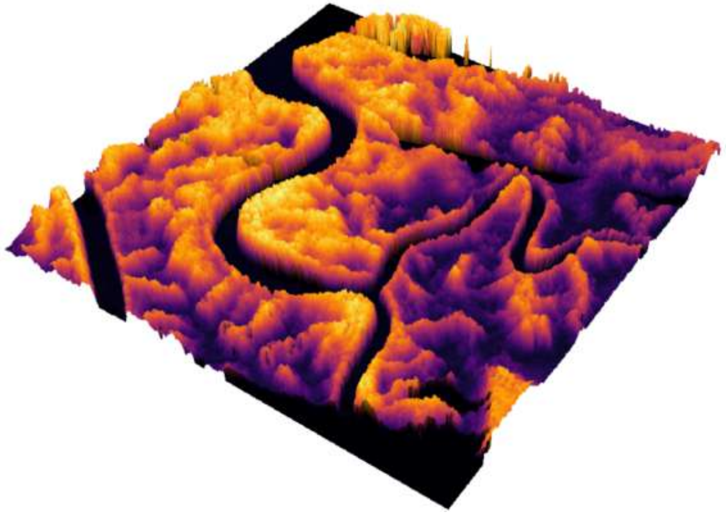
Sentinel-2



GEDI LIDAR

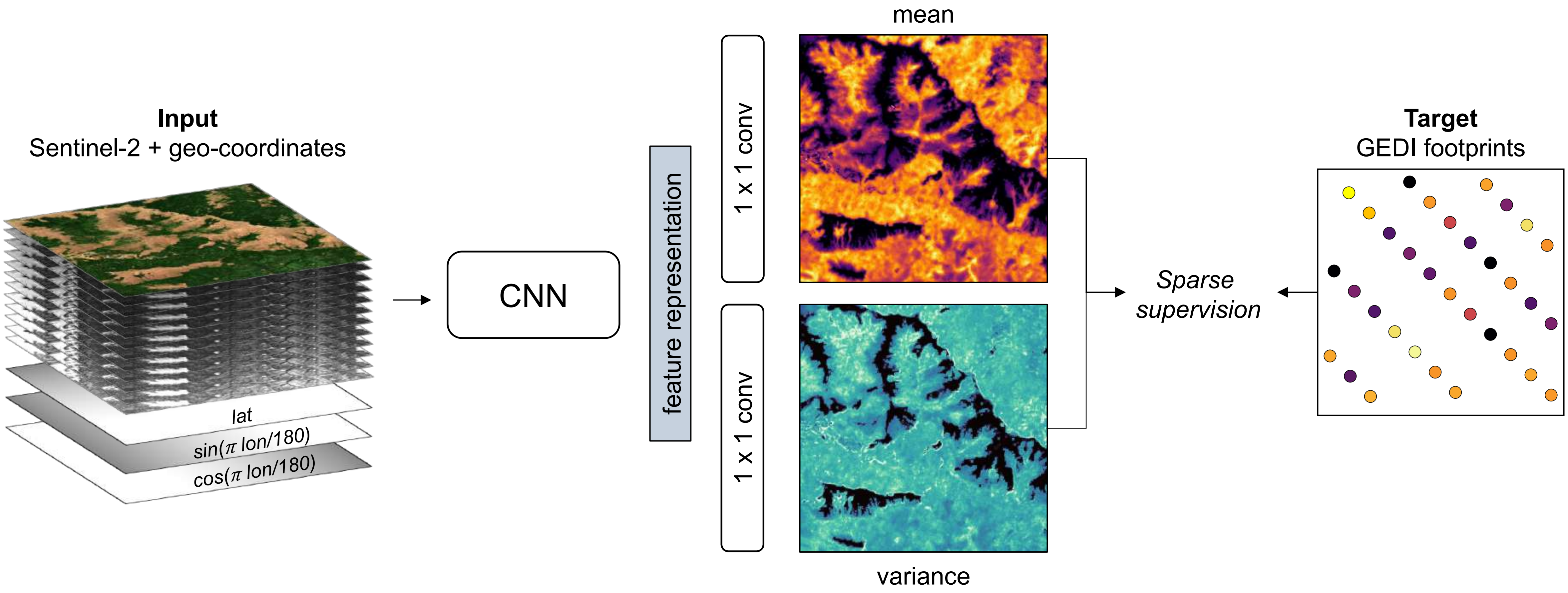


Dense canopy height maps



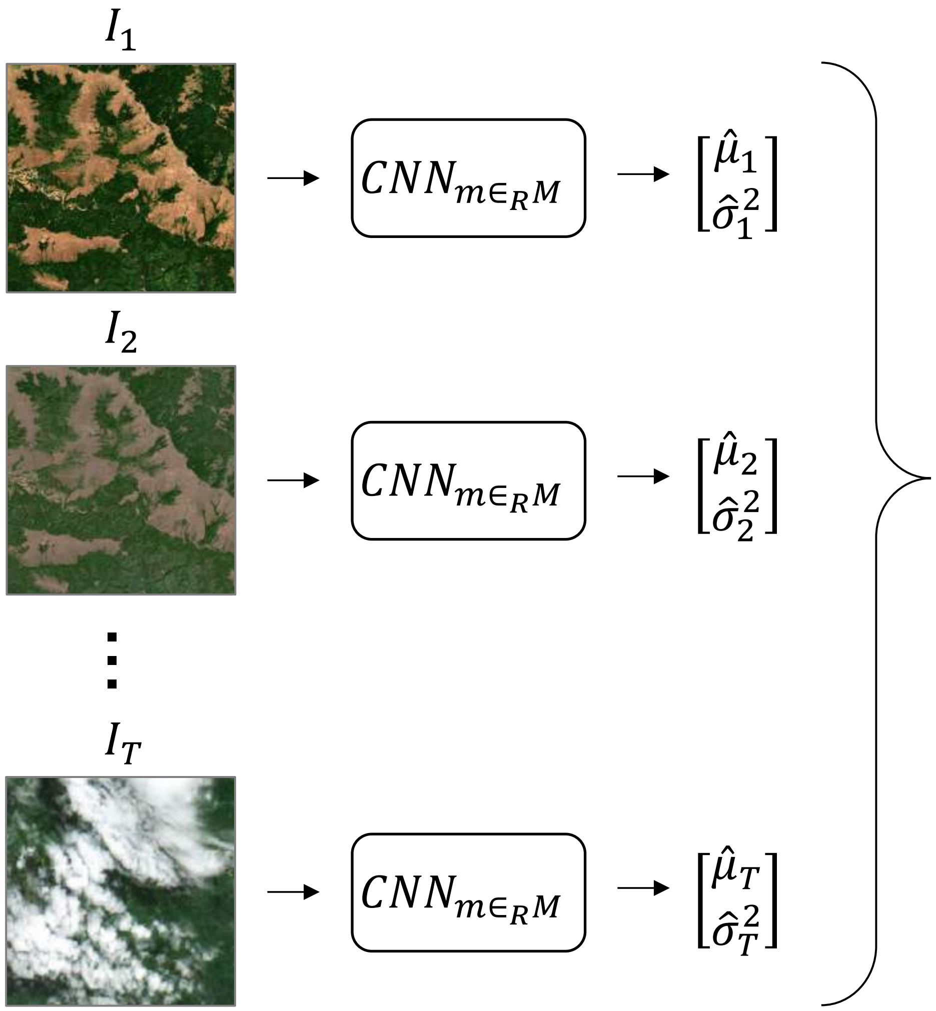
(3) Learning a global Sentinel-2 model with sparse GEDI supervision

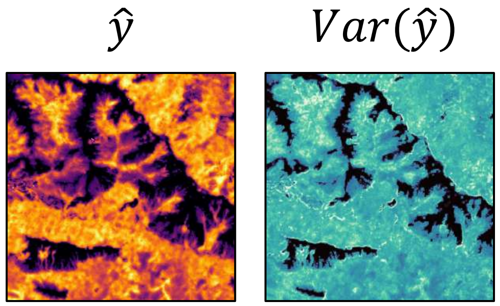
1. Adapt the **uncertainty estimation** approach
2. Train a **global Sentinel-2 regression model** using **sparse supervision from GEDI**



Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2022). A high-resolution canopy height model of the Earth. arXiv preprint arXiv:2204.08322. (Under review)

(3) Inference and ensembling of repeated predictions





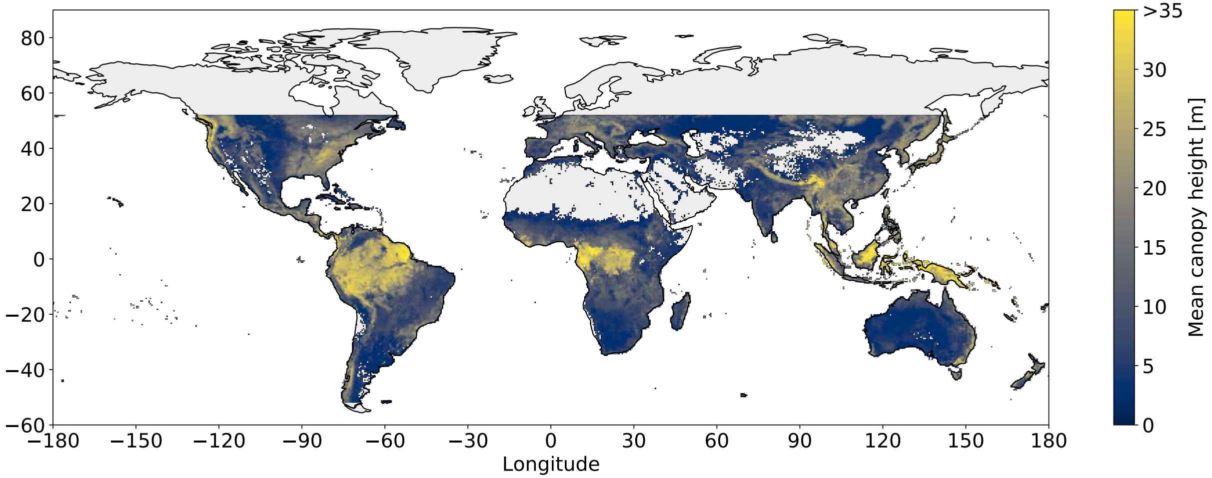
- Aggregated multiple observations to obtain complete maps and use redundancy
- Inverse variance weighting to reduce the impact of uncertain estimates.

Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2022). A high-resolution canopy height model of the Earth. arXiv preprint arXiv:2204.08322. (Under review)

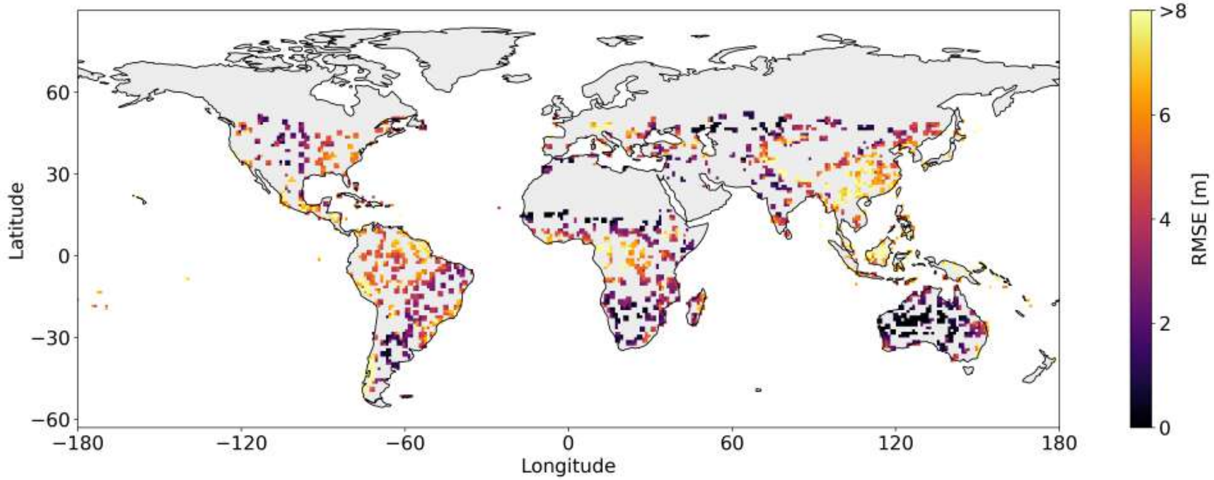
(3) Reference dataset: combining GEDI with Sentinel-2 at global scale

- **600 x 10^6 GEDI shots** with **canopy top height** estimates (April - July 2019 & 2020)
- Combined with **>13,000 Sentinel-2** tiles to create a **global training dataset**
- **Regional train-val-split** at the Sentinel-2 tile-level (i.e. 100 km x 100 km regions)

Mean canopy height from sparse GEDI
at 0.5 degree resolution (approx. 55 km)



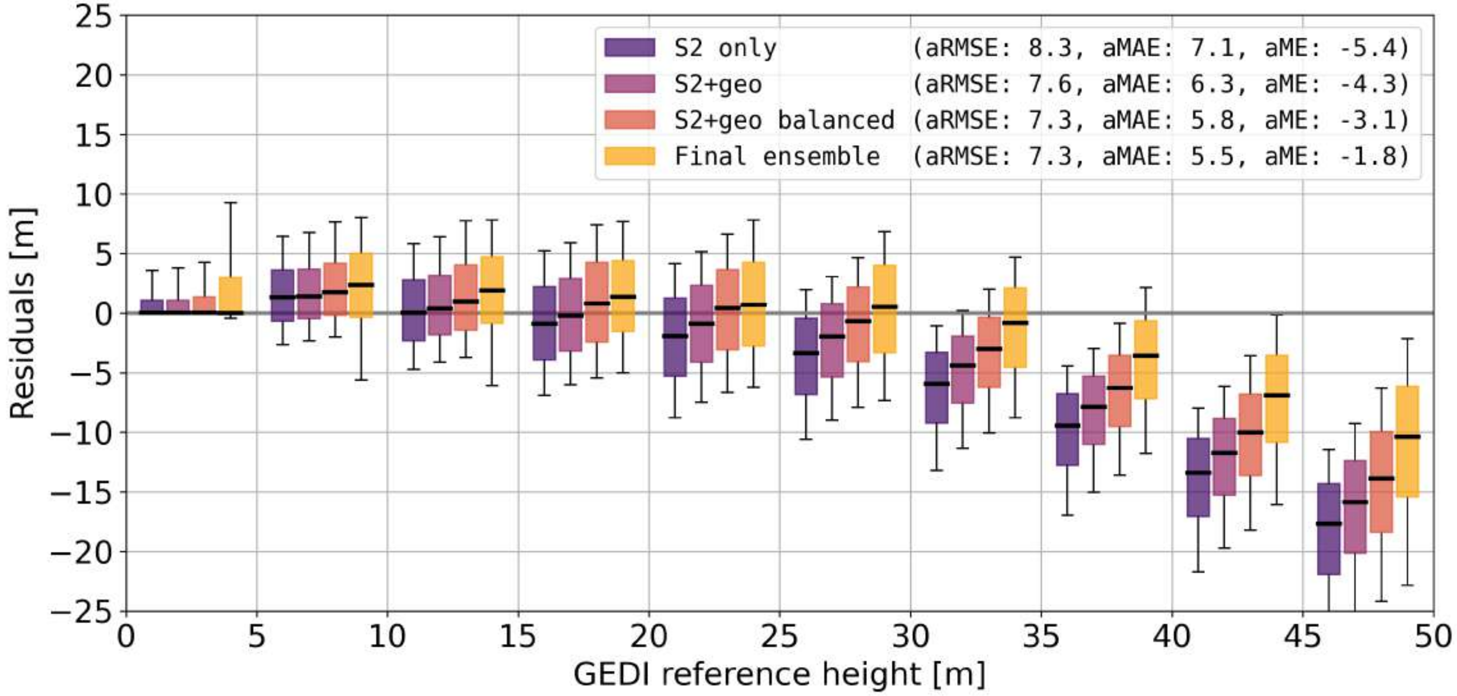
Hold-out validation regions
(RMSE within 0.5 degree cells)



Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2022). A high-resolution canopy height model of the Earth. arXiv preprint arXiv:2204.08322. (Under review)

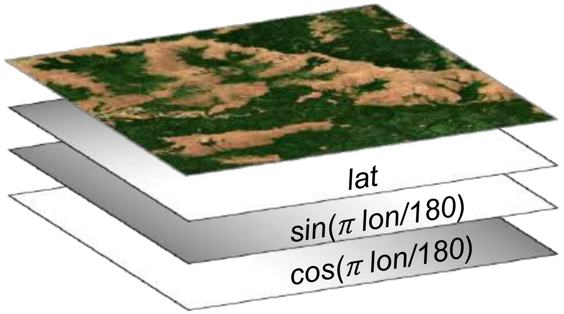
(3) Comparison of different strategies

Residuals within 5-meter height intervals

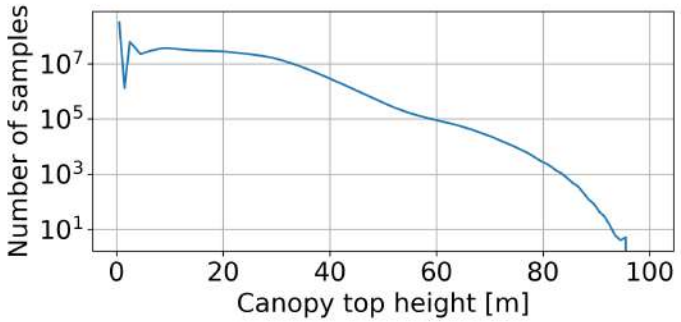


Final model: **6.0 m RMSE, 1.3 m ME**

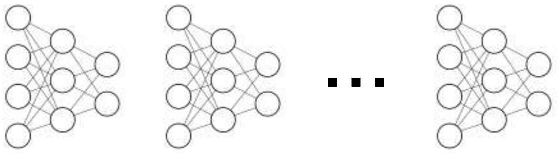
1. Geographical coordinates as inputs



2. Fine-tuning with a balanced loss



3. Ensembling

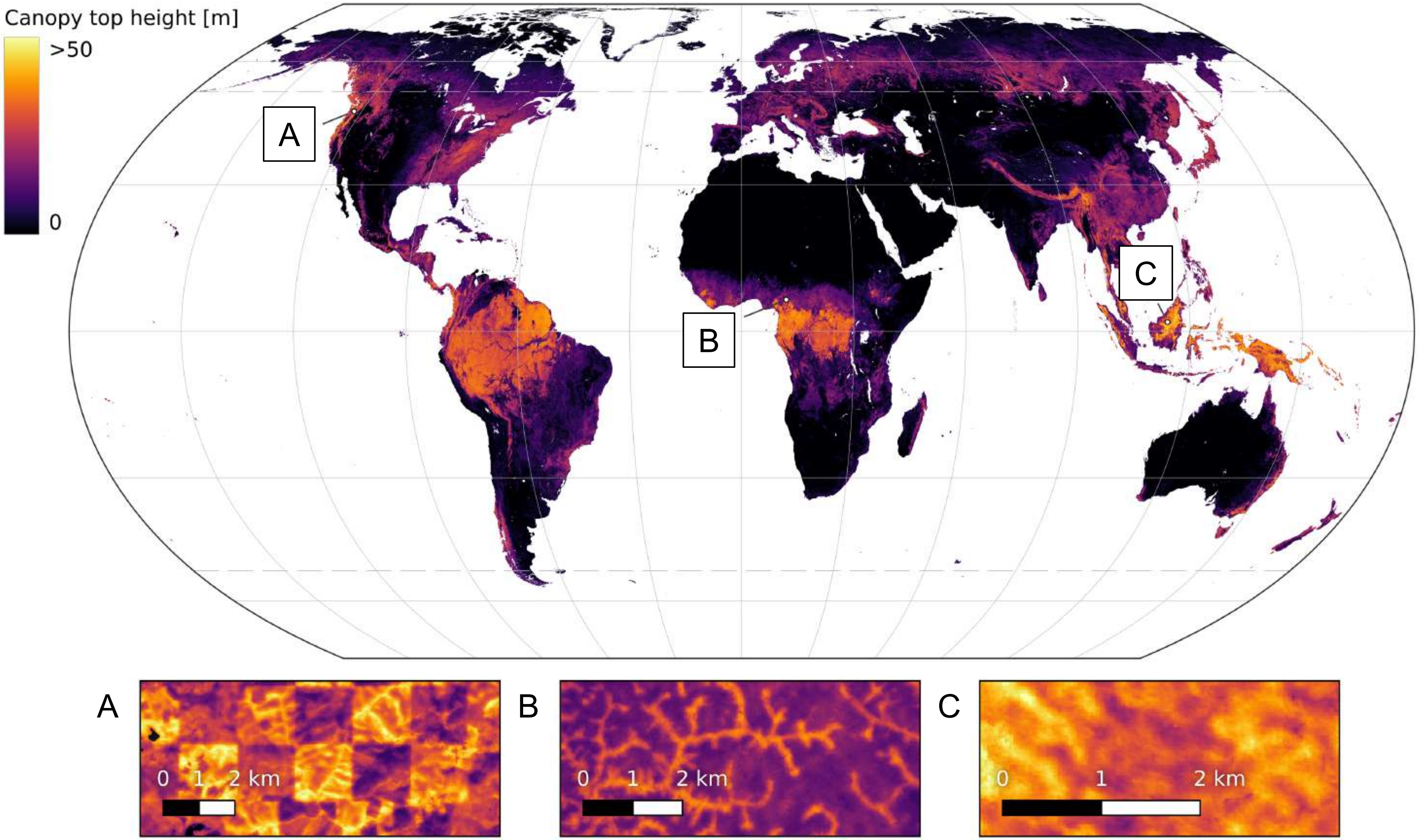


Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2022). A high-resolution canopy height model of the Earth. arXiv preprint arXiv:2204.08322. (Under review)

(3) Global canopy height map for 2020 estimated from Sentinel-2

- **18k Sentinel-2 tiles** processed to cover the global landmass (without Antarctica)

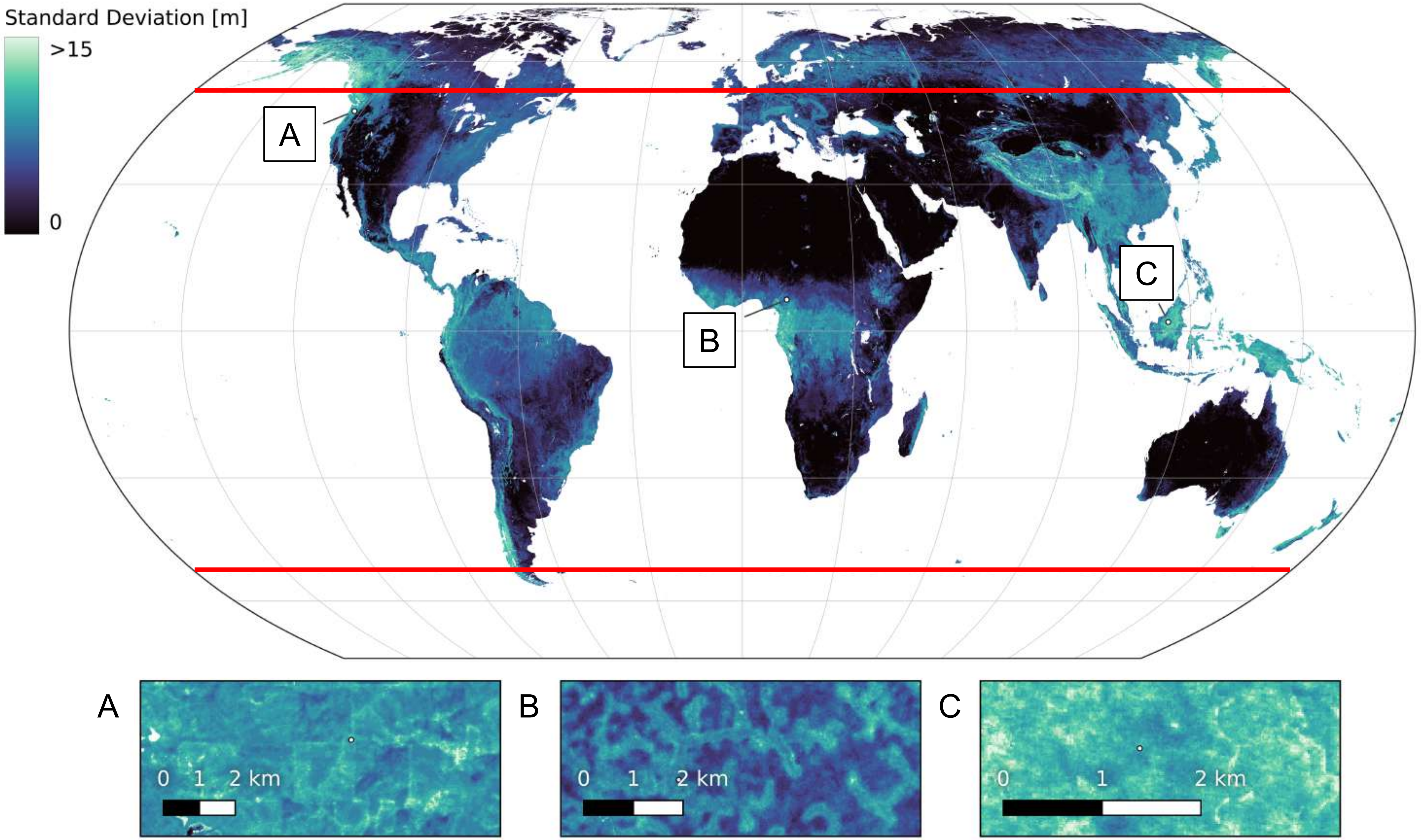
- Deployed on the **10 images** (dates) with lowest cloud cover per tile and "relevant" orbits.
  ≈160 TB,
  10 days (≈3 GPU-years)



Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2022). A high-resolution canopy height model of the Earth. arXiv preprint arXiv:2204.08322. (Under review)

(3) Predictive uncertainty

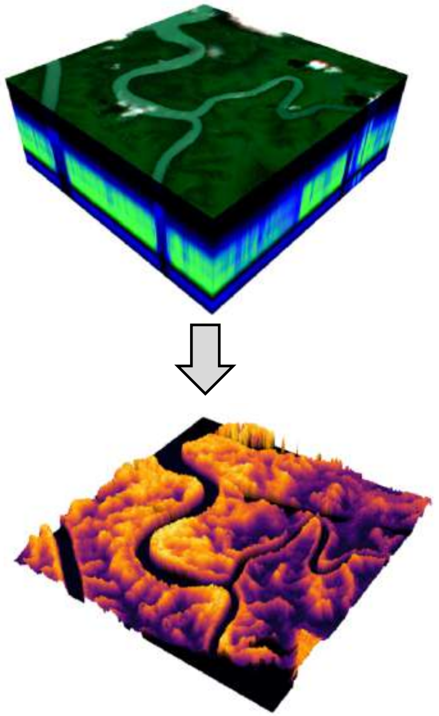
- **Epistemic uncertainty** is modelled with a **deep ensemble** of five CNNs.
- **Aleatoric uncertainty** results from the weighted average of the estimated variances.



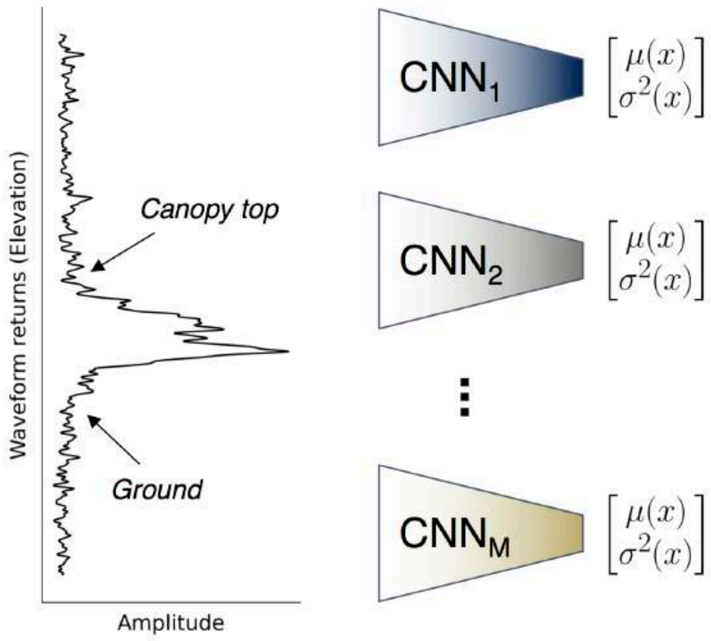
Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2022). A high-resolution canopy height model of the Earth. arXiv preprint arXiv:2204.08322. (Under review)

Overview

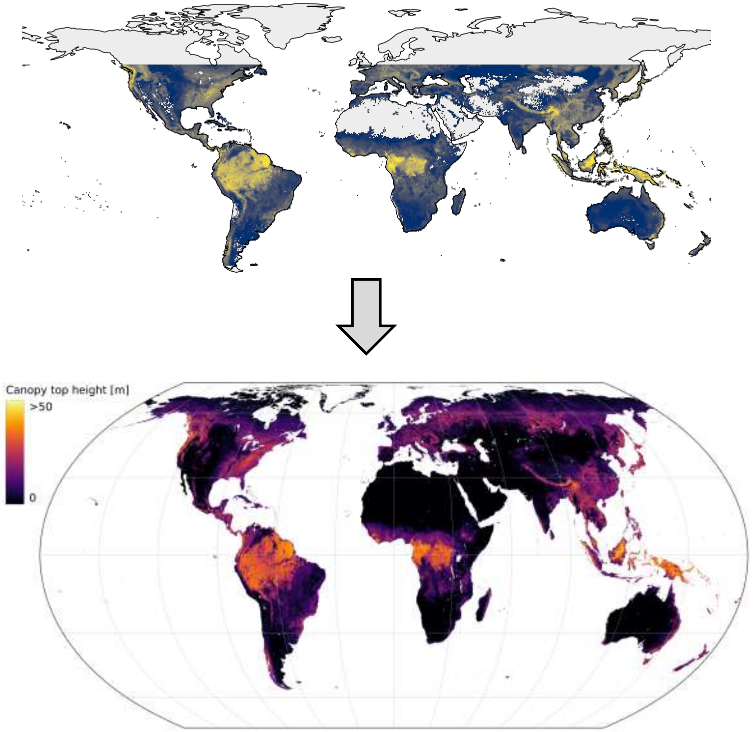
(1) Canopy height estimation from Sentinel-2 optical images



(2) Global canopy height regression and uncertainty estimation from GEDI spaceborne LIDAR

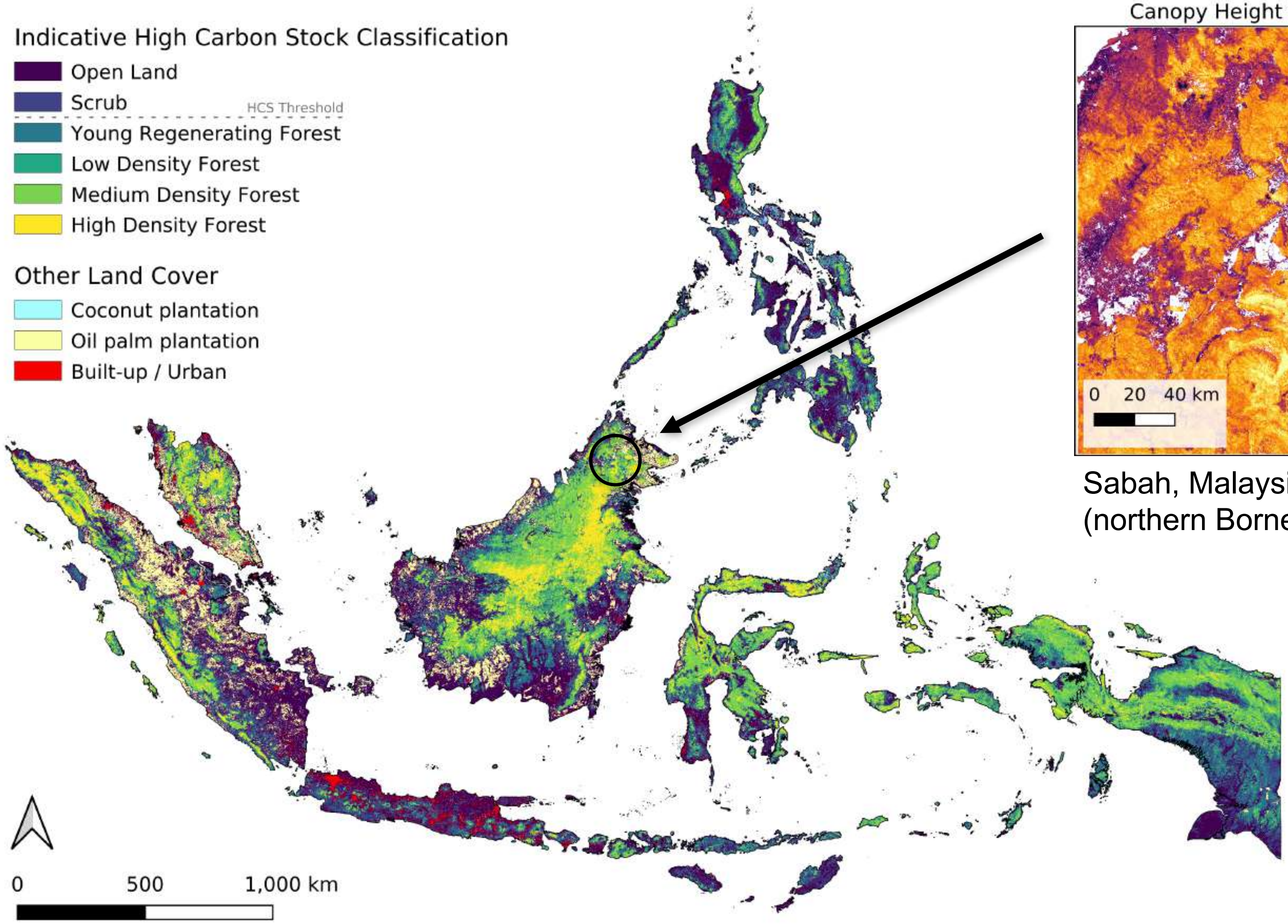


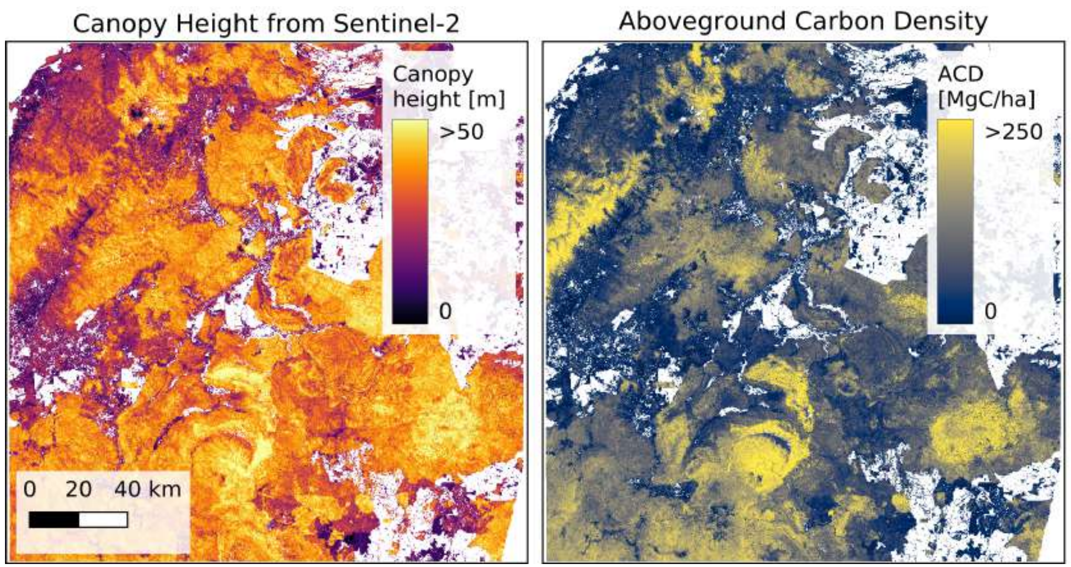
(3) A high-resolution canopy height model of the Earth



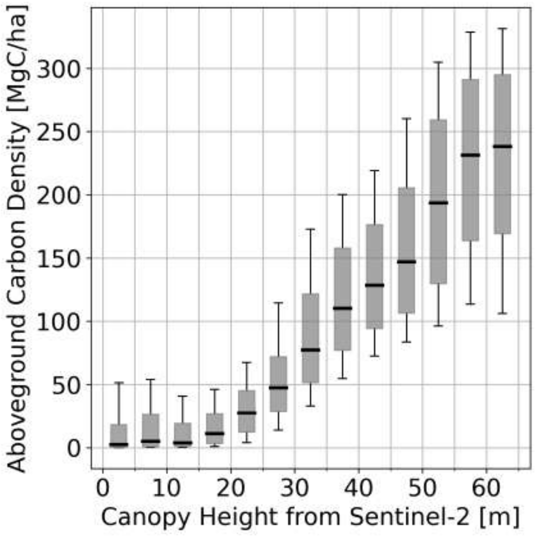
→ Potential applications

Potential for biomass and carbon stock estimation



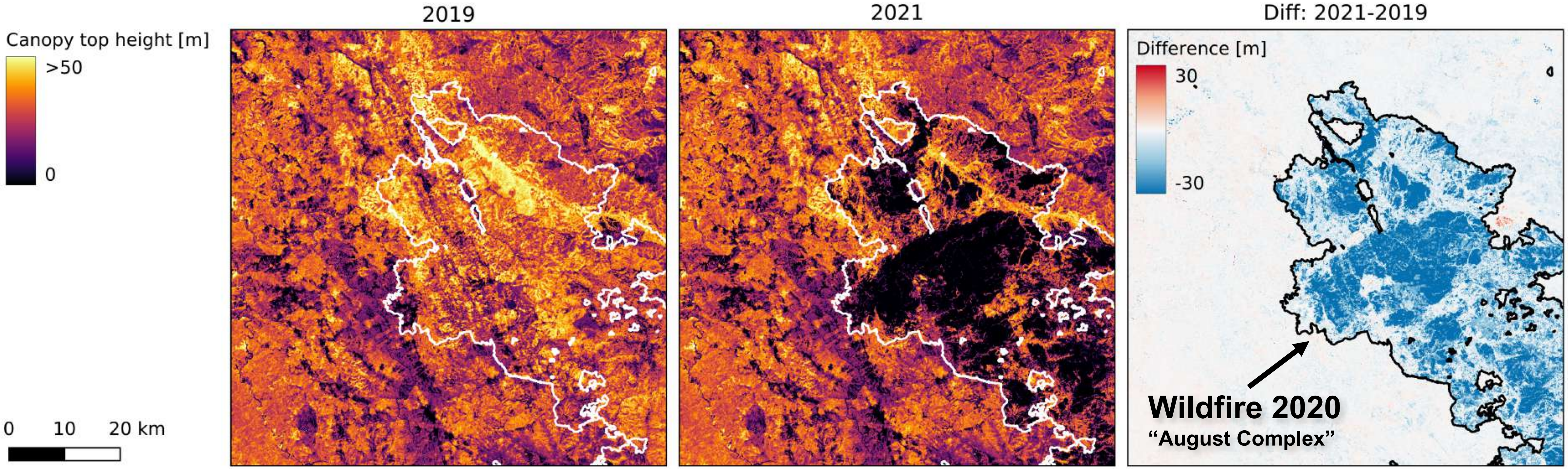


Sabah, Malaysia
(northern Borneo)



Lang, N., Schindler, K., & Wegner, J. D. (2021). High carbon stock mapping at large scale to guide forest conservation in tropical Southeast Asia. (Under review)

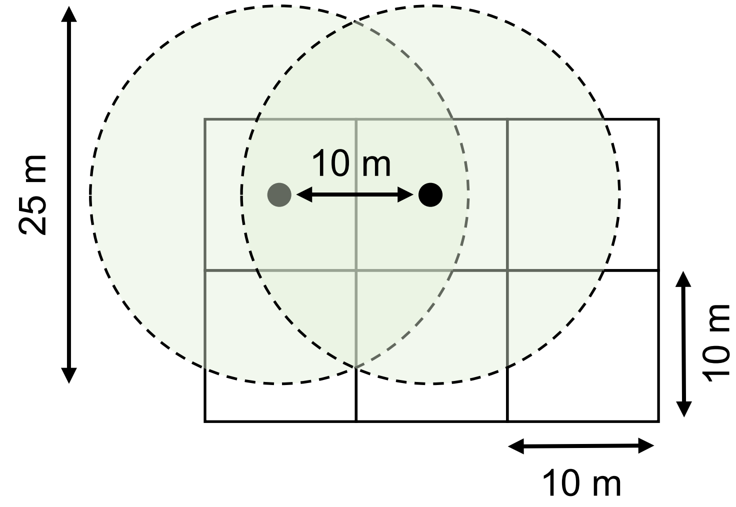
Potential for change analyses



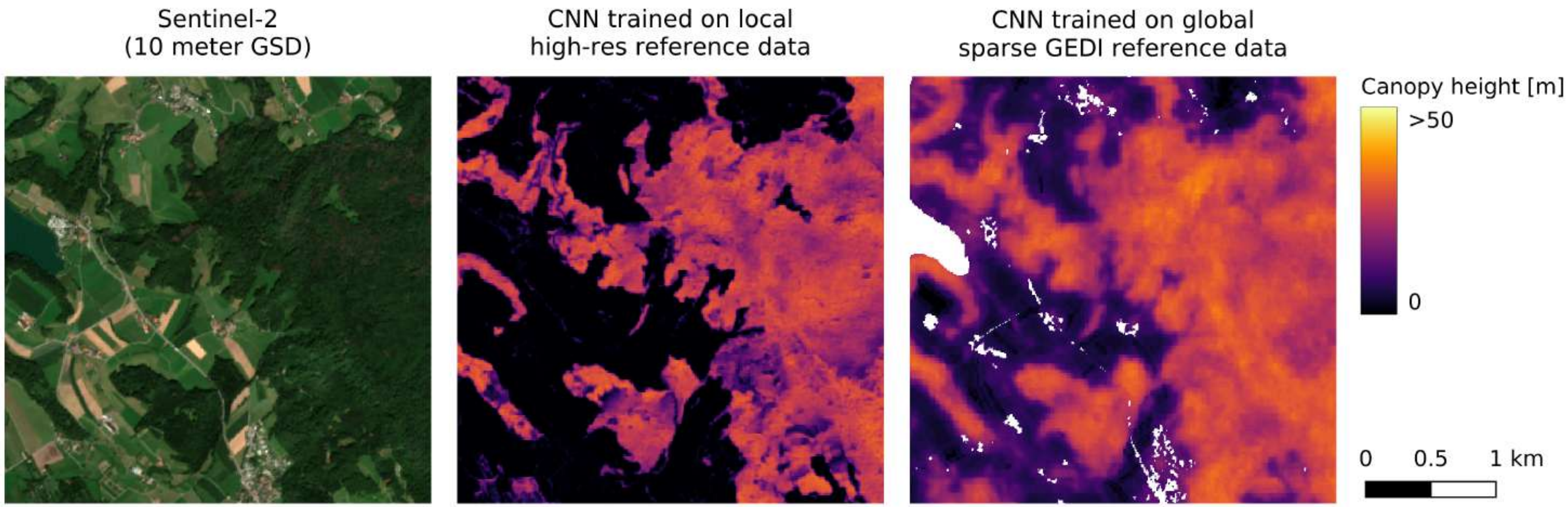
Mendocino National Forest in northern California, approx. 300 km north of San Francisco

Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2022). A high-resolution canopy height model of the Earth. arXiv preprint arXiv:2204.08322. (Under review)

Limitation: Spatial resolution of dense predictions



- Models trained on GEDI reference data yield a reduced effective resolution
- Possible reasons for reduced resolution:
  - Max height within 25-meter footprint
  - Sparse supervision
  - Geolocation uncertainty
- *Future work: How can we preserve small structures that are visible in the input?*



Conclusion and Outlook

- ✓ Sentinel-2 images are useful to map canopy height
- ✓ Deep learning is key to extract predictive features
- ✓ Uncertainty estimates can indicate errors

Open research questions

How can we ...

... preserve spatial resolution with sparse supervision?

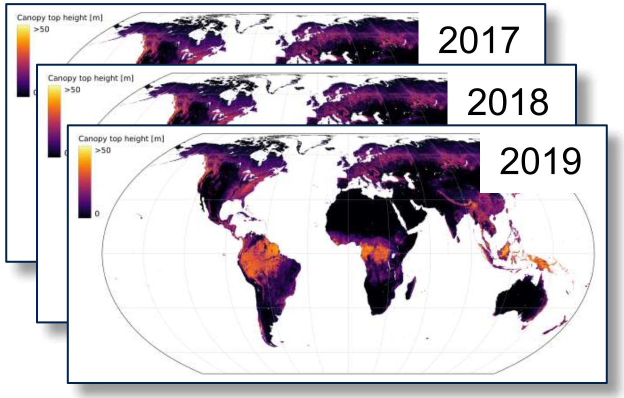
... reduce ambiguity and saturation in the predictions?

... improve geographic and temporal generalization?

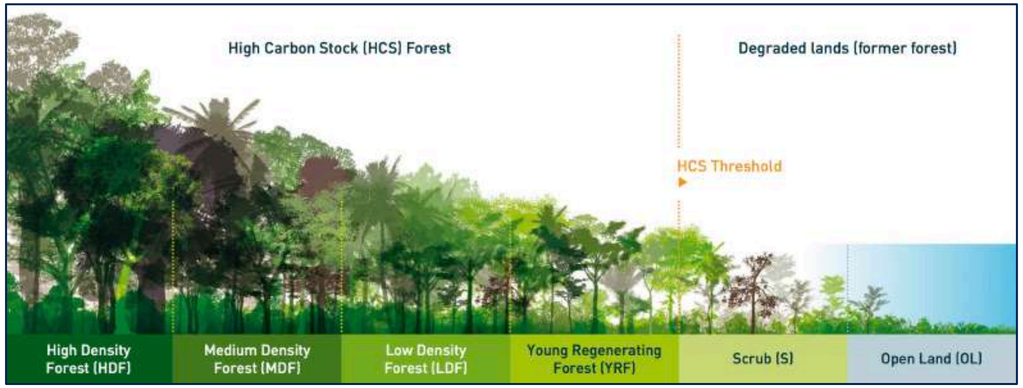
... reduce the amount of needed training data?

Exciting future directions:

Temporal analyses

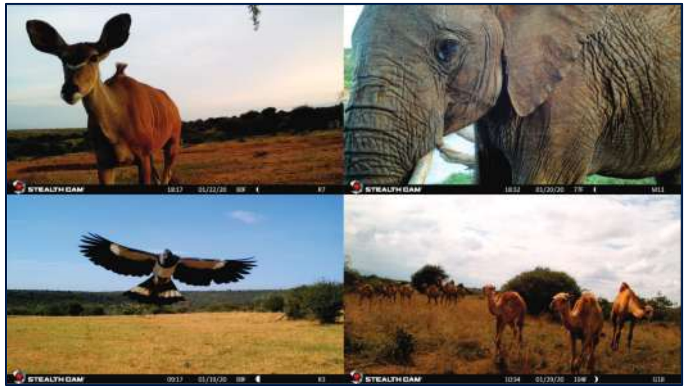


Biomass / Carbon stock estimation



highcarbonstock.org

Biodiversity modelling



Beery (2021)

Contact: Personal: langnico.github.io | Group: prs.igp.ethz.ch

References

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

N. Lang, N. Kalischek, J. Armston, K. Schindler, R. Dubayah, and J. D. Wegner. Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. Remote Sensing of Environment, 268:112760, 2022b.

N. Lang, K. Schindler, and J. D. Wegner. High carbon stock mapping at large scale with optical satellite imagery and spaceborne LIDAR. arXiv preprint arXiv:2107.07431, 2021.

N. Lang, W. Jetz, K. Schindler, and J. D. Wegner. A high-resolution canopy height model of the Earth. arXiv preprint arXiv:2204.08322, 2022a.

Contact: Personal: langnico.github.io | Group: prs.igp.ethz.ch