GRÜNBLICK

AI Powered Forest Biomass Estimation Service

Dr. Vytautas Jančauskas, Dr. Kalifou Rene Traore, Juan Pablo Espejo Belmonte, Dr. Daniela Espinoza-Molina

Data Science for Earth Observation Department, Earth Observation Center, German Aerospace Center

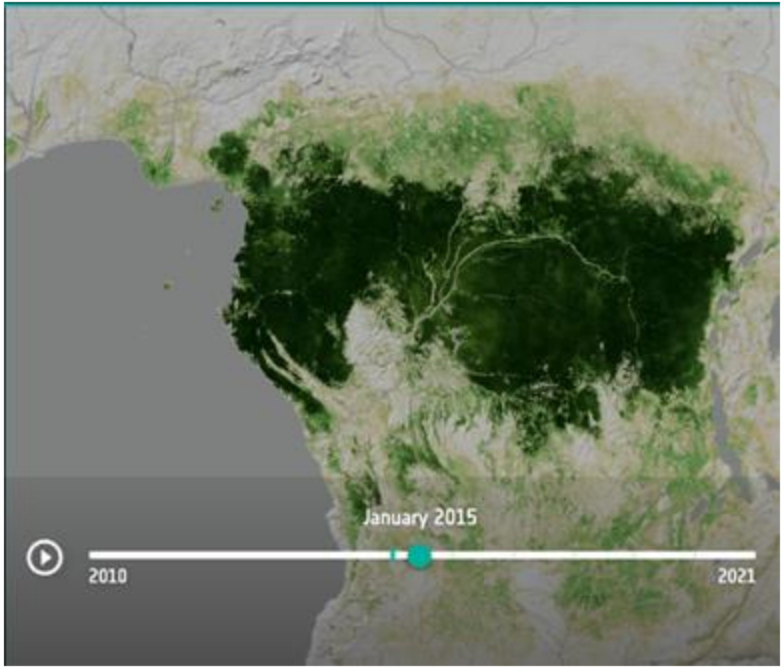
DLR

Grünblick – AI-Powered Forest Biomass Estimation

- Accurate biomass maps are essential for climate, biodiversity, and sustainable forestry
- Combines **multi-sensor Earth Observation** (Sentinel-1 SAR & Sentinel-2 MSI)
- Uses **deep learning (U-Net)** for pixel-level above-ground biomass (AGB) estimation
- **Key features:** multi-modal fusion + self-supervised pretraining → higher accuracy
- Modular, scalable pipeline with training, inference, and interactive user service
- Validated on **BioMassters (Finland)** and extended to **British Columbia**
- **Next steps:** uncertainty quantification + global deployment

Introduction Part 1

- Above-Ground Biomass (AGB) is the total mass of all living plant material above the soil surface, including stems, trunks, branches, bark, seeds, and leaves. It explicitly excludes below-ground parts such as roots.
- Forest biomass estimation is crucial for climate change mitigation, biodiversity assessment, sustainable forest management, and supply chain monitoring.
- AI and Big Data from satellites are transforming the field, combining remote sensing, machine learning, and ecological modeling.

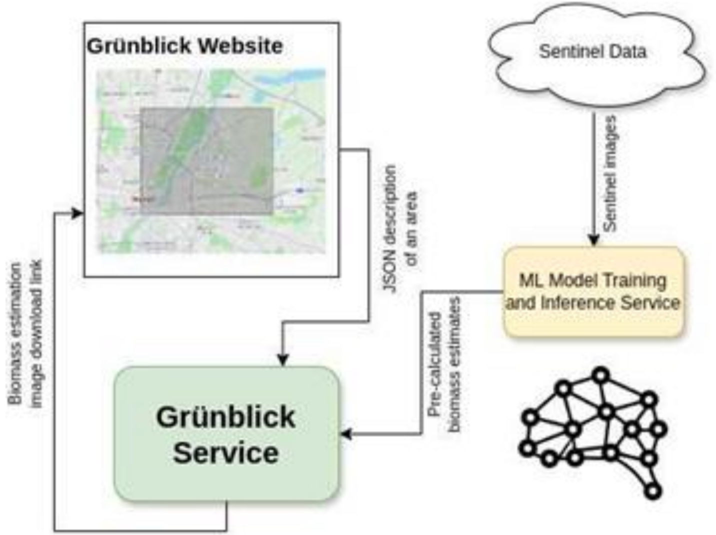


Introduction Part 2

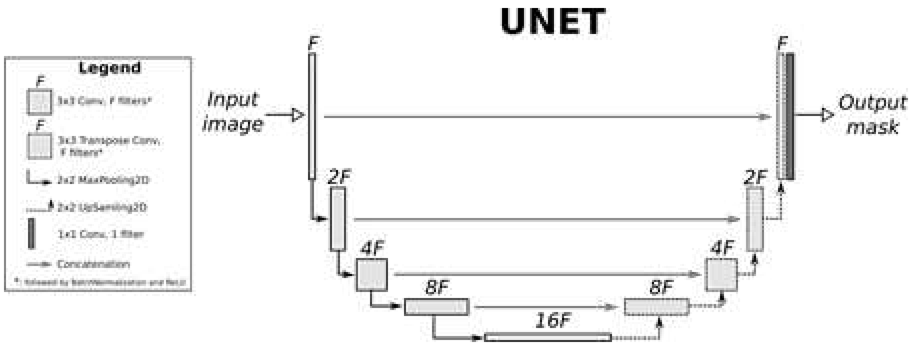
- Recent studies show improved above-ground biomass (AGB) estimation through multi-source satellite fusion and deep learning.
- Existing initiatives (ESA CCI, FAO Open Foris, BioMassters) provide valuable datasets but often lack resolution or flexibility compared to AI-powered systems.
- Grünblick is introduced as an AI-driven service that integrates Sentinel-1/2 data and deep learning for precise, scalable forest biomass estimation.

Grünblick Pipeline: Grünblick Overview

- Biomass estimation service under development at DLR
- Goal: continuous, large-scale forest biomass monitoring
- Based on Sentinel-1 SAR & Sentinel-2 MSI data
- Designed to be scalable & adaptable to new data streams



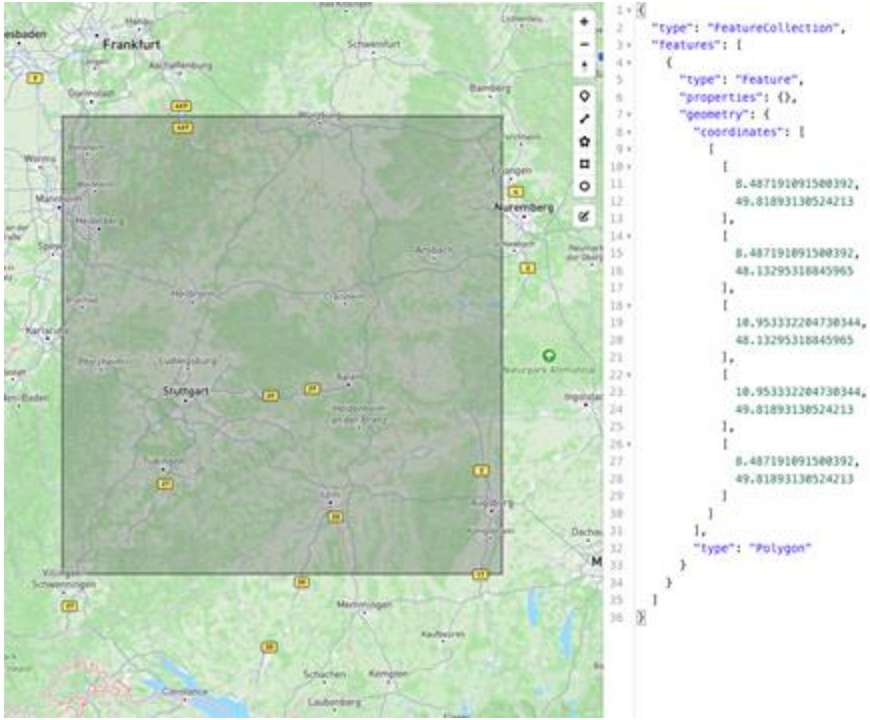
Grünblick Pipeline: Pipeline Architecture



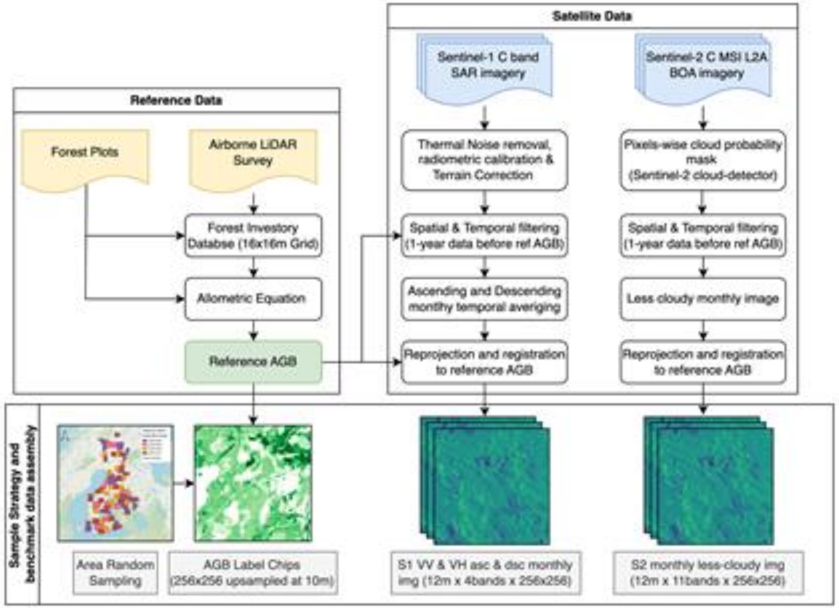
- Two main components:
- ML Model Training & Inference (biomass prediction models, Python, CI/CD, GEE data access)
- Front-end (caching, visualization, delivery)
- Models trained on co-registered imagery + ground-truth AGB
- Operational deployment with geographic cache for efficiency

Grünblick Pipeline: User Interaction

- Users submit JSON request specifying area of interest
- System outputs raster biomass map for that region
- On-demand delivery through interactive service
- Supports decision-making in forestry & related sectors



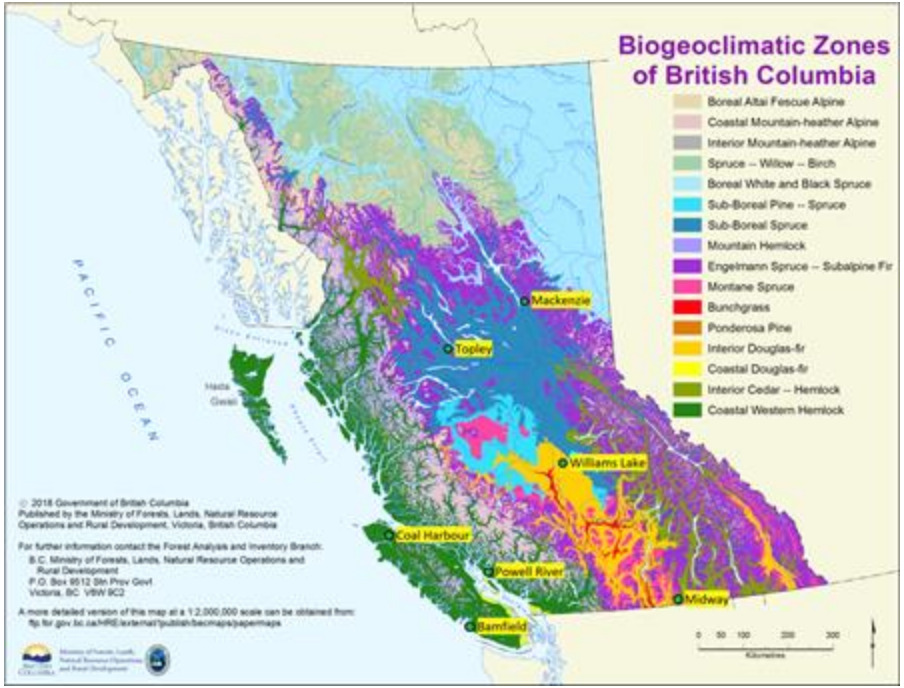
Datasets: BioMassters



- Public benchmark for AGB estimation (Finland, 2016–2021)
- Derived from LiDAR-based forest inventory plots
- Pixel-level biomass components: stem, branch, foliage, bark
- ~13,000 samples (2560×2560 m patches, 10 m resolution)
- 310,000+ paired satellite–ground truth samples
- Stratified sampling avoids temporal bias
- Provides rich ecological labels, standardized format, public availability
- Role: benchmarking & transferability for Grünblick

Datasets: British Columbia

- Region chosen for ecological heterogeneity and diverse forest types
- Satellite data:
  - Sentinel-2 MSI (B2–B4, B8, B11/B12 bands)
  - Sentinel-1 SAR (all-weather structural info)
- Ground truth: BC Vegetation Resources Inventory (VRI)
- 5.9M forest stands, ~5.5 GB spatial data
- Detailed biomass attributes
- Enables testing model generalizability in complex environments
- Supports future scaling of Grünblick to diverse ecosystems



Biomass Estimation Methodologies (I)

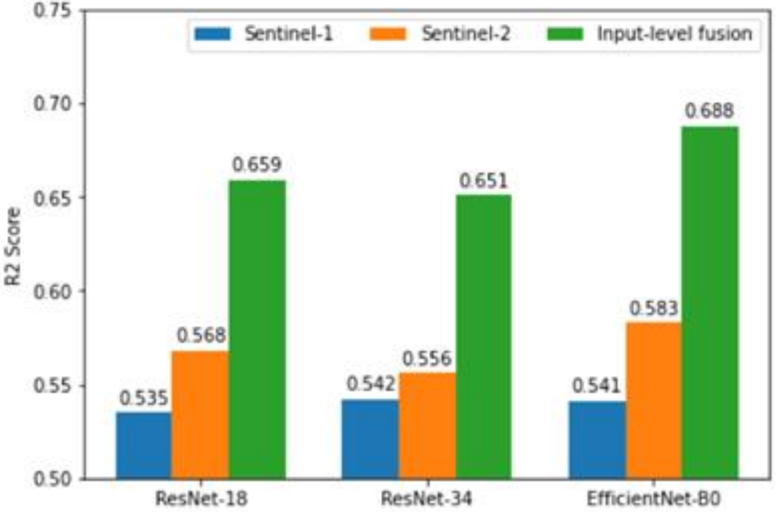


Fig. 2. Performance of U-Net regressors using different sensor inputs and backbones. Models trained for 50 epochs on the BioMassters dataset.

- Pixel-wise AGB regression with U-Net architecture
- Flexible encoder backbones: VGG, ResNet, EfficientNet
- Supports single-sensor inputs (SAR or MSI)
- Multi-sensor fusion (Sentinel-1 + Sentinel-2) for richer feature representation
- Transfer learning with self-supervised pretraining (SSL4EO-S12)

Biomass Estimation Methodologies (II)

- Models predict peak AGB for given timestamps (mono-temporal)
- Channel-wise stacking of SAR + MSI enables improved accuracy
- Pretrained weights improve results, especially with limited labeled data
- Future work:
  - Uncertainty quantification for confidence intervals
  - Multi-temporal estimation for dynamic forest monitoring

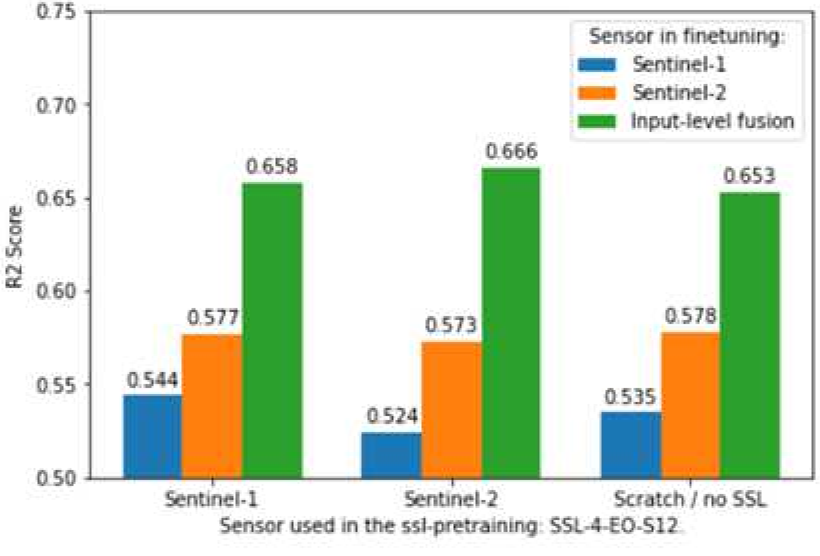


Fig. 3. R^2 for U-Net regressors (ResNet-50) trained with random or SSL-initialized weights.

Results

- Multi-sensor fusion (S1 + S2) outperforms single-sensor inputs
- Example (EfficientNet-B0): R^2 improved from ~0.54–0.58 → 0.69
- Efficient architectures (EfficientNet-B0) can match or beat deeper ResNets
- Self-supervised pretraining (SSL) improves accuracy, especially when modality aligns
- Fusion: Random init $R^2 = 0.65$ → SSL R^2 up to 0.67
- Overall: Grünblick achieves higher accuracy with fusion + SSL, without added complexity