Forest Stand Volume Estimation by Species from Sentinel-2 and LiDAR Data Using Regression Models

Gunta Grūbe
Institute of Electronics and Computes Science (IECS)
Email: gunta.grube@edi.lv

05.10.2022.

Baltic Satellite Service
NACIONĀLAIS ATTĪSTĪBAS PLĀNS 2020
EIROPAS SAVIENĪBA
Eiropas Reģionālās attīstības fonds
IEGULDĪJUMS TAVĀ NĀKOTNĒ
ELEKTRONIKAS UN DATORZINĀTŅU INSTITŪTS
INSTITUTE OF ELECTRONICS AND COMPUTER SCIENCE

Outline

- Motivation
- Main workflow and used data
- Workflow for forest stock estimation
- Preparing and exploiting regression model
- Results of estimated forest stock volumes
- Conclusions

Why it's essential to estimate forest stock volume?

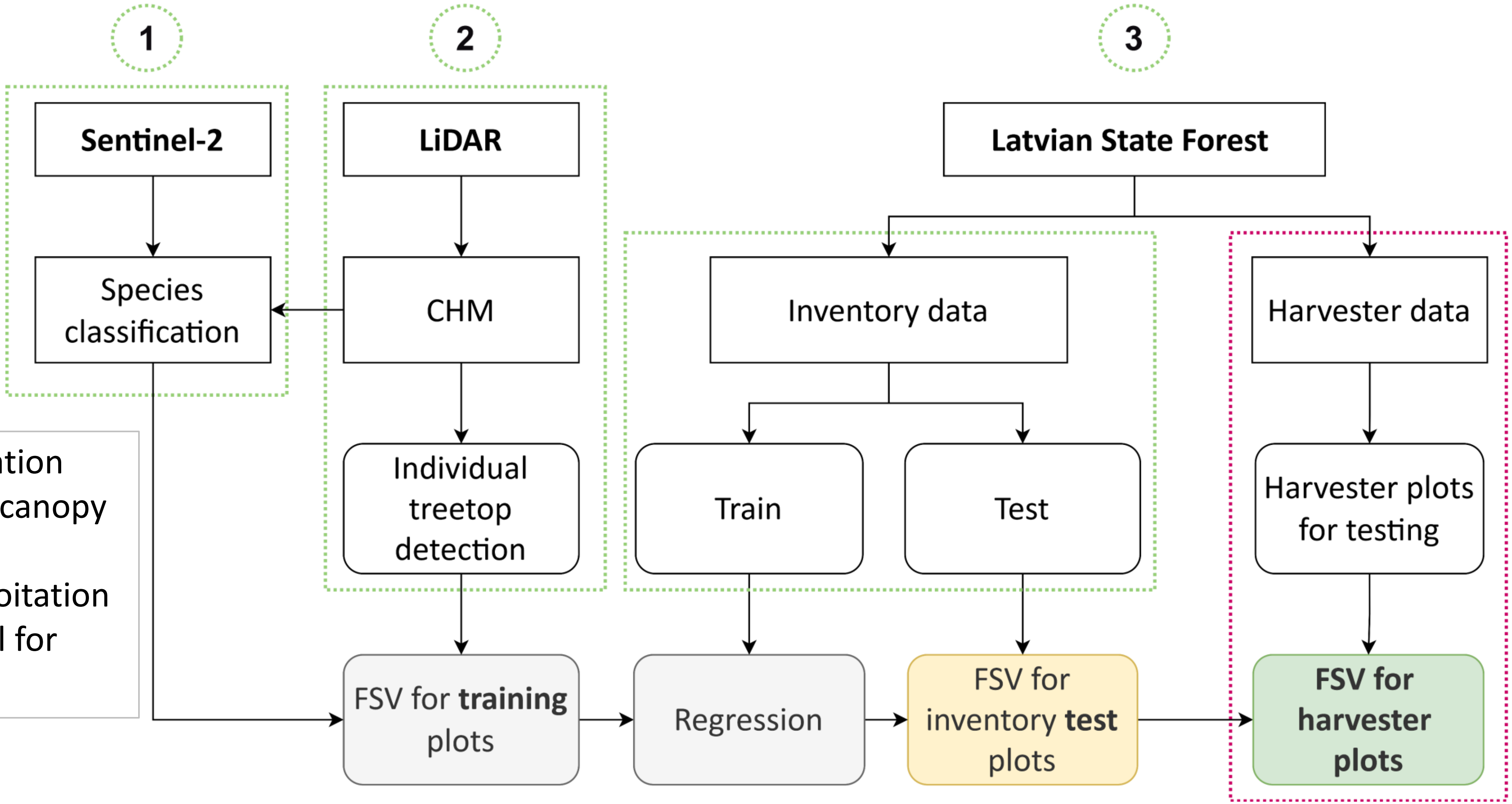
- Forest inventory has so far relied on field surveys
- Relatively expensive
- Requires a manual collection of field data

Why it's essential to estimate forest stock volume?

<u>Our approach</u> offers to determine the tree species, tree height, and stock volume of forest stands using only Satellite and LiDAR data.

Less expensive and faster then traditional methods!

The main workflow for forest stock estimation



1) Tree species classification
2) Identification of tree canopy peaks
3) Preparation and exploitation of a regression model for each species.

General workflow of the forest stock volume (m3/ha) estimation method

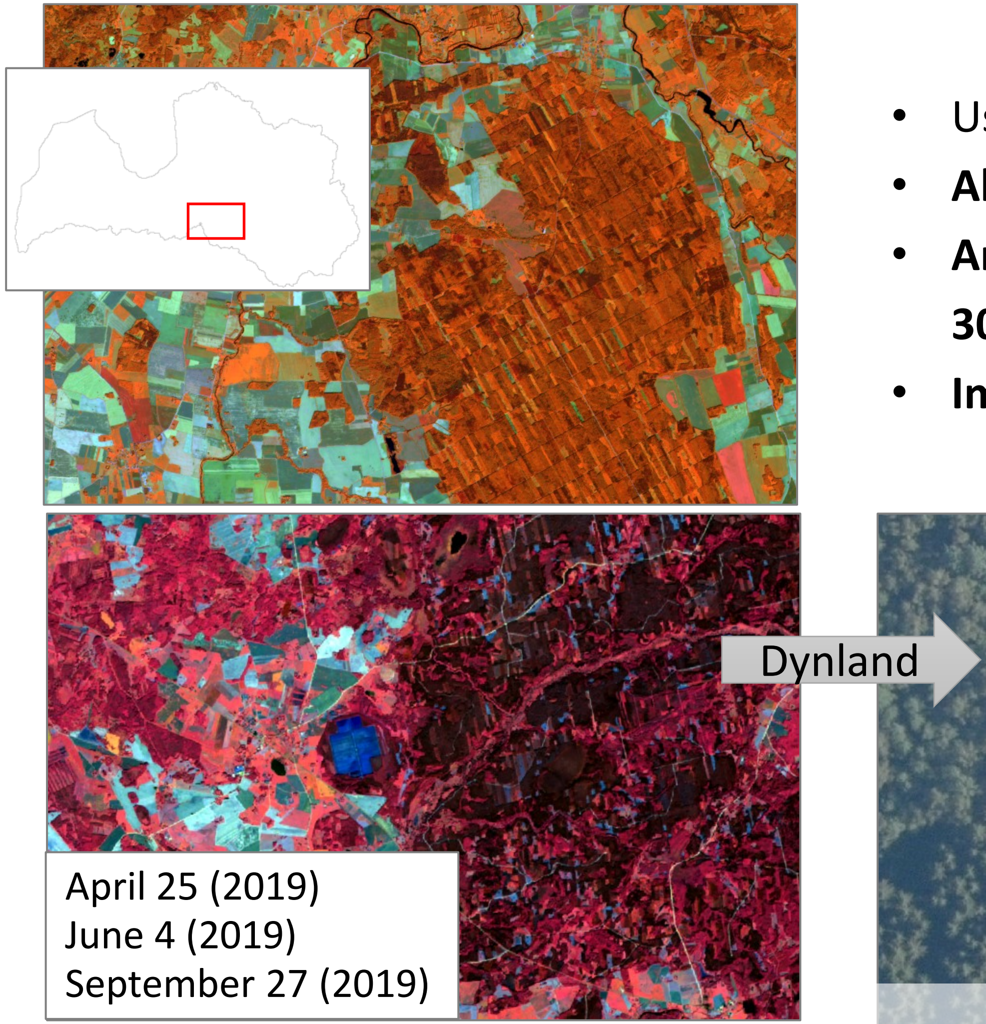
*FSV – forest stock volume (m3/ha)

ELEKTRONIKAS UN DATORZINĀTNU INSTITŪTS | INSTITUTE OF ELECTRONICS AND COMPUTER SCIENCE

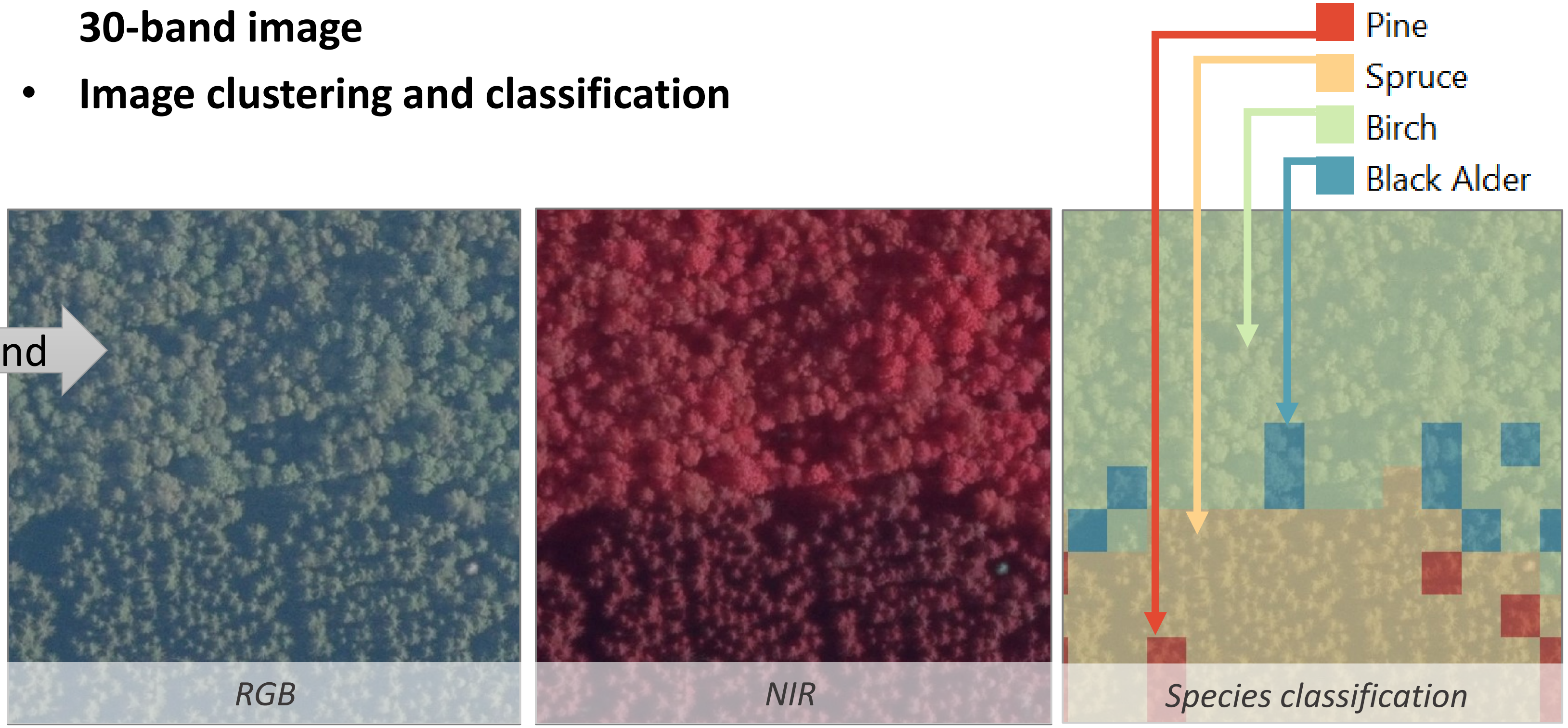
Sentinel-2 multispectral images

Dynland

- Usage of multispectral images with 10 spectral bands
- **All the bands were resampled to the 10-m resolution**
- **And the images were combined to process them together as a 30-band image**
- **Image clustering and classification**



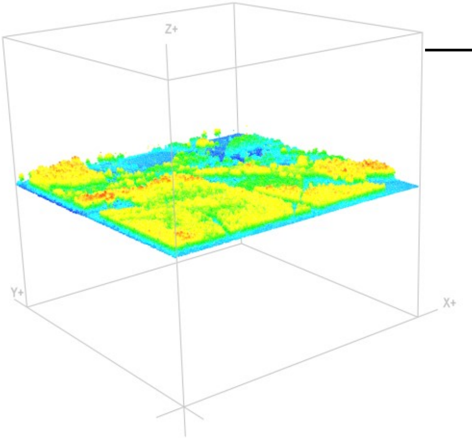
Sentinel -2 multispectral images for the study area located in the Zemgale region of Southeast Latvia



Visualization of tree species classification using *Dynland*

Light Detection And Ranging (LiDAR) data I

- **LiDAR data** were obtained from the Latvian Geospatial Information Agency
- The data of the study site were obtained in 2015 and 2017
- Classified point clouds were used for DTM and CHM creation
- We chose the individual tree detection algorithm of FUSION CanopyMaxima



LiDAR point clouds (.las)

↓

Point cloud classification

↓

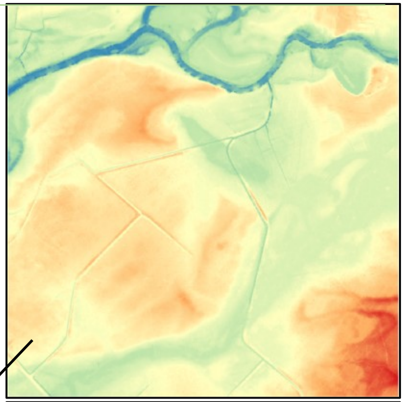
Spatial resolution 1 m → Digital terrain model (DTM)

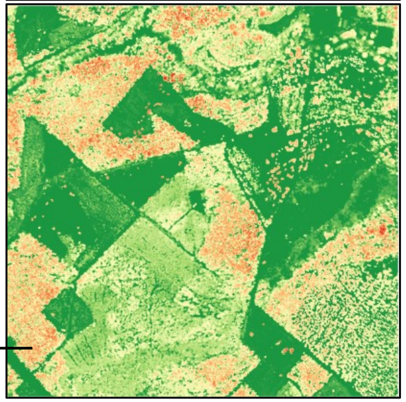
↓

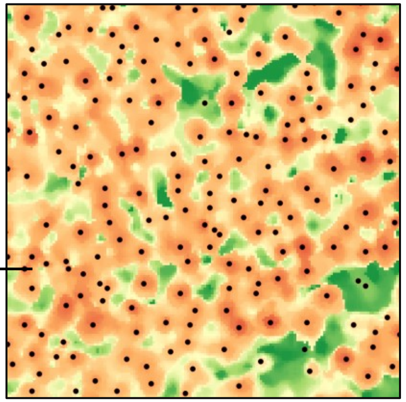
5x5 median filter → Spatial resolution 0.5 m → Canopy Height Model (CHM)

↓

CanopyMaxima tree detection







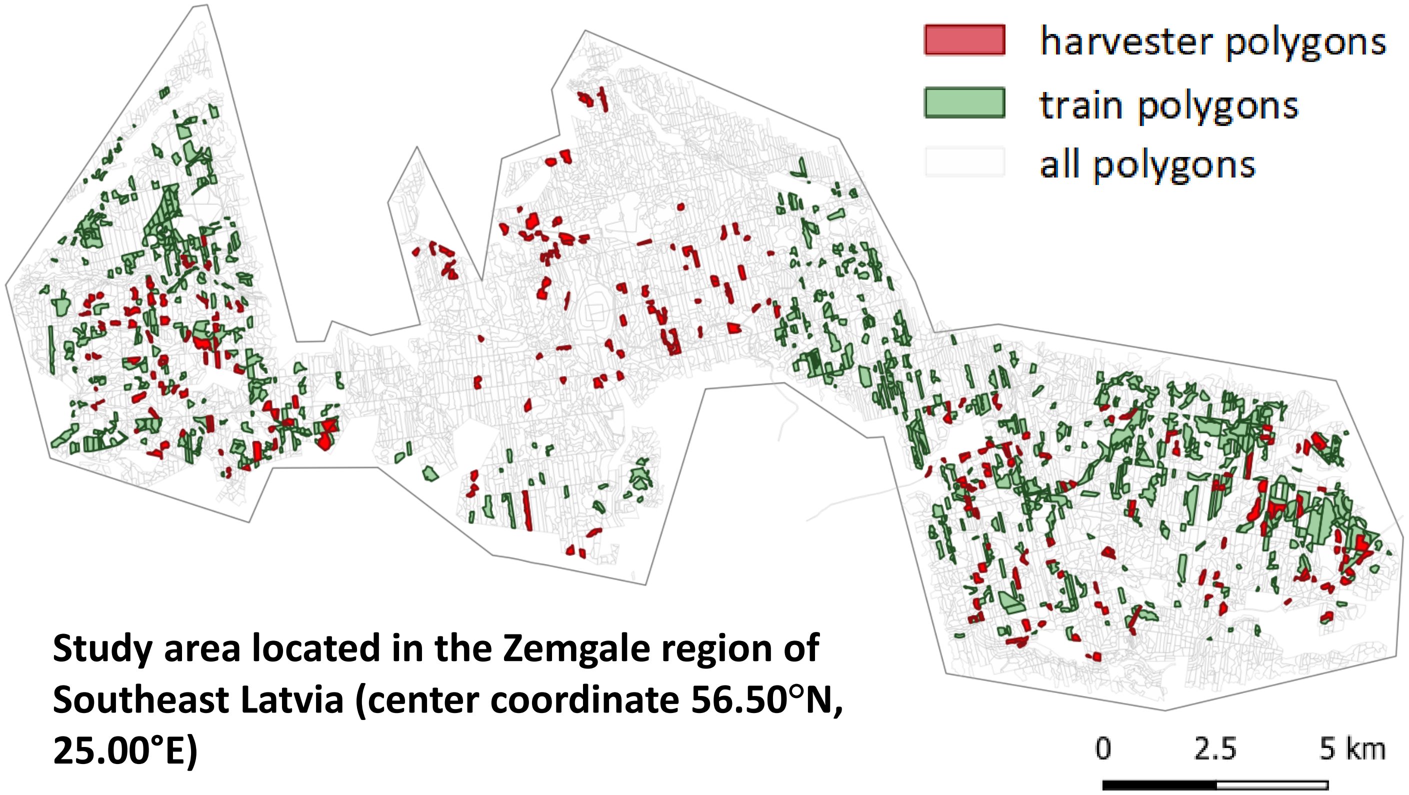
The result is individual trees with X, Y coordinates and height (m)

Latvia's State Forests inventory and harvester data

- Regular Forest inventory data was filtered by stand coefficient, density, and age group
- 971 forest stands (with dominant sp. > 80%) were used for **training**
- The approach proposed assumes that some sparse and partially imprecise forest inventory data
- 278 clear-cut harvester polygons (2020/2021) were used for **validation**

	Number of polygons used	
	Training	**Harvester**
Pine	660	80
Spruce	108	146
Birch	77	30
Black Alder	118	220
Total	**971**	**278**

- Different acquisition times
- Additionally filtered clear-cut forest stands (clearcutting mask)



Study area located in the Zemgale region of Southeast Latvia (center coordinate 56.50°N, 25.00°E)

Preparing and exploiting regression model I

- We used 971 inventory plots to prepare a regression model.
- Before performing the regression, each inventory plot's **basal area (m2)** and **forest stock volume (m3/ha)** were computed as follows [1, 2].
- **Calculated for each species separately** (dominant inventory species [S10]).

$$G = 0.7854 \cdot \left(\frac{H}{100}\right)^2 \cdot N$$

$$V = \frac{k \cdot G \cdot (H + 4)}{A}$$

G - basal area (m2)
H - average tree height (m) used instead of diameter breast height (DBH)
N - number of trees

V - forest stock volume (m3/ha);
k - species-specific coefficient (pine: 0.390, spruce: 0.415, birch: 0.385, black alder: 0.400)
A - the area of the plot (ha).

[1] *G. Prieditis, I. Smits, I. Arhipova, S. Daais, D. Dubrovskis, "Allometric Models for Predicting Tree Diameter at Breast Height," Energy Syst. Sustain, 4, pp.105-110, 2012.*
[2] *Liepa. Pieauguma maciba (Increment Science). Jelgava, LLU, 123 p, 1996.*

Preparing and exploiting regression model II

We used a 2nd-degree polynomial relationship (PR) between response and predictor variables:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1^2 + \beta_4 X_1 X_2 + \beta_5 X_2^2 + \varepsilon$$

Y - the dependent variable,
E(*Y*) - the expected value of *Y*
β_0 - intercept
$\beta_1, \beta_2, \ldots, \beta_5$ - regression coefficients of predictors X_1, X_2
ε - residual error.

Y – inventory first storey stock volume (m3/ha)

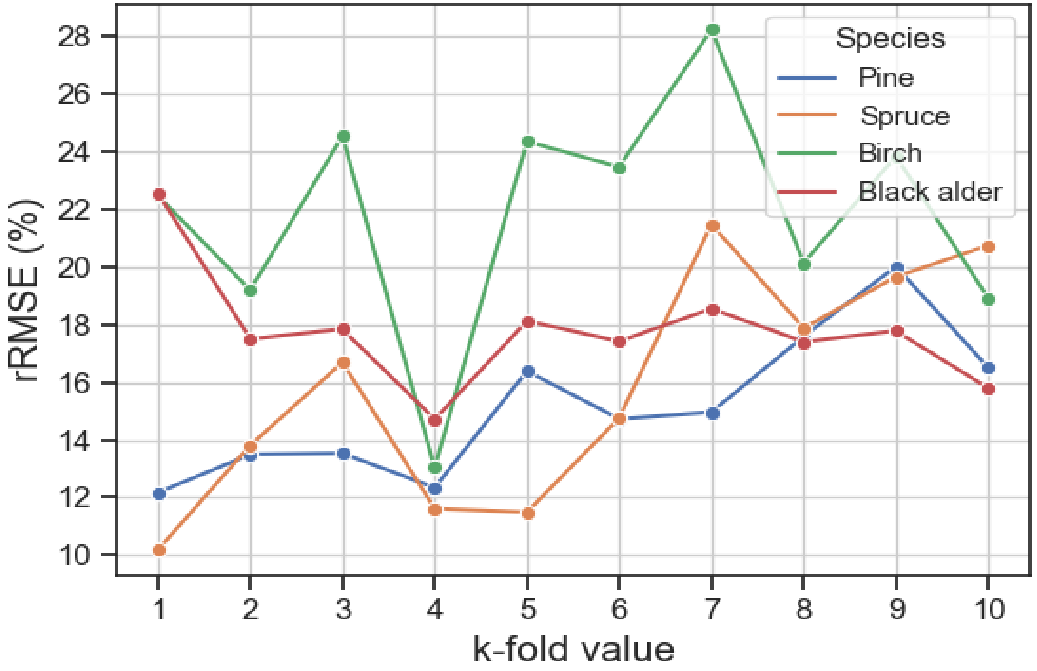
$\boldsymbol{X_1}$ – first storey stock volume (m3/ha) and

$\boldsymbol{X_2}$ – first storey average height obtained from individual trees and LiDAR CHM

Species	Polygon count	Regression Coefficients					
		β_0	β_1	β_2	β_3	β_4	β_5
Pine	660	-116.45	1.80	11.40	0.002	-0.07	0.19
Spruce	108	-220.24	1.30	27.25	0.002	-0.05	-0.16
Birch	77	140.74	1.15	-9.19	-0.001	-0.02	0.52
Black Alder	118	858.76	-1.31	-58.27	0.001	0.05	1.45

K-Fold Cross-validation of a regression model

- Cross-validation to compare results from different groups of predictor variables.
- 1382 selected plots (filtered by age group, stand density, and species coefficient) were **cross-validated against inventory forest stock volume**

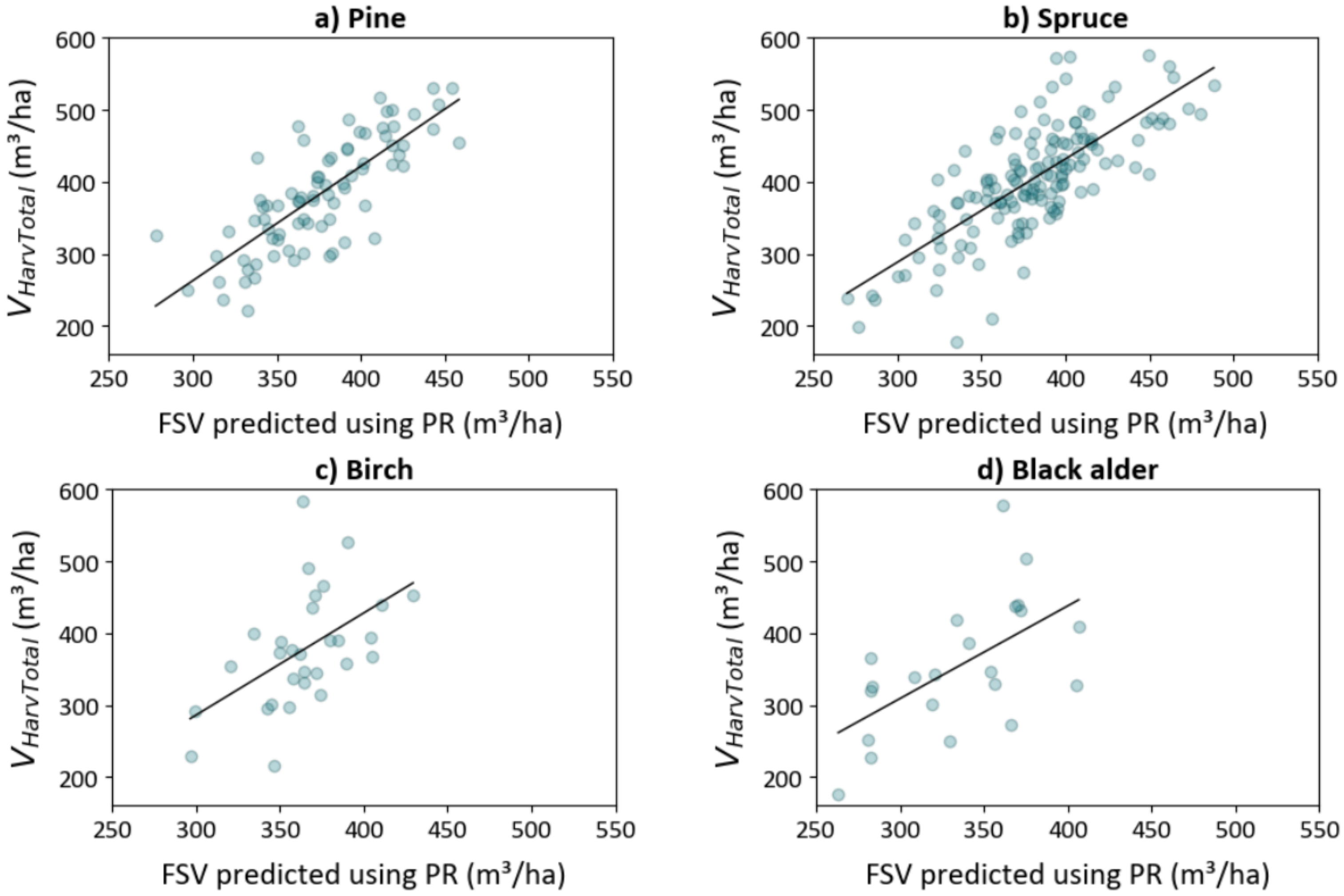


Sensitivity of the rRMSE values to different 10-fold values fitted with the PR method

	kFold value	
	5	**10**
Pine	13%	15%
Spruce	17%	16%
Birch	20%	22%
Black Alder	18%	18%

The sensitivity of the rRMSE values for 5 and 10-fold values fitted with the PR method

Results of estimated forest stock volumes I



Scatterplots of a total harvester (V_{Harv_Total}) vs total predicted FSV (m3/ha) using PR in test polygons, by dominant species

The FSV results of the *RMSE (m3/ha)*, and *rRMSE* that were computed by the PR model

Species	Count	RMSE	rRMSE
Pine	80	52	**14%**
Spruce	146	59	**15%**
Birch	30	71	**19%**
Black Alder	22	77	**22%**

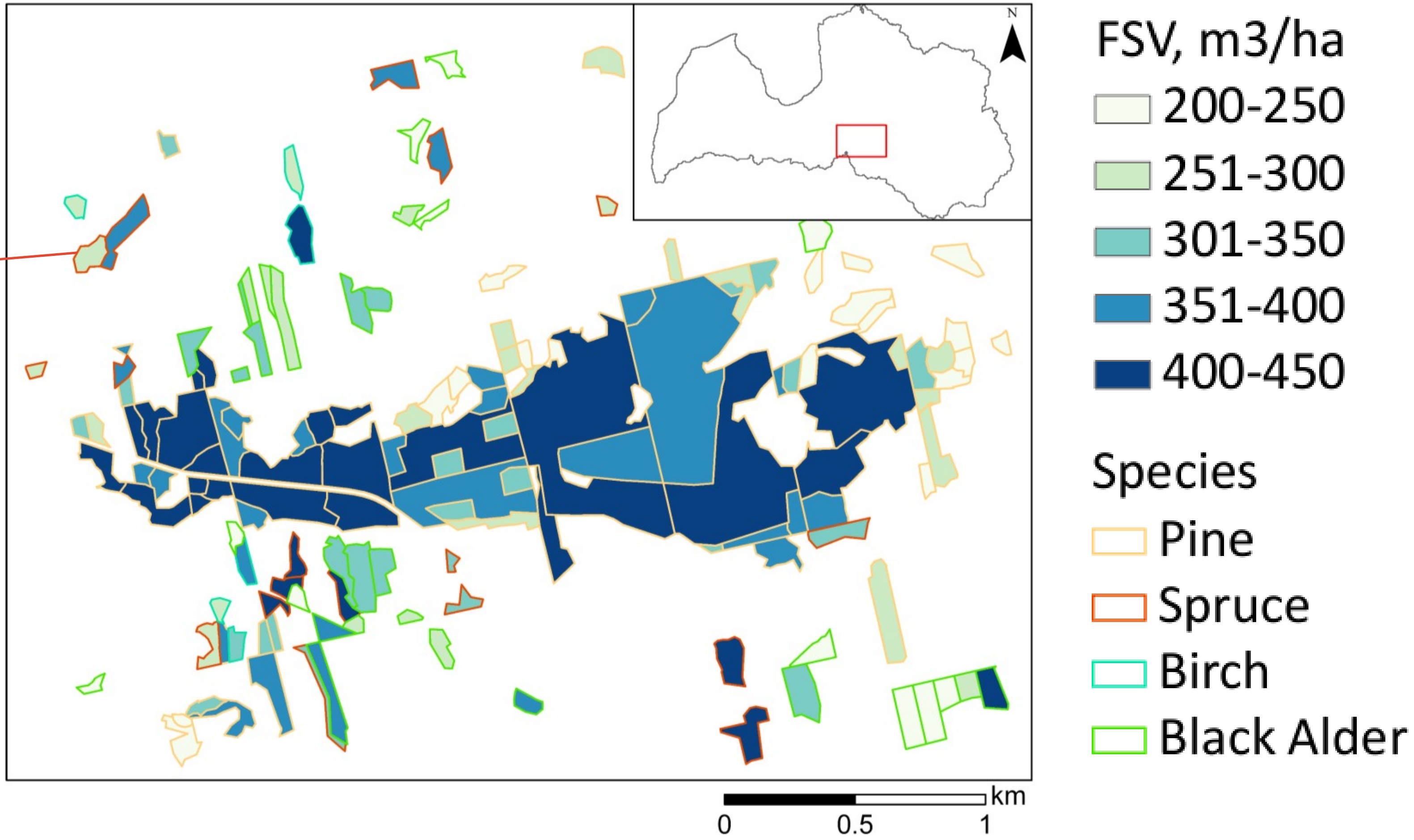
- The results of the correlation, RMSE, and rRMSE were computed by the PR method for 278 harvester polygons.
- Most accurate for pine and spruce.
- Assessed pure stands.

Results of estimated forest stock volumes II

Example:

- Species: **Spruce**
- Estimated forest stock volume: **298 m3/ha**

(harvester FSV: 310 m3/ha)



FSV (m3/ha) prediction map derived from LiDAR CHM, Sentinel-2 multispectral images using sparse and outdated forest inventory data

Conclusions

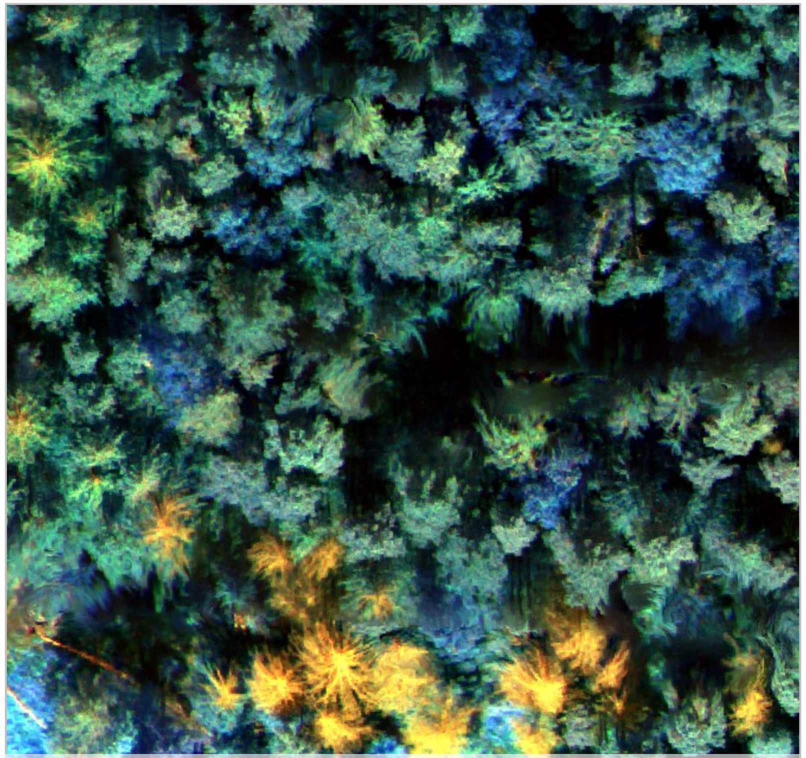
Our approach is based on the recently proposed *Dynland* algorithm, identification of tree canopy peaks, and using a regression model for stock volume estimation.

Dynland

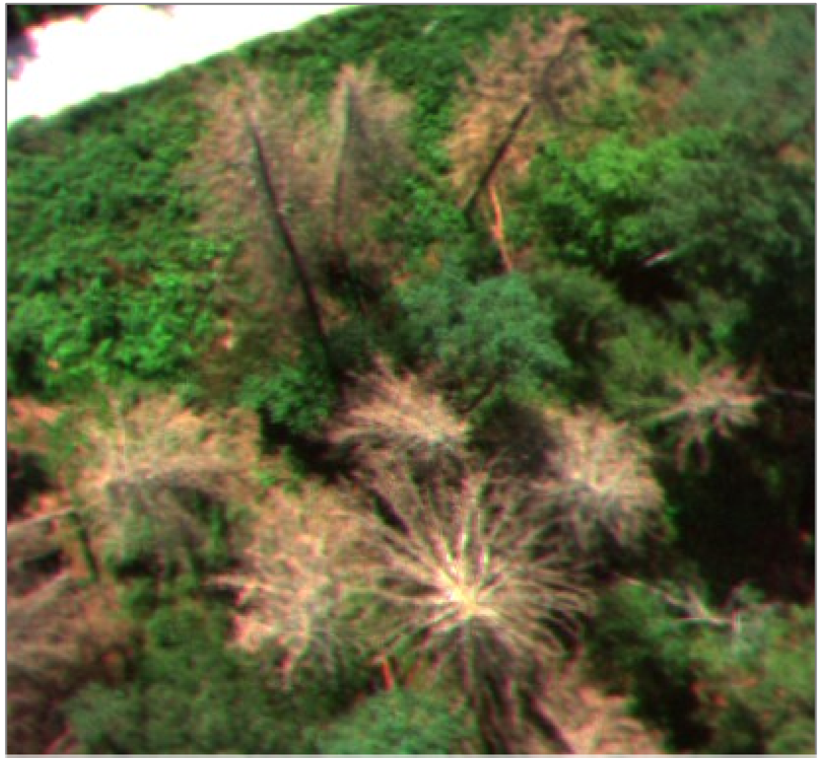
- This study demonstrates **forest structural attribute assessment** for four dominant species (pine, spruce, birch, and black alder).
- Our experiments showed that a lower *RMSE* can be achieved using **polynomial regression.**
- The proposed approach facilitates the estimation of other inventory parameters of forest stands separately for each species of interest.

Further work

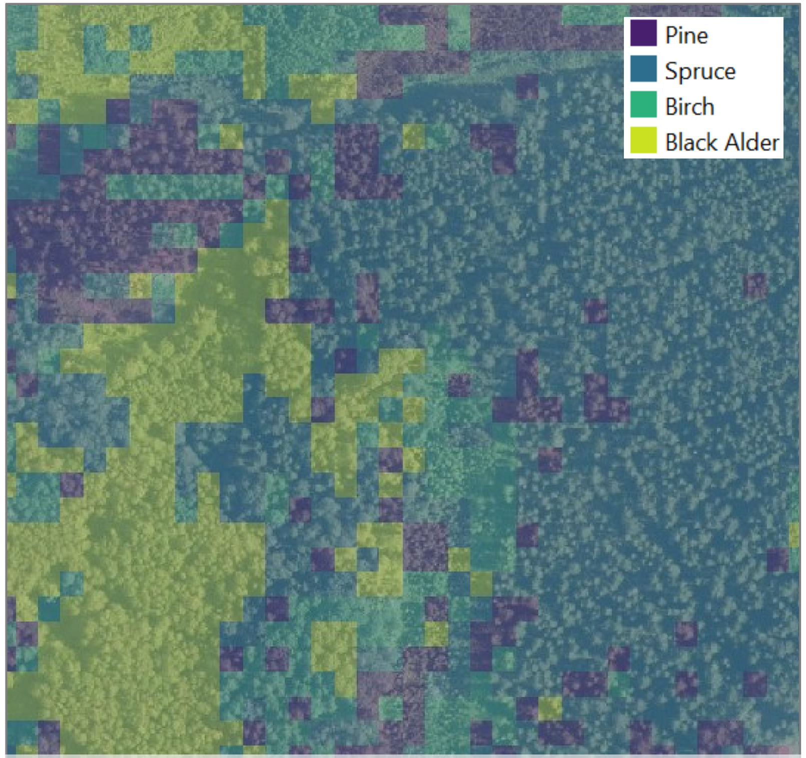
- ✓ Will include an examination of the potential of using **UAV-collected data** for forest inventory.
- ✓ The potential of **classification and FSV estimation for additional species** should also be explored.
- ✓ Stock volume estimation on larger territories **in Latvia, and Sweden**.



Drone-based hyperspectral image



Drone-based RGB image



Dynland species classification

Thank you!

Questions!

Visit and contact www.edi.lv
Gunta Grūbe
gunta.grube@edi.lv

ERDF-funded project entitled "Satellite remote sensing-based forest stock estimation technology (WoodStock)" (Project No. 1.1.1.1/18/A/165) and "Remote sensing based system for forest risk factor monitoring (Forest Risk)" (Nr.1.1.1.1/21/A/40)

ELEKTRONIKAS UN DATORZINĀTŅU INSTITŪTS
INSTITUTE OF ELECTRONICS AND COMPUTER SCIENCE

14 Dzerbenes St., LV-1006, Riga, Latvia