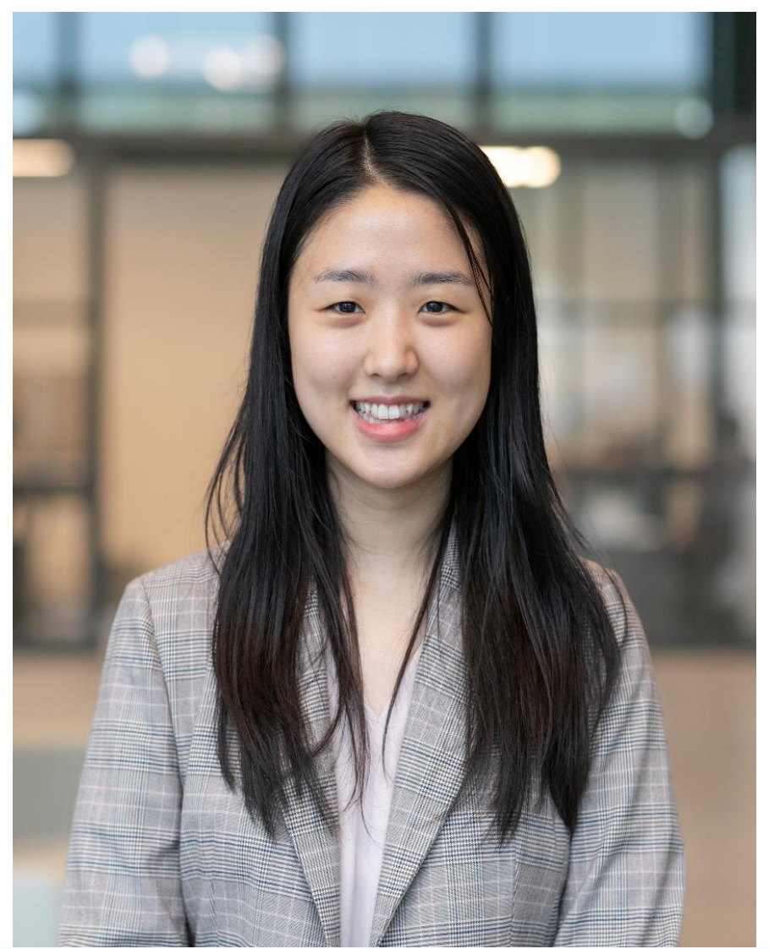
Contact

Yeji Lee (M.Sc Candidate)
Georg-August-Universität Göttingen
Ecosystem Analysis and Modelling

Email: yeji.lee0701@gmail.com
Linkedin →



Master Thesis (in Progress, Here for feedback & suggestion)

Automated Whorl Detection in Norway Spruce: Optimizing Clustering Hyperparameters on TLS Point Clouds

Yeji Lee, Thomas Hay

Background & Motivation

- **Branch whorls** — circular branch patterns — are biologically important indicators of tree growth and age.
- Detecting whorls from 3D tree scans (TLS data) offers a **non-destructive** alternative to manual methods.
- Manual annotation is **slow and error-prone**

➔ I propose an **automated pipeline** using **stem-projected point clouds** and clustering algorithms (Jarvis-Patrick clustering in progress; more methods planned)

Goal

: Identifying clustering algorithms and hyperparameters that best reflect human perception of whorls.

Materials & Tools

- Manually-segmented 25 highland trees and 16 lowland trees
- Quantitative Structure Model(QSM) generated with Rtwig
- Software: R, Python, CloudCompare

Preprocessing steps

(1) **Down-sample** to reduce noise & computation

(2) **Project** each point to its nearest stem cylinder (QSM) to get
- z: height from ground
- d: horizontal distance to stem

(3) Define a **donut-shaped region** (blue in Figure 1) to select **candidate whorl points**.

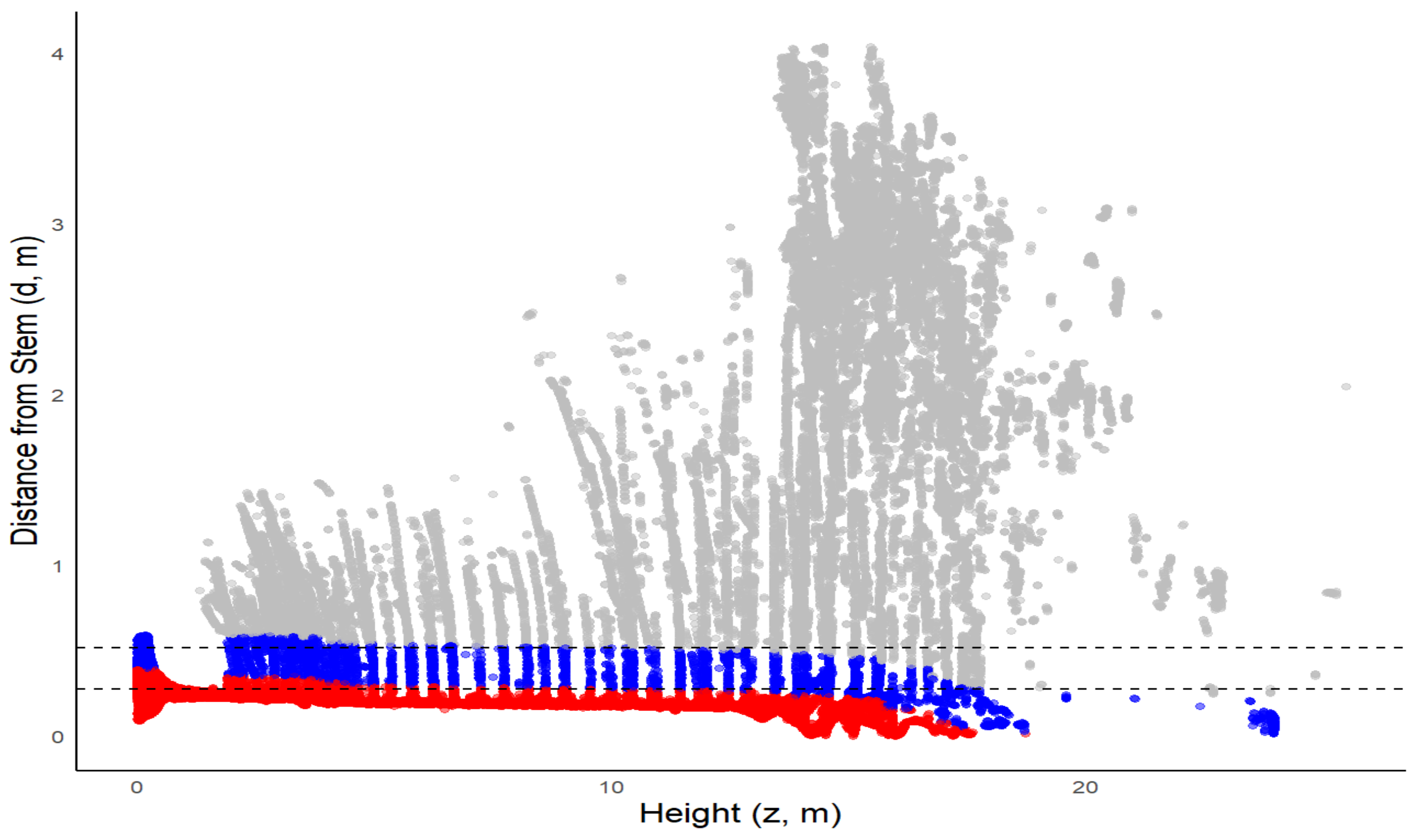


Figure 1. Donut shape Filtering based on Stem Distance (Gray: too far, Blue: In donut, Red: too near, Black dotted-line: mean value of near limit and far limit)

Whorl detection with JP Clustering

(1) Use JP clustering (dbscan::jpclust) in (z, d) space.

(2) Connect points that share enough **mutual nearest neighbors**.

(3) Key Parameters
- k: number of nearest neighbors
- kt: threshold of shared neighbors to form a cluster

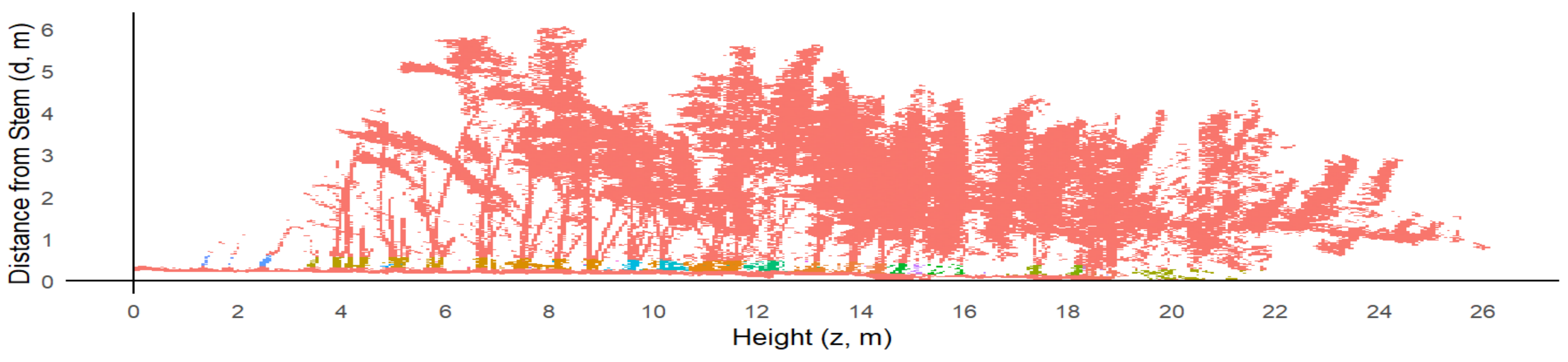


Figure 2. Clustered data plotted by height (z) and distance (d) from tree stem Different clusters are marked in different colors

Parameter Optimization

To match automatic whorl detection with manual references, I optimized the JP parameters (k and kt) using **GenSA (Generalized Simulated Annealing)**.

1. Ground truth (manual reference):

(1) Normalize branch point heights (ground = min z)

(2) Group branches into whorl using 20 cm vertical threshold

(3) Compute average height per group (whorl center)

(4) Extract
- Lowest whorl height
- Mean & SD of whorl spacing
- Number of whorls

2. Optimization:

(1) Run JP clustering with a (k, kt) pair

(2) Extract predicted whorl positions

(3) Calculate $RMSE = \sqrt{(\hat{d}_{whorl} - d_{ref})^2}$

(4) Penalize invalid (e.g., kt > k) or failed runs

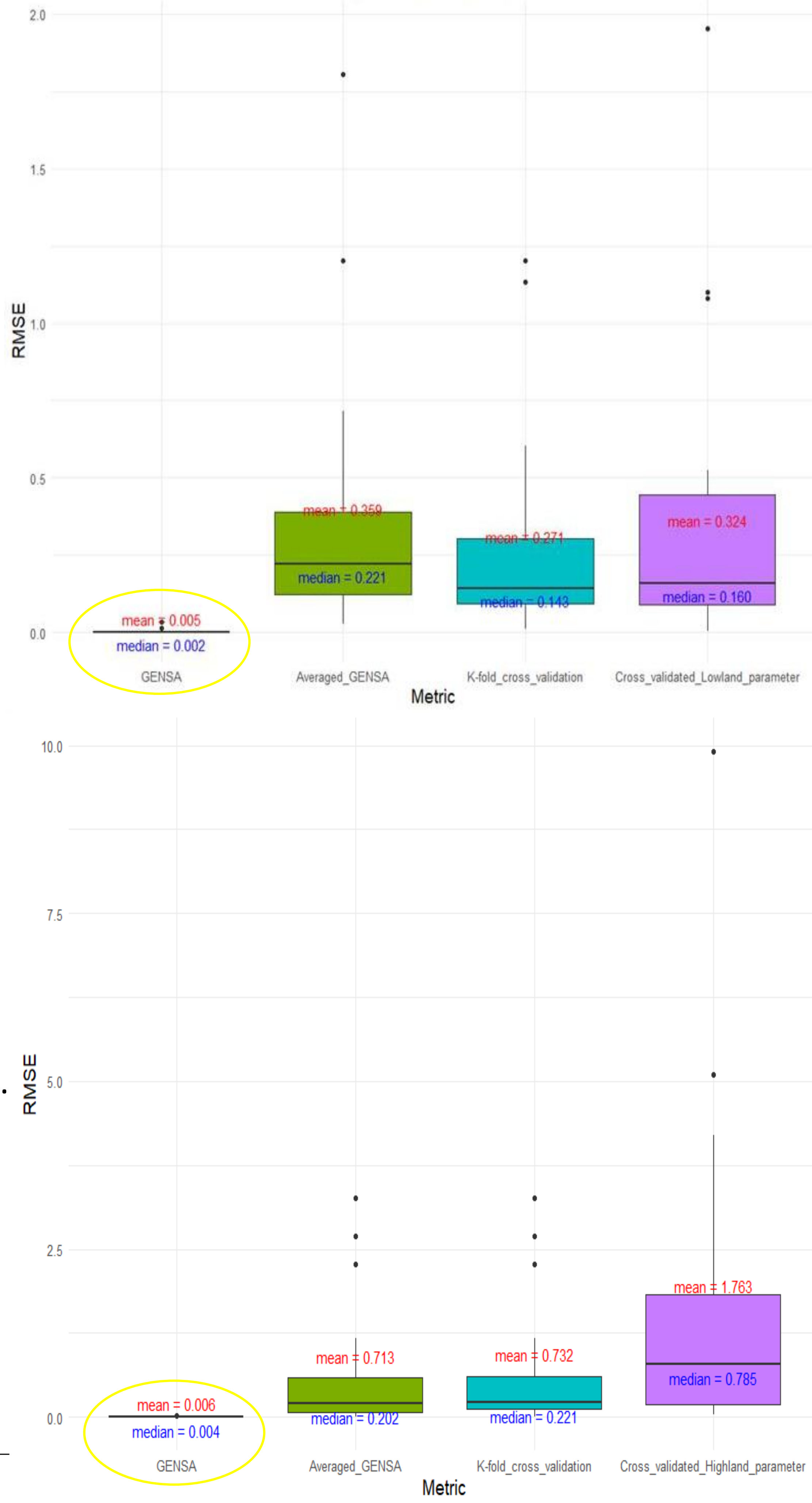


Figure 3. RMSE Boxplot by different JP Parameter Selection Approach (Top: Highland, Bottom: Lowland)

Open Questions

❖ What are the trade-offs between RMSE and clustering quality (e.g., silhouette) & how can they be jointly optimized?

❖ How sensitive are clustering-based whorl detections to QSM errors?

❖ Can biologically meaningful features (e.g., branch angle, length, diameter) improve clustering accuracy beyond spatial (z, d) information alone?

❖ What are the minimum point cloud density and quality thresholds required to achieve reliable whorl detection results?

Evaluation

(1) Generalization Test

- Compared Tree-specific VS Generalized JP parameters (averaged, k-fold, cross-site) as described in Figure 3.
- Tree-specific gives lowest RMSE → Best fit
- Generalized performs slightly worse but acceptable
- Cross-site yields higher RMSE → limited transferability

(2) Sillhoutte Index

- Measures cluster quality (–1 to 1), higher the better
- Most scores are below 0 (Figure 4) → weak separation
- Suggested parameters don't clearly distinguish tree groups

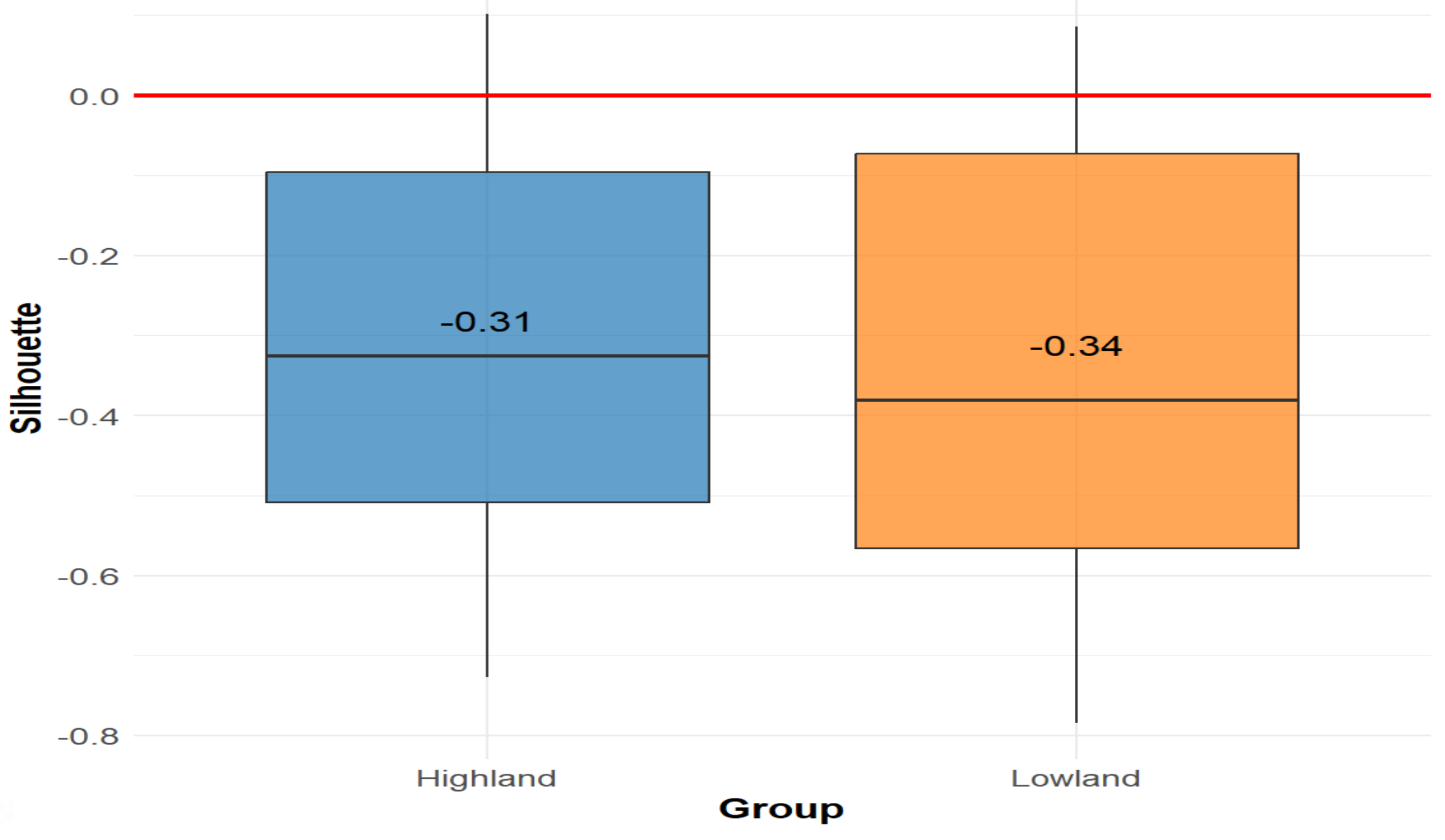


Figure 4. Silhouette values of JP parameter derived from GenSA

Moving Forward

Refining the clustering method or exploring alternative approaches could help improve the robustness and clarity of the detected whorls.

- Exploring **SNN (Shared Nearest Neighbor) clustering** as a more noise-robust alternative (in Progress)
- **Multi-Objective Differential Evolution**
- **NSGA-II** (Non-dominated Sorting Genetic Algorithm II)

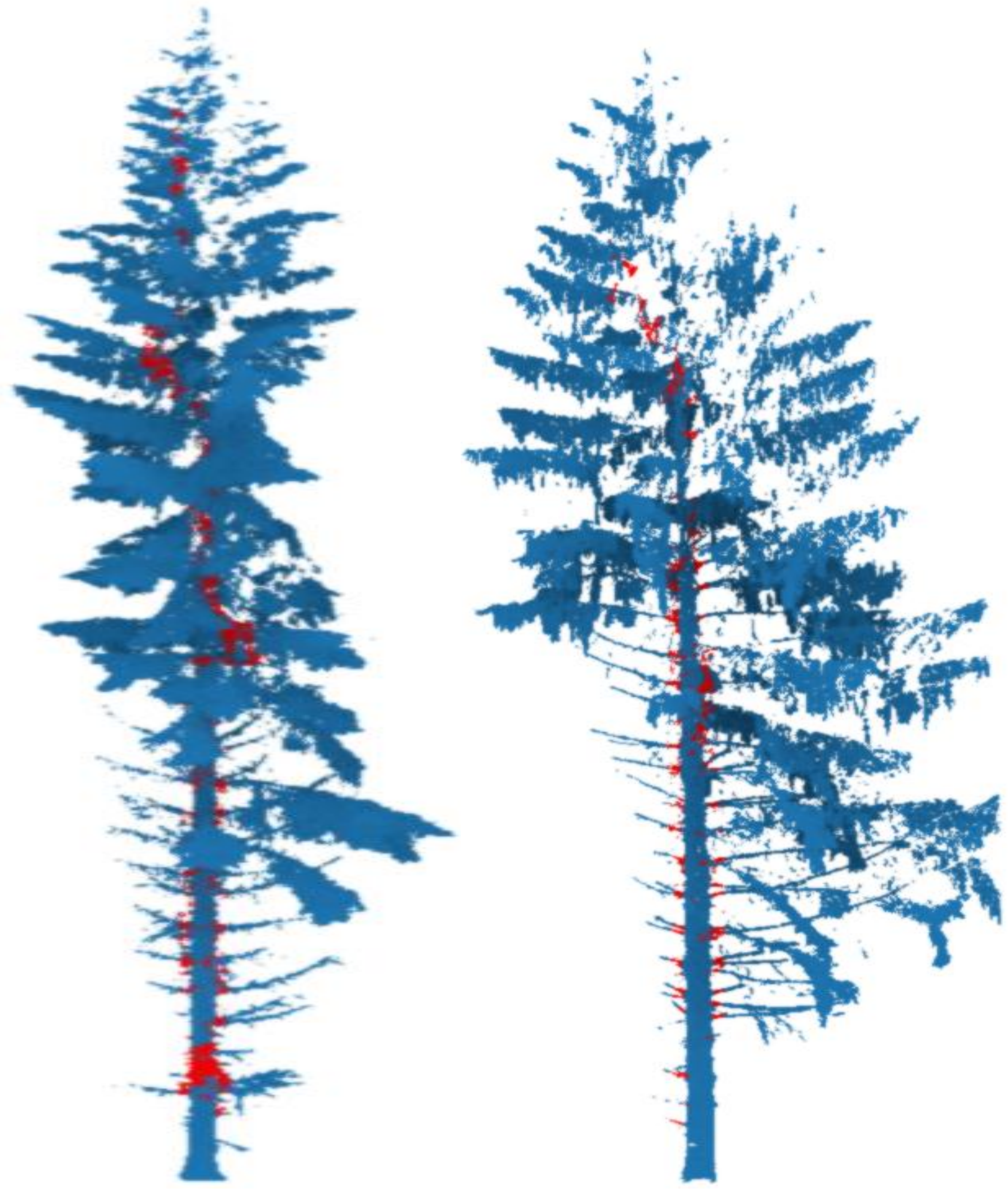


Figure 5. JP Clustered Tree with GenSA Parameter (blue: non-whorl, red: whorl, left: bad , right: good)

Interdisciplinary Summer School on Precision Forestry Using LiDAR Technology 2025