Data quality analysis after hyperspectral LiDAR sequentially mapping trees

Shao Dong[1] and Yi Lin[1, *]

[1]Institute of Remote Sensing and Geographic Information System, School of Earth and Space Sciences, Peking University, Beijing 100871, China – 2301210158@stu.pku.edu.cn,
* Correspondence yi.lin@pku.edu.cn

Keywords: Hyperspectral LiDAR, Data quality analysis, Time series analysis, PROSPECT-D, Sensitivity analysis.

Abstract

Light detection and ranging (LiDAR), as an innovative remote sensing tool, not only captures target reflectance but also provides its morphological parameters. Traditional single/multi-band LiDAR and multispectral LiDAR (MSL) are presently employed in applications such as 3D modeling and plant biochemical parameter inversion albeit with effectiveness limited. Moreover, hyperspectral LiDAR (HSL) distinguished by its expanded array of spectral detection channels and enhanced spectral resolution, has proven more effective in meeting these requirements and also exhibits superior capabilities in both feature and land cover classification tasks. Nevertheless, point clouds acquired through HSL frequently exhibit quality deficiencies, including uneven density and excessive noise. Meanwhile, there exists a notable absence of technical specifications and operational standards governing the measurement protocols for HSL systems globally. To address this gap, this study constructed a systematic analysis framework of data quality in hyperspectral point clouds and endeavors to qualitatively analyse 30 tree point clouds continuously scanned with Finnish Geospatial Research Institute (FGI) 8-band hyperspectral laser scanner. Furthermore, this research validated the theoretical feasibility of employing the 8-band HSL system for inversion processes aimed at quantifying chlorophyll leaf content. Apart from detecting the time-varying patterns of reflectance within birch canopy point clouds, the results of this study also effectively pinpointed the band exhibiting heightened noise level of the HSL system, demonstrating the efficacy of our proposed quality analysis methodology. The endeavor presented in this study can serve as a cornerstone for advancing hyperspectral LiDAR across a diverse array of related remote sensing and earth observation applications.

1. Introduction

According to the 2020 Global Forest Resources Assessment by the Food and Agriculture Organization of the United Nations, the world's forest area was 4,059 million hectares, constituting 31.1% of the land area. However, this proportion has been steadily declining over the years (FAO 2020). Trees, integral components of the global ecosystem carbon cycle (Liu et al., 2015), significantly influence the entire cycling process through their morphological structure and biochemical composition. Effectively obtaining these attributes of tree canopies has become a prominent research focus.

Currently, LiDAR stands as the mainstream method for comprehensively capturing the structural attributes of trees. It enables accurate ranging and localization, providing high-precision vertical structure information and reliable three-dimensional data on forests' canopy characteristics—information challenging to obtain through traditional hyperspectral remote sensing. LiDAR has evolved from multi-band LiDAR to multispectral LiDAR and hyperspectral LiDAR. LiDAR applications in plant structural attribute acquisition include single-tree segmentation (Shen and Cao, 2017), tree species classification (Sugumaran and Voss, 2007), biomass estimation (Luos et al., 2017; Ju et al., 2022), among other fields.

Hyperspectral LiDAR, as a novel earth observation technology, captures spectral intensity information of different bands on laser foot points, producing 3D hyperspectral point cloud data (Gong et al., 2021). This data type, distinct from passive hyperspectral image data and single-wavelength LiDAR point cloud data, facilitates biomass estimation and simultaneous acquisition of spectral and positional information for each laser scanning point (Luos et al., 2017; Ju et al., 2022). Nevertheless, the hyper-continuous spectrum laser pulses employed by hyperspectral LiDAR exhibit low energy, posing challenges in achieving hyperspectral fast imaging under low-light conditions. This limitation may lead to uneven point cloud density in the acquired data, thereby impacting the precision of subsequent analyses and modeling efforts (Uchida et al., 2020). Furthermore, intrinsic sensor-generated noise contributes to the point cloud, necessitating a comprehensive analysis of the acquired point cloud data to ensure data quality.

Presently, global research on hyperspectral LiDAR is experiencing rapid development, with a focus on two main directions: the advancement of prototype systems and the expansion of application functionalities (Lin et al., 2019). Notably, the Finnish Geospatial Research Institute introduced the world's first full-waveform hyperspectral LiDAR system in 2012, marking a significant milestone in remote sensing (Hakala et al., 2012). This system produces 3D point cloud data with spectral backscattered reflectance, extending the realm of imaging spectroscopy to include spectral 3D sensing. In 2014, Niu Zheng's team at the Aerospace Information Research Institute, Chinese Academy of Science, developed a 32-band hyperspectral LiDAR spanning 538.0 to 910.0 nm, representing the first international 32-band hyperspectral LiDAR system in China with fully independent intellectual property rights (Sun et al., 2014). Another 32-band hyperspectral LiDAR prototype, with a spectral resolution of 12 nm, created by Gong Wei's team at Wuhan University, has been widely applied in monitoring various nitrogen application levels in vegetation and chlorophyll inversion in vegetation (Du et al., 2016; Sun et al., 2018).

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-1-2024-49-2024 | © Author(s) 2024. CC BY 4.0 License.

However, the effective utilization of data collected by hyperspectral LiDAR systems necessitates thorough data quality analysis to ensure accuracy, reliability, and applicability of point cloud data. This involves assessing and validating acquired hyperspectral point cloud to identify potential issues that might impact data quality (Wu et al., 2016). Subsequent improvement measures can then be implemented to enhance or rectify the accuracy of point cloud data based on the insights obtained from the quality analysis.

The objective of this study is to analyse the quality of hyperspectral point cloud obtained from the same birch at different times using the HSL system developed by Finnish Geospatial Research Institute. In addition to investigating the intrinsic factors influencing the quality of reflectance data across individual bands, this study contributes new insights to the analysis of data quality within the domain of hyperspectral LiDAR, which is currently lacking established technical specifications and operational standards. Additionally, the paper seeks to envision and validate the capability of the HSL system inverting chlorophyll content through the utilization of the PROSPECT-D model (Féret et al., 2017).

2. Materials and methods

2.1 HSL System

The Hyperspectral Laser Scanner system developed by Finnish Geospatial Research Institute operates within a laser wavelength range of 420-2100 nm. It features an average output power of 41 mW, a measurement rate of 5.3 kHz, a pulse width of $\leqslant$ 1 ns, a collecting optics of view of 0.2°, a scanning resolution of 0.1° in the horizontal direction, and a scanning resolution of 0.02° in the vertical direction. The detection center comprises eight wavelengths: 545.0 nm, 641.2 nm, 675.0 nm, 711.0 nm, 741.5 nm, 778.4 nm, 978.0 nm, and 1292.4 nm. In addition, the Full-Width at Half Maximum per band is 20 nm.

Each point in the resulting point cloud from the HSL system measurements carries intensity information at these eight different wavelengths. The point cloud encompasses 13 classes that have been classified, including the reference panels, reflective spheres, and most notably, the canopy foliage and woody constituents of the drooping birch. This hyperspectral LiDAR system has exerted significant influence in remote sensing research and various other fields requiring three-dimensional detection and identification, such as civil engineering, archaeology, materials processing, and other specialized domains (Nevalainen et al., 2014). The proposal and application of this system have left a profound and far-reaching impact.

2.2 Data Description

The dataset acquired by the hyperspectral LiDAR system comprises hyperspectral point clouds from 30 individual scans conducted in southern Finland (Kirkkonummi, 60°09′40″N, 24°32′48″W) on September 11-12, 2013. Each scan, lasting 15 minutes, was conducted approximately once every 1 hour, with a shortened interval of 40 minutes near sunset and sunrise.

The point cloud dataset encompasses a total of 13 feature types; one of the most dominant feature types is a birch (Figure 1) and the intensity values of each point have undergone calibration. The calibration process involved distance correction of point intensities. This was achieved by measuring the reflectance of a reference panel at different distances. An interpolation curve was fitted to the reference points, and this curve was then used to interpolate point intensities for all other points, thereby performing the distance correction. Range calibration specifically occurred at a distance of 12.19 m from the scanner. Points beyond this distance correction range had their intensities set to zero on each wavelength channel.

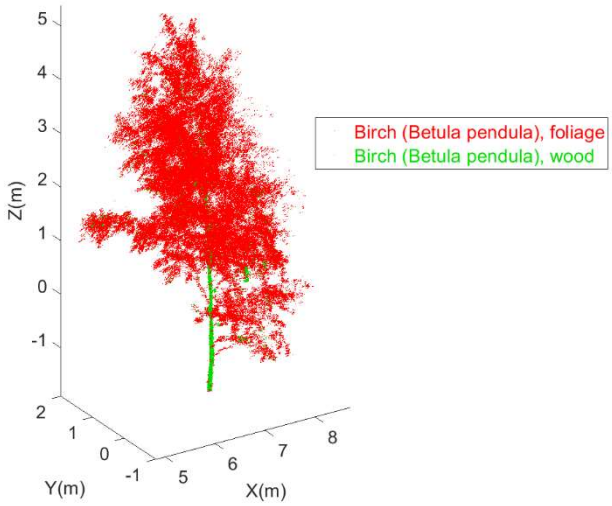

Figure 1. Display of the birch’s point cloud data acquired by the HSL scanning at 09/11/2013 13:10. The scanner was located at the coordinate origin (0,0,0).

Following distance correction, points exhibiting negative intensity values due to dark current or other instrument-related noise were rounded up to zero. It should be acknowledged, notably, that the intensity values in the 8th channel contain significant noise and are unsuitable for analytical purposes.

2.3 Methods

For the obtained 30 sets of hyperspectral point clouds of the same birch, we initially selected different types of features with data measured at 13:10 on 09/11/2013 for the Pearson correlation coefficient calculations of the reflectance data between the 8 bands, which was used to obtain the characterization of the reflectance of the 8 bands of this HSL system. Subsequently, we explored the variations in reflectance of birch leaves at different times and bands through time series analysis. Additionally, we tentatively explored the possibility of inverting the biochemical content of leaves through this HSL system. This exploration was based on the PROSPECT-D model and sensitivity analysis.

2.3.1 Pearson Correlation Coefficient: The Pearson correlation coefficient, commonly referred to as the linear correlation coefficient, serves as a valuable metric for assessing the degree of correlation between two continuous variables, denoted as s and k. This coefficient represents a quantitative measure of the correlation existing between these variables. Hence, in this investigation, we employ the Pearson correlation coefficient to quantify the correlation among the reflectance values across each spectral band. For this analysis, we utilize data extracted from the hyperspectral point cloud acquired during the measurement at 13:10 on 09/11/2013.

The calculation of the Pearson correlation coefficient involves the implementation of the following formula:

$$\rho_{s,k} = \frac{cov(s,k)}{\sigma_s \sigma_k} \quad (1)$$

where, $cov(s,k)$ signifies the covariance between the continuous variables s and k, while σ_s and σ_k represent their respective standard deviations. This formulation explicitly elucidates that the Pearson correlation coefficient is determined by the ratio of the covariance to the product of the standard deviations of the two variables.

2.3.2 Time Series Analysis: In this study, our primary focus is on scrutinizing the variations in reflectance within leaf point cloud data obtained at different time points across each spectral band. The objective is to discern the impact of measurement time on reflectance values across the bands. This approach not only facilitates the identification of moments during measurement with discernible noise and anomalies, it can also enable an exploration of varying noise levels across different datasets. Additionally, we employ a normalization process on the reflectance values to mitigate disparities arising from changes in lighting conditions at different time points. This ensures that the comparison results are more robust and comparable, enhancing the reliability of our findings.

2.3.3 PROSPECT Model: Building upon the quality analysis of hyperspectral point cloud data, we endeavored to assess the feasibility of inversely estimating leaf chlorophyll content. This assessment involved utilizing the PROSPECT-D model in conjunction with reflectance data derived from the hyperspectral point cloud.

The relationship between plant biochemical components and productivity is crucial (Grove et al., 2012). Chlorophyll, for instance, plays a key role in converting solar energy into plant biomass through photosynthesis, impacting plant growth, development, and yield. Methods for inversely determining the content of biochemical components mainly involve physical and empirical models. Physical models, such as the PROSPECT model (Jacquemoud and Baret, 1990), and the leaf-scale optical property model based on the spectrally invariant properties (leaf-SIP) model (Wu et al., 2021), are complex, requiring numerous challenging-to-obtain input parameters. Empirical methods, like the GSV-L model (Ma and Fang, 2023) and Partial Least Squares Regression (PLSR) model (Geladi and Kowalski, 1986), offer simplicity but face challenges related to data universality and accuracy, dependent on training set selection.

The PROSPECT model is an extension of the "plate model" (Allen et al., 1969), which examines leaf reflectance and transmittance based on the leaf structural parameter N and various biochemical parameters. It is capable of simulating leaf reflectance and transmittance across the visible to mid-infrared wavelength bands. The latest iterations of the PROSPECT model include PROSPECT-D and PROSPECT-PRO (Féret et al., 2021). Both versions incorporate input parameters such as the leaf structural parameter N, chlorophyll content (C_{ab}), carotenoid content (C_{ar}), anthocyanin content (Ant), brown pigment, and equivalent water thickness (C_w).

Notably, PROSPECT-D differs by including the dry matter content as a single input parameter, Leaf Mass per Area (LMA, or C_m). In contrast, PROSPECT-PRO introduces a more detailed breakdown of C_m into two primary components: protein content and the carbon-based component (CBC). The CBC encompasses cellulose, lignin, starch, and all non-structural carbohydrates, with minimal contribution to leaf nitrogen content. Specifically, C_m is represented as the sum of CBC and protein content. This distinction enhances the model's capability to capture the intricate biochemical composition of leaves.

2.3.4 Global Sensitivity Analysis: Our primary objective in incorporating sensitivity analysis is to examine the responsiveness of each input parameter in the PROSPECT-D model to the resulting reflectance. This analysis spans the entire range of detection bands within the HSL system. The overarching goal is to assess the theoretical feasibility of employing the HSL system for inversion to obtain plant biochemical attributes, specifically aiming at the estimation of chlorophyll content in leaves. Previous studies, such as Li and Wang (2011) and Sun et al. (2017), have attempted sensitivity analyses on reflectance using the PROSPECT-4 model in different wavelength ranges.

Global sensitivity analysis (the Sobol method) revolves around the fundamental concept that the model output Y can be expressed as the sum of individual contributions and various cross-contributions arising from multiple input variables $X = \{X_1, X_2, \dots, X_n\}$, defined as $Y = f(X)$ (Sobol, 1993; Sobol, 2001). This method presupposes that the input values are independent and uniformly distributed within the unit hypercube, adhering to the relationship:

$$Y = f_0 + \sum_{i=1}^{n} f_i(X_i) + \sum_{i<j}^{n} f_{ij}(X_i, X_j) + \cdots + f_{1,2,\dots,n}(X_1, X_2, \dots, X_n) \quad (2)$$

where, f_0 represents a constant value that fulfills the condition:

$$f_0 = E(Y) \quad (3)$$

The values of $f_i(X_i)$ and $f_{ij}(X_i, X_j)$ conform to:

$$f_i(X_i) = E(Y|X_i) - f_0 \quad (4)$$

$$f_{ij}(X_i, X_j) = E(Y|X_i, X_j) - f_0 - f_i - f_j \quad (5)$$

where $i, j = 1, 2, \dots, n$ and $i \neq j$. The same rationale is applicable to the remaining multiple cross-terms. This structured approach enables the systematic exploration and quantification of the sensitivity of the model output to individual input variables and their interactions, providing valuable insights into the overall impact of each variable on the model's response.

3. Results

To assess the quality of the 30 hyperspectral point clouds acquired by the hyperspectral LiDAR system, Pearson correlation coefficient matrices were calculated separately for three types of point clouds: all feature classes (including unclassified point clouds), canopy-only point clouds, and trunk-only point clouds (Benesty et al., 2008). The correlation coefficient matrix serves as a tool to characterize the linear relationship between the bands of the Finnish HSL system, containing correlation coefficients between the reflectance of the eight bands. Reflectance correlation coefficients provide a statistical indicator describing the similarities and differences in factors affecting reflectance variation between two bands, with values typically ranging between -1 and 1. Theoretically, as the wavelengths between the bands detected by the LiDAR become closer, the reflectivity correlation increases, and the correlation decreases as the wavelengths of separated bands are farther apart. Generally, a higher correlation coefficient between bands indicates more consistent factors affecting reflectivity size in each band.

Table 1 theoretically suggests that the reflectivity correlation is larger when the wavelengths between detected bands are closer, gradually decreasing as the wavelengths of the separated bands are farther apart. In general, the larger correlation coefficient between bands indicates the more consistent factors affecting reflectivity size in each band.

For the Pearson correlation coefficient matrix of all point clouds (Table 1), it is observed that the correlation coefficients between the reflectance at 545.0 nm and 711.0 nm are higher than those at 545.0 nm and 675.0 nm, attributed to the inclusion of the reflectance of the green leaf point cloud. Correlation coefficients decrease as the wavelengths of the separated bands become larger, except for an anomaly in the correlation coefficients between band 1 (545.0 nm) and the other bands.

In the case of leaves (Table 2), the correlation coefficients of reflectance at 545.0 nm and 711.0 nm are higher than those at 545.0 nm and 641.2 nm. Subsequently, the correlation coefficients gradually decrease, but those of 545.0 nm with 741.5 nm and 778.4 nm remain higher than those of 545.0 nm with 675.0 nm. This is attributed to the strong absorption band of chlorophyll in leaves between 640.0 nm and 680.0 nm, resulting in a larger correlation coefficient for reflectance at its bands. It is evident that the reflectance influencing factors of leaves at 711.0 nm, 741.5 nm, and 778.4 nm are similar to those at 545.0 nm, mostly related to the structure of plant chloroplast cells. Additionally, the reflectance correlation coefficients of 675.0 nm with 778.4 nm and 978.0 nm are slightly higher than those of 675.0 nm with 741.5 nm, as the plant cell wall structure at 778.4 nm and 978.0 nm increases leaf reflectance, while 675.0 nm is situated in the chlorophyll strong absorption band, indicating a greater effect of cell structure at 675.0 nm on reflectance compared to weak chlorophyll absorption.

	545.0 nm	641.2 nm	675.0 nm	711.0 nm	741.5 nm	778.4 nm	978.0 nm	1292.4 nm
545.0 nm	1	0.9250	0.8965	0.9153	0.8355	0.8215	0.7774	0.3491
641.2 nm		1	0.9806	0.9114	0.7850	0.7790	0.7658	0.3600
675.0 nm			1	0.8789	0.7476	0.7445	0.7405	0.3481
711.0 nm				1	0.9530	0.9484	0.9262	0.4134
741.5 nm					1	0.9875	0.9621	0.4240
778.4 nm						1	0.9662	0.4299
978.0 nm							1	0.4515
1292.4 nm								1

Table 1. Pearson correlation coefficient matrix for reflectance between 8 bands of all features in the point cloud measured at 13/09/11 13:10

	545.0 nm	641.2 nm	675.0 nm	711.0 nm	741.5 nm	778.4 nm	978.0 nm	1292.4 nm
545.0 nm	1	0.8219	0.7395	0.8558	0.7883	0.7623	0.6940	0.1533
641.2 nm		1	0.9244	0.8668	0.7260	0.7204	0.6993	0.1814
675.0 nm			1	0.8162	0.6840	0.6888	0.6878	0.1756
711.0 nm				1	0.9490	0.9442	0.9162	0.2134
741.5 nm					1	0.9843	0.9575	0.2186
778.4 nm						1	0.9641	0.2232
978.0 nm							1	0.2448
1292.4 nm								1

Table 2. Pearson correlation coefficient matrix for reflectance between 8 bands of the foliage in the point cloud measured at 13/09/11 13:10

	545.0 nm	641.2 nm	675.0 nm	711.0 nm	741.5 nm	778.4 nm	978.0 nm	1292.4 nm
545.0 nm	1	0.9597	0.9485	0.9417	0.9278	0.9211	0.8904	0.4702
641.2 nm		1	0.9946	0.9918	0.9839	0.9811	0.9616	0.5191
675.0 nm			1	0.9924	0.9870	0.9854	0.9683	0.5216
711.0 nm				1	0.9949	0.9937	0.9805	0.5295
741.5 nm					1	0.9966	0.9868	0.5361
778.4 nm						1	0.9887	0.5307
978.0 nm							1	0.5471
1292.4 nm								1

Table 3. Pearson correlation coefficient matrix for reflectance between 8 bands of the wood in the point cloud measured at 13/09/11 13:10

https://doi.org/10.5194/isprs-annals-X-1-2024-49-2024

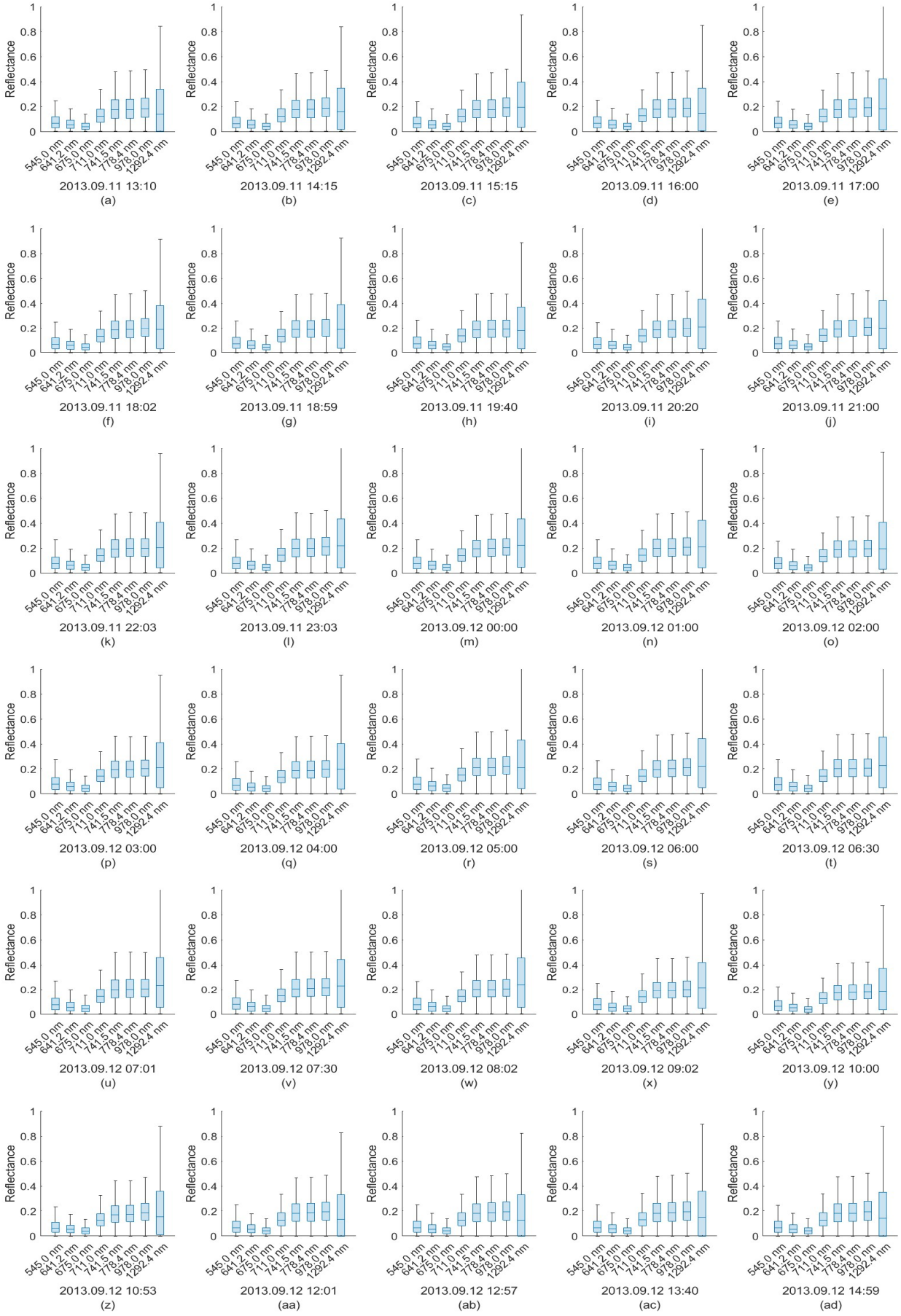

Figure 2. Boxplots of reflectance per band at different moments in the hyperspectral point cloud data of the birch foliage.

Regarding the birch stems (Table 3), as chlorophyll and water content are much smaller than in leaves, they lack the spectral properties of green plants. Consequently, correlation coefficients between bands gradually decrease as the wavelengths become farther apart. However, the correlation coefficients between the first seven bands are very large, suggesting similar factors affecting reflectance between each band.

Across all Pearson correlation coefficient matrices in Table 1, 2, 3, the correlation coefficients between the 8th band (1292.4 nm) and the other bands are relatively small due to noise effects. There is a significant difference in the size of correlation coefficients with the other bands.

Considering the significance of the reflectance of the birch canopy among the 13 feature classes, an analysis was conducted to examine the impact of different time points on the reflectance of hyperspectral point clouds of the canopy, primarily foliage. Some box plots were generated to illustrate the reflectance of hyperspectral point cloud data from the birch canopy detected at 30 different times across eight bands (Figure 2). Each box plot represents the reflectance of the birch's canopy (mainly foliage) for all point clouds in each band at the current moment in time. Additionally, it illustrates the trend of reflectance over the range of the HSL system through the mean reflectance value in each band. We had excluded the outliers from the presented results, enabling a clearer comparison of the temporal impact on reflectance changes in different bands. The outliers, denoting data points significantly distant from the overall distribution, were identified to ensure the robustness of the dataset (Zou and Djokic, 2020).

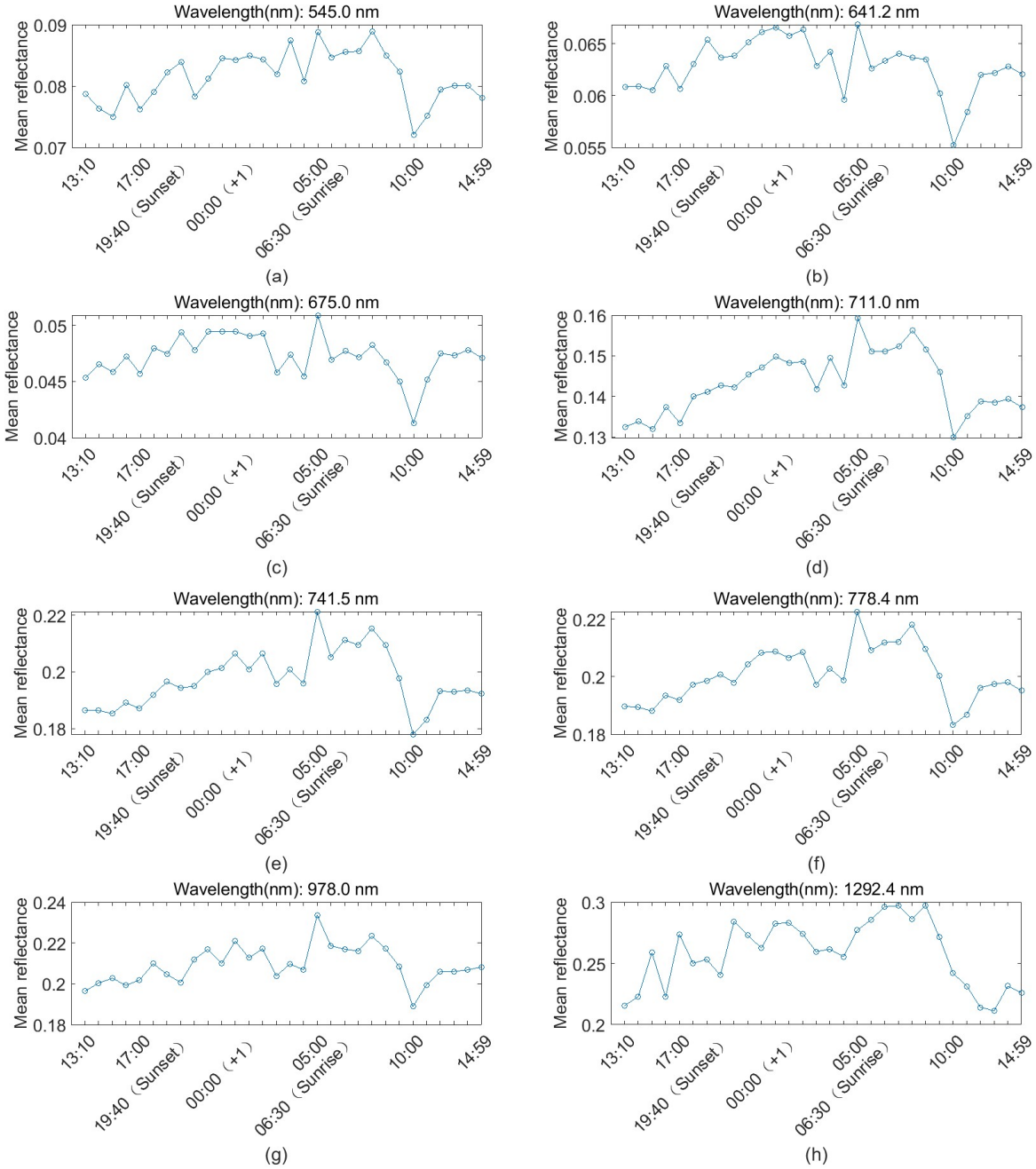

Figure 3. Broken line graphs of leaf reflectance at different times in each band.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-1-2024-49-2024 | © Author(s) 2024. CC BY 4.0 License.

In addition to the box plots, some broken line graphs depicting the change in the median reflectance of each band over the 30 time points were generated (Figure 3). This visual representation aids in understanding how the reflectance values in each band fluctuate over the different moments, providing valuable insights into temporal variations in hyperspectral point cloud data of the birch canopy. Moreover, this point cloud data visualization method serves to troubleshoot whether one or all of the bands of the point cloud data acquired at each point in time have quality issues. This can be in comparison to the data measured at other moments in time, and it also aids in determining the trend of the overall reflectance over time.

The hyperspectral point cloud data of the birch canopy foliage acquired at different times (Figure 2), suggest that overall trends in reflectance changes remain consistent and align with the spectral characteristics of green leaves in the visible-near-infrared wavelength range. For instance, the reflectance of the point cloud at all moments reaches a minimal value at 675.0 nm, followed by a steep increase in reflectance. Surprisingly, when connected the median reflectance of each band and displaying it in the form of a line graph, it becomes evident that the reflectance of different bands at each moment follows the same trend (Figure 3). All bands reach their peak at 05:00 on the next day and reach the minimum at 10:00 on the next day.

A free software tool for Global Sensitivity Analysis (GSAT) package (Cannavó, 2012) in MATLAB (MathWorks, Inc.) was primarily utilized to analyse the leaf structural index (N), chlorophyll content (C_{ab}), carotenoid content (C_{ar}), anthocyanin content (Ant), brown pigment, equivalent water thickness (C_w), and dry matter content (C_m) in the range of 545.0-1292.4 nm. Sensitivity to reflectance output from the PROSPECT-D model was assessed. Figure 4 presents the result of the first-order global sensitivity analysis of N and each leaf biochemical constituent to reflectance, and it covers the spectral range of the HSL system detection.

The results in Figure 4 indicate that C_{ab} and N play dominant roles in the spectral range (545.0-1292.4 nm) detected by this HSL. In the visible range (approx.580-700 nm), chlorophyll exhibits the greatest sensitivity to output reflectance. In the near-infrared range (approx.720-1292.4nm), the sensitivity of the leaf structural parameter to output reflectance is the greatest. This proves that plant biochemical parameters such as chlorophyll content and the leaf structural parameter N can be theoretically inverted from the hyperspectral point cloud data acquired by this HSL system within its detection range.

Furthermore, we employed the specific absorption coefficients for chlorophyll within leaves from the PROSPECT-D model code provided by Féret et al.’s (2017) study, assessing the potential of the HSL system for inverting chlorophyll content. These coefficients were derived from extensive datasets encompassing hundreds of leaf species. The absorption coefficients of chlorophyll were utilized in the range of 545.0-1292.4 nm against the reflectance of the tree canopy in each band. The outcomes of this analysis are presented in Figure 5.

Figure 5 demonstrates that the specific absorption coefficient of chlorophyll is wavelength-dependent, with distinct absorption peaks at approximately 420 and 680 nm. Additionally, this result shows the reflectance of leaves in the birch canopy measured by the HSL system, indicating that the second absorption peak of chlorophyll is covered by the channel of this HSL system. This highlights the ability to invert chlorophyll content using the PROSPECT-D model, providing an advantage in determining the chlorophyll content of foliage utilizing the hyperspectral point cloud data acquired by this HSL system.

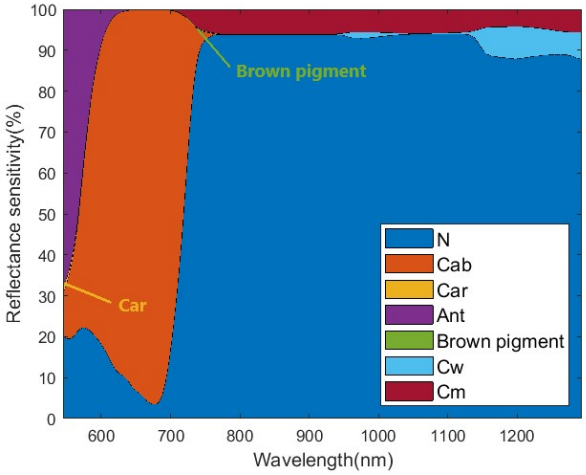

Figure 4. The first-order global sensitivity analysis of the input parameters of the PROSPECT-D model to the output reflectance in the detection band range of the Finnish HSL system.

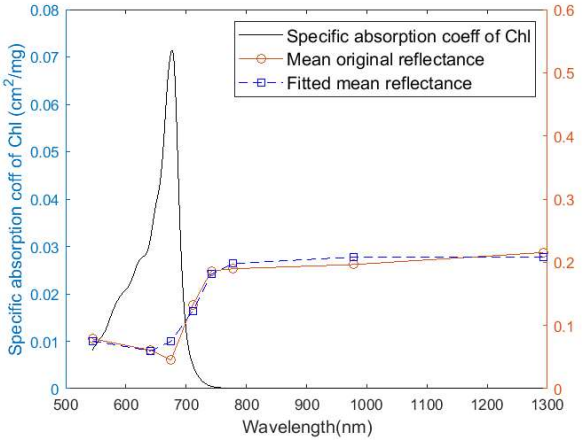

Figure 5. Specific absorption coefficient of Chlorophyll within the detection range and the mean reflectance per band of leaves measured by the HSL.

4. Discussion

This study assessed the quality of an 8-band hyperspectral point cloud dataset obtained from the HSL system developed by the Finnish Geospatial Research Institute. Additionally, we explored the potential for chlorophyll inversion using this dataset. Employing the Pearson correlation coefficient matrix, we scrutinised the relationships between various feature types across different spectral bands. Furthermore, we conducted an analysis of the factors influencing the magnitude of these correlation coefficients. Our findings revealed that the spectral range of chlorophyll absorption exerts a notable impact on the magnitude of correlation coefficients among the bands. Moreover, we observed a substantial rise in noise levels within the eighth band when compared to the remaining seven bands.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-1-2024-49-2024 | © Author(s) 2024. CC BY 4.0 License.

By analysing the time series of reflectance data from this dataset, we identified a significant correlation between the reflectance of the birch canopy (predominantly composed of leaves) across each band and the plant's respiration-photosynthesis dynamics. The process of photosynthesis resulted in a reduction in mean canopy reflectance at wavelengths of 545.0 nm (Figure 3(a)), 641.2 nm (Figure 3(b)), and 675.0 nm (Figure 3(c)). Excluding the 8th band, which was notably impacted by noise, the first seven bands exhibited a consistent trend: the peak mean reflectance was observed at 05:00 AM, with a subsequent decline to the minimum at 10:00 AM. Given that our data were collected during the local summer season, respiration tended to be inhibited by low temperatures from 2:00 to 5:00 AM, with peak activity occurring around 10:00 AM, when the respiration was most appropriate. As temperatures continued to rise, respiration was further suppressed, aligning closely with fluctuations in average reflectance across various bands and leaf respiration patterns. It has been shown that both CO_2 exchange and conductance in leaves are controlled by circadian rhythms (Qin et al., 2024). However, further investigation is warranted to determine whether this correlation stems from stomatal regulation or internal carbon exchange processes within the leaf.

This study has two main shortcomings. First, the hyperspectral point cloud data measured by the HSL system were not precisely categorized, particularly in the case of point clouds categorized as canopy foliage. There existed a small number of point clouds within this category where the real species were branches, having significantly different spectral properties from leaves. This discrepancy can impact reflectance, and the lack of strict differentiation could result in the insufficient accuracy of using hyperspectral point cloud data measured by the HSL system for the inversion of plant biochemical composition.

Second, the results in Figure 4 and Figure 5 theoretically demonstrate the capability to use the PROSPECT-D model for the inversion of chlorophyll content. However, there is still a mismatch in the spectral domains. The detection domain of the HSL system is 545.0-1292.4 nm, whereas the input wavelength of the PROSPECT-D model is 400-2500 nm. While previous studies have explored PROSPECT model inversion utilizing unique sensitivity domains for each biochemical parameter (Li and Wang, 2011), the HSL system lacks coverage across much of the reflectivity spectrum. Consequently, during inversion, it becomes necessary to assign biochemical parameters outside the detection range to values that are reasonable but not precisely accurate. This discrepancy may impact the inversion of other biochemical components, and future efforts should focus on improving the estimation accuracy of biochemical parameters in a wider detection spectral range.

Traditional LiDAR methods have well-established data quality evaluation systems and operational standards (Babbel et al., 2019; Wang et al., 2020). However, as of now, hyperspectral LiDAR measurements lack corresponding technical specifications and operational standards. Most hyperspectral LiDAR systems are currently in the laboratory prototype stage and have not been widely commercialised. Consequently, assessing the quality of hyperspectral LiDAR data often relies on users determining the suitability of the data based on specific task requirements. The methodology employed in this study, for analysing the quality of hyperspectral point cloud data, can serve as a reference for future evaluations of point cloud data obtained from hyperspectral LiDAR systems. In the future, it is essential to enhance the quantitative aspects of quality assessment alongside qualitative analysis of hyperspectral LiDAR-acquired point clouds. This way will facilitate more accurate evaluation of data quality while enhancing it from various dimensions.

5. Conclusion

In this study, we analysed the quality of hyperspectral point cloud data obtained from the same birch at different moments by the self-developed HSL system of Finnish Geospatial Research Institute, and constructed a systematic methodology for analysing the data quality of HSL. By exploring the correlation between changes in Pearson's correlation coefficient values of reflectance in different bands and canopy composition, and analysing the relationship between changes in reflectance in each band of the hyperspectral point cloud and time, we have concluded that the reflectance of canopy leaves measured at different moments exhibits a consistent trend of change. Meanwhile, the reflectance of birch canopies shows a uniform pattern of change across different moments at the same waveband, demonstrating a significant correlation with the plant's photosynthesis-respiration, among other findings. Furthermore, this study also integrated reflectance data from the hyperspectral point cloud of the birch foliage across eight wavelength bands (545.0-1292.4 nm) with the PROSPECT-D model to examine the feasibility of inverting vegetation parameters by using the HSL system. The results demonstrate that the foliage hyperspectral point cloud data obtained by the HSL system can be used to the inversion of leaf biochemical parameters, and the chlorophyll content can be accurately estimated using the PROSPECT model.

Acknowledgements

The work was financially supported by the National Key Research and Development Program of China (No. 2022YFE0112700) and the National Natural Science Foundation of China (No. 32171782).

References

Allen, W.A., Gausman, H.W., Richardson, A.J., Thomas, J.R., 1969. Interaction of isotropic light with a compact plant leaf., *Josa*, *59*(10), 1376-1379.

Babbel BJ, Olsen MJ, Che E, Leshchinsky BA, Simpson C, Dafni J., 2019. Evaluation of Uncrewed Aircraft Systems’ Lidar Data Quality. *ISPRS International Journal of Geo-Information*, *8*(12):532.

Benesty, J., Chen, J., Huang, Y., 2008. On the importance of the Pearson correlation coefficient in noise reduction. *IEEE Transactions on Audio, Speech, and Language Processing*, *16*(4), 757-765.

Cannavó, F., 2012. Sensitivity analysis for volcanic source modeling quality assessment and model selection. *Computers & geosciences*, *44*, 52-59.

Du, L., Gong, W., Shi, S., Yang, J., Sun, J., Zhu, B., Song, S., 2016. Estimation of rice leaf nitrogen contents based on hyperspectral LIDAR. *International Journal of Applied Earth Observation and Geoinformation*, *44*, 136-143.

FAO. 2020. Global Forest Resources Assessment 2020 – Key findings. Rome. https://doi.org/10.4060/ca8753en

Féret, J.B., Berger, K., De Boissieu, F., Malenovský, Z., 2021. PROSPECT-PRO for estimating content of nitrogen-containing

leaf proteins and other carbon-based constituents. *Remote Sensing of Environment*, *252*, 112173.

Feret, J.B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidel, L.P., Ustin, S.L., Le Maire, G., Jacquemoud, S., 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote sensing of environment*, *112*(6), 3030-3043.

Féret, J.B., Gitelson, A.A., Noble, S.D., Jacquemoud, S., 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment*, *193*, 204-215.

Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. *Analytica chimica acta*, *185*, 1-17.

Gong, W., Shi, S., Chen, B., Song S., Niu, Z., Wang, C., Guan, H., Li, W., Gao, S., Lin, Y., Sun, J., Yang, J., Du, L., 2021. Development and prospect of hyperspectral LiDAR for earth observation[J]. *Journal of Remote Sensing*, *25*(1):501-513.

Grove, S., Haubensak, K.A., Parker, I.M., 2012. Direct and indirect effects of allelopathy in the soil legacy of an exotic plant invasion. *Plant Ecology*, *213*, 1869-1882.

Hakala, T., Suomalainen, J., Kaasalainen, S., Chen, Y., 2012. Full waveform hyperspectral LiDAR for terrestrial laser scanning. *Optics express*, *20*(7), 7119-7127.

Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. *Remote sensing of Environment*, *34*(2), 75-91.

Ju, Y., Ji, Y., Huang, J., Zhang, W., 2022. Inversion of forest aboveground biomass using combination of LiDAR and multispectral data. *JOURNAL OF NANJING FORESTRY UNIVERSITY*, *46*(1), 58.

Li, P., Wang, Q., 2011. Retrieval of leaf biochemical parameters using PROSPECT inversion: A new approach for alleviating ill-posed problems. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(7), 2499-2506.

Lin, Y., Zhang, M., Zhang, L., Jiang, M., 2019. Exploration of the Angular Effect in Hyperspectral LiDAR Spectrum-Location-Synchronous Data Collection. *Remote Sensing Technology and Application*, *34*(2), 225-231.

Liu, Q., Yang, L., Liu, Q., Li, J., 2015. Review of forest above ground biomass inversion methods based on remote sensing technology. *J Remote Sens*, *19*(1), 62-74.

Luo, S., Wang, C., Xi, X., Pan, F., Peng, D., Zou, J., Nie, S., Qin, H., 2017. Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. *Ecological Indicators*, *73*, 378-387.

Ma, T., Fang, H., 2023. GSV-L: A general spectral vector model for hyperspectral leaf spectra simulation. *International Journal of Applied Earth Observation and Geoinformation*, *117*, 103216.

Nevalainen, O., Hakala, T., Suomalainen, J., Mäkipää, R., Peltoniemi, M., Krooks, A., Kaasalainen, S., 2014. Fast and nondestructive method for leaf level chlorophyll estimation using hyperspectral LiDAR. *Agricultural and Forest Meteorology*, *198*, 250-258.

Qin, H., Sun, M., Guo, W., He, Y., Yao, Y., Resco de Dios, V., 2024. Time-dependent regulation of respiration is widespread across plant evolution. *Plant*, *Cell & Environment*, *47*, 408–415.

Shen, X., Cao, L., 2017. Tree-species classification in subtropical forests using airborne hyperspectral and LiDAR data. *Remote Sensing*, *9*(11), 1180.

Sobol, I. M., 1993. Sensitivity estimates for nonlinear mathematical models. *Math*. *Model*. *Comput*. *Exp*., *1*, 407.

Sobol, I. M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, *55*(1-3), 271-280.

Sugumaran, R., Voss, M. 2007. Object-oriented classification of LIDAR-fused hyperspectral imagery for tree species identification in an urban environment. In *2007 Urban Remote Sensing Joint Event* (pp. 1-6). IEEE.

Sun, G., Niu, Z., Gao, S., Huang, W., Wang, L., Li, W., Feng, M., 2014. 32-channel hyperspectral waveform LiDAR instrument to monitor vegetation: Design and initial performance trials. *Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications V* (Vol. 9263, pp. 533-539). SPIE.

Sun, J., Shi, S., Yang, J., Chen, B., Gong, W., Du, L., Mao, F., Song, S., 2018. Estimating leaf chlorophyll status using hyperspectral lidar measurements by PROSPECT model inversion. *Remote Sensing of Environment*, *212*, 1-7.

Uchida, T., Hasegawa, K., Li, L., Adachi, M., Yamaguchi, H., Thufail, F.I., Riyanto, S., Okamoto, A., Tanaka, S., 2020. Noise-robust transparent visualization of large-scale point clouds acquired by laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *161*, 124-134.

Wang, H., Yang, T., Wang, Z., 2020. Development of a coupled aerosol lidar data quality assurance and control scheme with Monte Carlo analysis and bilateral filtering. *The Science of the total environment*, *728*, 138844.

Wu, S., Zeng, Y., Hao, D., Liu, Q., Li, J., Chen, X., Asrar, G.R., Yin, G., Wen, J., Yang, B., Zhu, P., Chen, M., 2021. Quantifying leaf optical properties with spectral invariants theory. *Remote Sensing of Environment*, *253*, 112131.

Wu, Y., Bishop, I., Hossain, H., Sposito, V., 2006. Using GIS in landscape visual quality assessment. *Appl*. *GIS*, *2*(3), 18.1-18.20.

Zou, M., Djokic, S.Z., 2020. A review of approaches for the detection and treatment of outliers in processing wind turbine and wind farm measurements. *Energies*, *13*(16), 422.