International Journal of Distributed Systems and Technologies
Volume 16 • Issue 1 • January-December 2025

Assessing Forest Quality Using Multi-Source Satellite Remote Sensing Data:

A Case Study in Western Beijing's Mountainous Regions

Chen Bo
Beijing Vocational College of Agriculture, China

Shan Miao
Beijing Vocational College of Agriculture, China

Yun Zhao
Beijing Jingxi Forestry Administration Office, China

Jinyu Li
https://orcid.org/0009-0006-9657-8799
Beijing Academy of Forestry and Landscape Architecture, China

ABSTRACT

This study uses Sentinel satellite data to estimate forest quality over a large area, focusing on Beijing. By combining ground survey data with remote sensing, a random forest model predicts forest parameters. The results show a correlation coefficient of 0.60-0.76 and a relative root mean square error of 0.09-0.39. Average tree height and diameter at breast height (DBH) had the highest accuracy (75%-80%), followed by canopy density and plant number density (68%-75%). The spatial agreement between predicted and actual forest quality indicates the model's effectiveness.

KEYWORDS

Sentinel Data, Beijing, Forest Quality, Stand Factor, Random Forest

INTRODUCTION

In the context of global climate change, forests, as critical terrestrial ecosystems, play an essential role in mitigating carbon dioxide emissions and attaining carbon neutrality (Ghasemi et al., 2024; Xu, 2023). The concept of forest quality encompasses a multifaceted assessment that integrates ecological, societal, and economic dimensions. Research into forest and site quality enhances our understanding of forest functionality and valuation, facilitates the formulation of judicious management practices for foresters, and fosters the modernization and sustainability of the forestry sector (Lehmler et al., 2023). Traditional methods for assessing forest health are predominantly based on on-site investigations, which establish various indicators to evaluate forest attributes. While these approaches offer high accuracy, they are labor-intensive and resource-demanding, thus limiting their applicability over extensive areas. To address this limitation, remote sensing technologies provide a viable alternative for large-scale and rapid forest health assessments.

DOI: 10.4018/IJDST.383047

This study focused on stand-level attributes to assess the ecological functions of forests. By leveraging state-of-the-art remote sensing data and integrating it with ground-based survey information, a regression model was constructed to predict stand factors within the study area. A novel aspect of this approach was the application of the random forest (RF) algorithm to quantify the relative significance of individual feature variables. By analyzing the sensitivity of these variables to forest and site quality, stand quality and its components were evaluated, clarifying their spatial distribution patterns and interrelationships.

The main contribution of this research lies in having achieved high-precision estimates of forest quality using Sentinel-1 and Sentinel-2 satellite data through an optimized RF regression model. It also acknowledged the assumptions and limitations inherent in the study, such as the potential impact of relying on data from specific regions on the model's generalizability and the fact that only static analysis was conducted, without fully capturing the dynamic processes of forest and site quality.

The article begins by introducing the motivation and background of the study, followed by an overview of the proposed methodological framework. Detailed descriptions of variable selection and model construction processes are provided, including the use of the Boruta algorithm for feature selection and the RF algorithm for building regression models. Additionally, the paper elaborates on data acquisition and preprocessing, ground sample data collection, and feature factor extraction. Results demonstrated the inversion accuracy of different forest parameters and validated the effectiveness of the proposed method by comparing it with other regression models. Conclusions summarize the findings and propose future directions, emphasizing the importance of integrating multi-source remote sensing imagery, unmanned aerial vehicles (UAVs), and light detection and ranging (LiDAR) technology to improve the precision of forest quality assessments.

RELATED RESEARCH

De Moura Fernandes et al. (2023) evaluated the mapping performance of spectral and textural variables derived from MSI/Sentinel-2 imagery for vegetation segments in the Caatinga biome, northeastern Brazil. The study determined multiple indicators, including raw reflectance values, which yielded a high estimation accuracy. Bera et al. (2023) analyzed the efficacy of Sentinel-2 data in predicting canopy cover, revealing that its predictive power was inferior to that of a full Landsat-8 model, as Sentinel-2 distributed variable importance across a broader set of features. Muñoz et al. (2022) utilized a normalized difference index based on Landsat imagery to produce original forest cover maps, subsequently incorporating improved Sentinel-1 and ALOS PALSAR data fusion. This approach achieved a classification accuracy of 0.946 for forested areas and 0.92 for non-forest regions.

To improve the precision of forest parameter inversion, Zhen (2022) applied super-resolution reconstruction of Sentinel-2 images, combined with support vector machine (SVM)-recursive feature elimination, to select optimal variables for estimating canopy nitrogen content. The regression model developed for mangrove canopy nitrogen content prediction exhibited a fitting accuracy comparable to that of a resampled model. Lasko et al. (2024) employed a RF classifier, utilizing Sentinel-2 and low-density time-series images for automatic data training and the generation of a near-global land cover classification framework. The evaluation indicated that the regional adaptation threshold for land cover could accurately delineate land cover types under 9-level, 6-level, and 5-level classification schemes.

Zou et al. (2023) utilized nonlinear least squares and Bayesian methods to analyze dominant height data from *Pinus massoniana* stands, developing an approach that demonstrated superior fit for stand quality and potential productivity predictions. Wambsganss et al. (2022) assessed the influence of tree species mixing on fine root litter quality degradation, considering the interplay between species composition and site conditions. Their findings suggested that while species mixtures can modify fine root litter quality loss across large environmental gradients, their impact was relatively minor compared to that of environmental factors. Dong et al. (2022) constructed a site quality model using the