

Convolutional Neural Network Modeling for the Prediction of Site Index from DEM-based RGB Images: Case Study of Hinoki (*Chamaecyparis obtusa*) Plantation Forests in Nara Prefecture, Japan

Yuki Hirose^{1*}, Naoto Matsumura² 

Abstract: In the present study, we aimed to develop an efficient prediction model for the site index of hinoki (*Chamaecyparis obtusa*) plantations in Higashi Yoshino Village, Nara Prefecture, Japan. For this purpose, we trained a convolutional neural network (CNN) model, then investigated the accuracy of site index prediction and the reproducibility of topographic factors from digital elevation model (DEM) image data acquired using an aerial laser scanner. We also examined terrain red–green–blue (RGB) images, derived from the DEM, as an alternative form of input data. The terrain RGB images outperformed DEM images in terms of prediction accuracy and reproducibility for all evaluation indicators. Reproducibility analysis of elevation, slope, and orientation revealed particularly high accuracy (coefficient of determination > 0.95). Although further improvements to the model are needed, our results emphasize the practicality of using a CNN model combined with terrain RGB images for site index predictions. This approach has the potential to achieve prediction accuracy close to or even surpassing those of existing methods, while requiring fewer types of input data.

Keywords: Convolutional neural network (CNN), digital elevation model (DEM), geographic information system (GIS), site index prediction model, terrain RGB.

1. Introduction

An understanding of site productivity is essential for the development of appropriate forest management strategies. Among measures of forest productivity, the “site index”, calculated as the average height of dominant trees of a specific species at a particular index age (e.g., 40 years in Japan), is widely accepted (e.g., Noordermeer et al. 2020; Kodani et al. 2011). Many statistical and machine learning models have been developed to predict the site index, mainly with application to an entire country or region, using the topographic conditions of local forest land (e.g., weather, geology, topography, and soil) as explanatory factors (e.g., Minowa et al. 2009; Farrelly et al. 2011; Sharma et al. 2012; González-Rodríguez and Diéguez-Aranda 2020; Nakao et al. 2022). Recent studies have suggested that the use of deep-learning methods, which are rapidly evolving, in prediction models could improve model prediction accuracy compared with conventional models (Hirose and Matsumura 2018).

Japan has many private forests, but they are often small and under dispersed ownership. In this context, Japan is promoting “smart forestry” that uses high-resolution (e.g., small forest area, single-tree scale) and high-precision forest information, thereby leveraging the ongoing development and diffusion of remote-sensing technology (MAFF 2018; MAFF 2022). However, practical applications of the site index have been insufficient; problems arise from the applicable site scale, technical hurdles in model development, and low model prediction accuracy. Resolutions for these issues are needed to promote smart forestry.

Regarding site index prediction, many explanatory factors in prediction models applied in Japan have been composed of multiple types of numerical or categorical topographic factors (e.g., slope, aspect, and curvature) derived from geographic information system (GIS) processing of digital elevation model (DEM) data. Although, there are some commonalities depending on the ecological characteristics of each tree species, the optimal combination and importance of topographic factors differs according to study site and spatial scale (e.g., Morisada 2016; Nakao et al. 2022). Therefore, it

Received Dec. 25, 2023; Accepted May. 21, 2024

¹Nara Prefectural Office, Japan

²Mie University, Japan

*Corresponding Author: yuki.hirose64@gmail.com

©2024 The Authors. Published by FORMATH Research Society.

This is an open access article under a Creative Commons [Attribution 4.0 International] licence.

(<https://creativecommons.org/licenses/by/4.0/>) 

is necessary to select the topographic factors appropriate to the site; however, this generally requires significant time and effort from the perspective of topographic data acquisition and prediction model evaluation. González-Rodríguez and Diéguez-Aranda (2020) developed a high-performance site index model; however, they themselves noted that it was inappropriate to require many explanatory variables, including non-topographic factors. In practical site index prediction, there is a need to develop new methods that allow for the simple selection and acquisition of a smaller number of effective explanatory factors, while ensuring a certain level of accuracy.

Considering these circumstances, we focused on convolutional neural networks (CNNs), a type of deep learning, for precise and efficient site index prediction. Among machine learning techniques, CNNs—used in the field of image classification—can automatically process tasks that previously required expert-mediated feature extraction during model training. The classification accuracy of CNNs has reached a level similar to, or even higher than human accuracy, leading to breakthroughs in image classification tasks (e.g., Taigman et al. 2014; He et al. 2015). As an example of the application of DEM images to regression tasks (e.g., site index), Heo et al. (2019) used regression estimation of DEM-derived solar radiation, which was determined from a mathematical model that utilized a CNN model with DEM images as input data. For site index prediction, Hirose and Matsumura (2018) used deep learning but not specifically CNNs. CNNs offer the advantage of directly treating DEM images that spatially incorporate topographic factors, enabling both spatial and areal processing of the terrain structure of the target area. In the present study, we used a CNN model to predict the site index, then investigated its effectiveness considering conventional challenges to the prediction models. Specifically, we used only DEM image data as the explanatory factor (i.e., input data) with a CNN model—rather than using multiple DEM-derived topographic factors, conventionally regarded as explanatory factors—to predict the site index (i.e., output data).

2. Materials and Methods

The methods of this study are summarized and compared with existing methods in Figure 1.

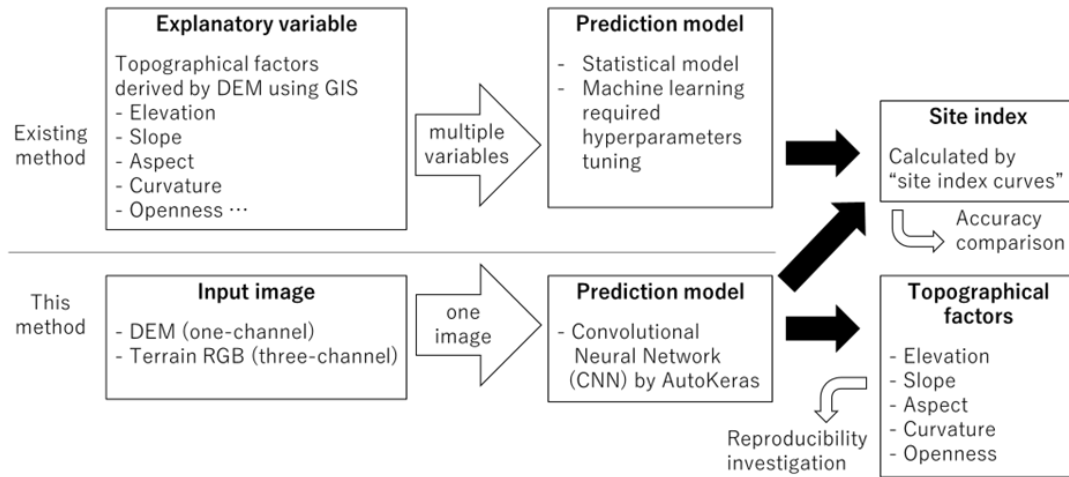


Figure 1. Overview of the methods.

2.1 Site

We conducted the study in Mugitani district (elevation: 854–1,241 m) in the Nara Prefectural Forest located in Higashi Yoshino Village, Nara Prefecture, Japan (Figure 2). This district is a public plantation forest area owned by Nara Prefecture, which has more accurate information regarding forest age, necessary for calculations of the site index, compared with nearby privately owned forests. The forest in this site is composed of hinoki (*Chamaecyparis obtusa*) with stand age ranging from 29 to 36 years old.

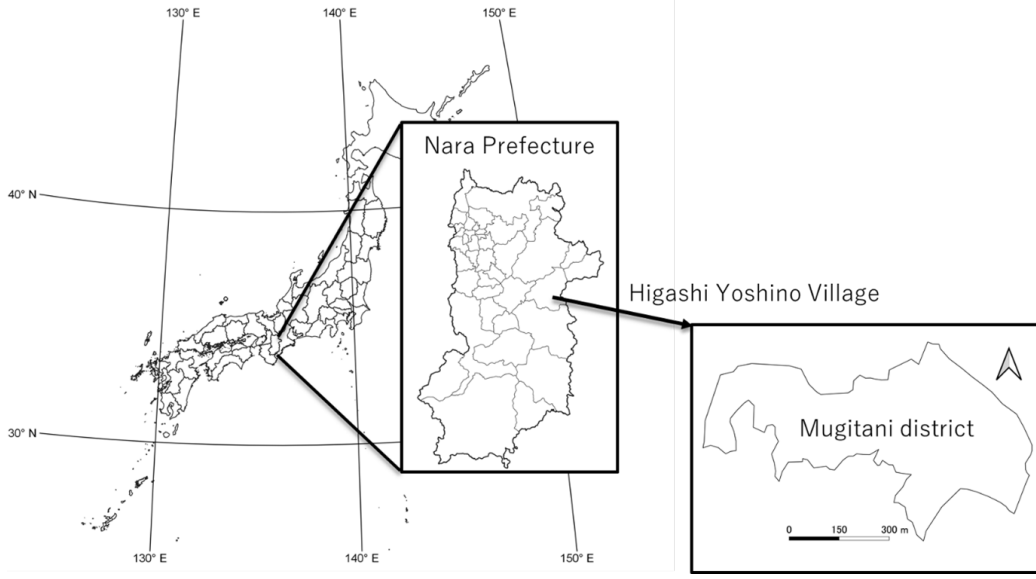


Figure 2. Location of the sample site.

2.2 Data

We used the following datasets for the present study.

1. Aerial laser survey (ALS) data for Nara Prefecture: DEM image dataset (resolution: 0.5 m).
The data was obtained from an ALS project (minimum datapoint density: 4 pt / m^2) conducted by Nara Prefecture in 2020.
2. ALS analysis data for Nara Prefecture: tree species mesh polygon data (mesh size: 10 m by 10 m) and treetop position point data with tree height values.
The data was derived from an analysis conducted by Nara Prefecture in 2021, using the 2020 ALS data. The accuracies of tree species classification and tree height were approximately 80% in the verification survey. Although there is room for discussion regarding the accuracy of tree height, in this study, we assume these values are accurate.
3. Forest planning map (polygon data) and forest register of the Nara Prefectural Forest.
The forest planning map records the spatial distribution of tree species, whereas the forest register records the age and other details of the tree species corresponding to the map.

The sample data were extracted from the above-described datasets as follows (Figure 3).

I. Extraction of hinoki DEM images

From dataset 1, we extracted DEM images of 10×10 pixels covering a 5×5 -m area. These images were selected from overlapping areas in which the tree species was classified as hinoki in datasets 2 and 3. We excluded meshes adjacent to other tree species, where the accuracy of tree species classification was expected to be low.

II. Calculation of the site index and extraction of DEM images for samples.

The site index was calculated for each individual tree using the site index curves for hinoki plantation forest in Nara Prefecture (Waguchi et al. 2019), based on the tree height from dataset 2, and forest age from dataset 3 for trees in the images described in step I. To ensure the quantity and quality of the sample data, we selected DEM images in which the number of treetops (i.e., standing trees) exceeded three and the average site index ranged from 8 to 19, which is the standard range of the site index curves. These criteria yielded a final total value of 4491 DEM images for sampling (Figure 4). Accordingly, an area of 5×5 m, within

hinoki plantation forest and containing at least three trees, was regarded as one forest area in this study. Although this differs from the IUFRO definition, which uses the average height of the tallest 100 trees per hectare, the need for abundant sample data for CNN training and the goal of collecting as many samples as possible within the study area led us to use the average tree height of all standing trees within each 5×5 m cell. Adopting such small-scale cells allowed for the collection of high-resolution data, which is particularly effective in Japan where small-scale, dispersed forest ownership is common. The site index was predicted from the DEM image (resolution: 0.5 m, number of pixels: 10×10) of this forest area.

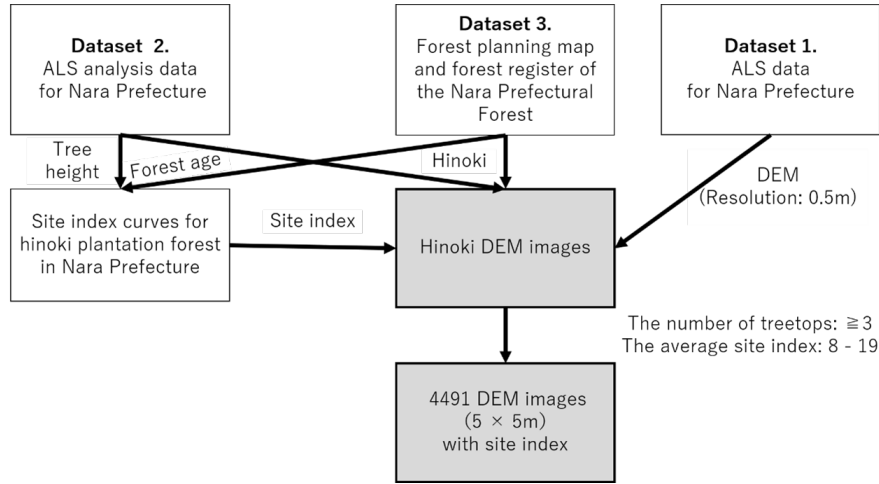


Figure 3. Flowchart of sample data acquisition.

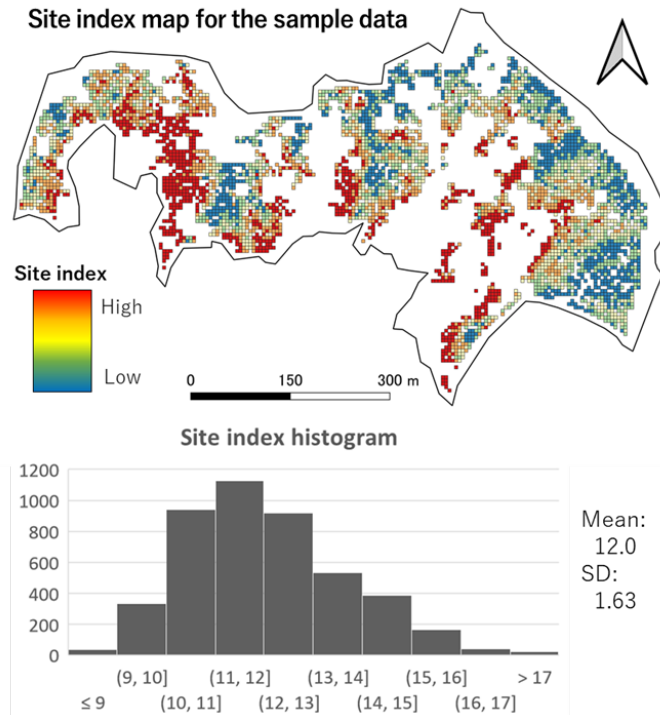


Figure 4. Overview of the sample data.

The DEM image as input image was smaller than those used in Heo et al. (2019)'s CNN regression method involving DEM images. The smaller size caused the differences in elevation values derived

from one-channel DEM images between pixels to be relatively minor, potentially complicating the extraction of terrain features within each image. Hence, in our study, we used both traditional one-channel DEM images, which represent actual elevation values in a 32-bit floating-point format ranging from approximately -3.4×1038 to 3.4×1038 , and terrain red-green-blue (RGB) image data, which encodes actual elevation values based on specific rules (Mapbox 2023). The terrain RGB images consist of three channels converted elevation values derived from one-channel DEM images into RGB values, each was represented as an 8-bit unsigned integer (totaling 24 bits) with values ranging from 0 to 255. The terrain RGB images have been utilized in both visual and analytical, ranging from styling terrain slope and hill shades to generating 3D terrain meshes for video games. The conversion of elevation values into RGB values is performed using the rio-rgb tool developed by Mapbox, an open-source Python library. The process is described as follows.

I. Normalization

Elevation value is adjusted by subtracting a predetermined base value (-10,000) and then normalizing the data by a specified interval (0.1). In this step, the normalized values are rounded to integer values.

$$[1] \quad data_{norm} = \frac{Elevation - (-10,000)}{0.1}$$

II. RGB color channel

RGB values are obtained using a floor function applied to $data_{norm}$.

$$[2] \quad R = \left(\left\lfloor \frac{\left\lfloor \frac{data_{norm}}{256} \right\rfloor}{256} \right\rfloor - \left\lfloor \frac{\left\lfloor \frac{data_{norm}}{256} \right\rfloor}{256} \right\rfloor \right) \times 256$$

$$[3] \quad G = \left(\left\lfloor \frac{\left\lfloor \frac{data_{norm}}{256} \right\rfloor}{256} \right\rfloor - \left\lfloor \frac{\left\lfloor \frac{data_{norm}}{256} \right\rfloor}{256} \right\rfloor \right) \times 256$$

$$[4] \quad B = \left(\frac{data_{norm}}{256} - \left\lfloor \frac{data_{norm}}{256} \right\rfloor \right) \times 256$$

For instance, elevation values from 854 to 1241 meters are represented by RGB values ranging from R:1, G:167, B:252 to R:1, G:183, B:26.

The terrain RGB images have higher sensitivity to elevation values compared to one-channel DEM images due to this pseudo base-256 as RGB encoding, enabling more detailed and diverse depictions of terrain variations and undulations associated with topographic factors. The detailed and diverse color information in RGB images is particularly well-suited for the convolutional layers of a CNN. The convolutional layers analyze the terrain features by filtering and accentuating specific patterns within the RGB data, such as variations in color intensity and texture contrasts (Zeiler and Fergus 2014). This advanced processing capability suggests that the CNN can extract more nuanced and detailed features necessary for site index prediction from the terrain, potentially enhancing its overall performance. Based on the conversion process, the relationship between elevation values and RGB values is as follows: $elevation(m) = -10,000 + ((R \times 256 \times 256 + G \times 256 + B) \times 0.1)$. Furthermore, to better accentuate the detailed terrain features of the terrain, we multiplied the elevation values by 10 and then used the rio-rgb tool to create terrain RGB maps (Figure 5).

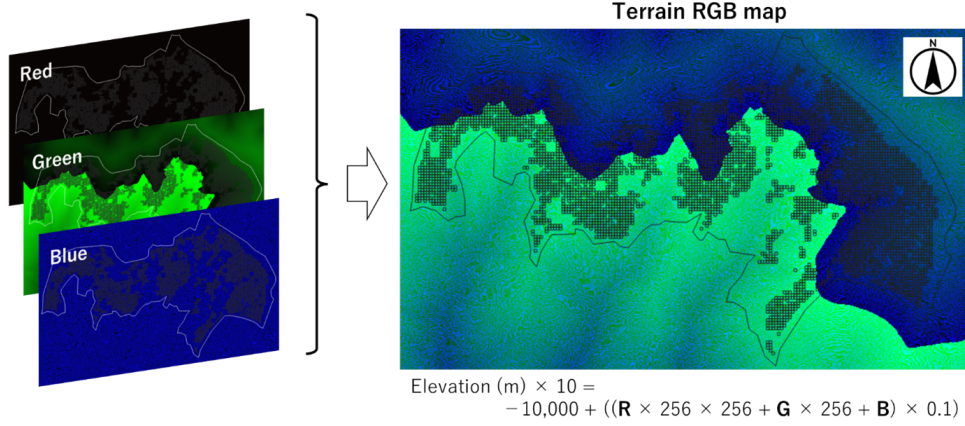


Figure 5. Terrain RGB map of the study site.

2.3 Model design

AutoKeras (DATA Lab at Texas A&M University, College Station, TX, USA) is based on the Keras neural network library. It automates the processes of model architecture search, hyperparameter tuning and model selection while encapsulating the end-to-end process from raw datasets to trained machine-learning models in a simple, flexible interface (Chollet et al. 2023). We used AutoKeras to construct a CNN model that could predict site index (i.e., output data) using either DEM or terrain RGB images (i.e., input data).

Considering the variability inherent in each process during model development, the best model was determined from 50 iterations of model architecture search and hyperparameter tuning, using the training data. Predictive simulations were then executed using the test data. The best models selected from the training process were subsequently assessed based on the average results from 10 simulations for each task.

Furthermore, using AutoKeras, we anticipate that optimal model selection, which has traditionally required significant time and effort, will be automated by its search algorithm to allow future applications of our method in other regions.

2.4 Evaluation method

To assess the generalization performance of the trained model, the sample data (4491 samples) were divided into two groups: 95% (4266 samples) were used as training data and 5% (225 samples) were used as test data. Prediction results from the test data were used to compute the coefficient of determination (R^2), root mean square error (RMSE), and relative RMSE (i.e., $RMSE/standarddeviationoftestdata$). R^2 was calculated using the formula provided by the scikit-learn developers (2023):

$$[5] \quad R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

where y represents the actual observed values in the dataset; $\hat{y}$ denotes the values predicted by the model; y_i and $\hat{y}_i$ are individual observed and predicted data points, respectively; $\bar{y}$ signifies the average of the observed values; and n is the total number of data points in the analysis.

During hyperparameter adjustment, 10% of the training data were used as validation data (427 samples), and the training was repeated to minimize RMSE.

To compare our CNN-based model with existing prediction models, we performed linear regression and deep-learning model (excluding CNN) analyses. The prediction models were based on either elevation as the sole explanatory variable, or on five topographic factors as explanatory variables, using the same approach established for CNN input data. The deep-learning model was constructed using AutoKeras under the same conditions as the CNN.

2.5 Reproducibility of the topographical factors

Before prediction of the site index, we investigated the ability of the CNN model to reproduce topographic factors from the DEM or terrain RGB images. The following topographic factors were selected based on previous studies: elevation, slope, aspect, curvature (i.e., degree of unevenness), and openness (i.e., degree of exposure over a 100-m radius). The values of each terrain factor at the pixel located in the center of the DEM image or terrain RGB image (i.e., input data) of the sample data were predicted (Figure 6). The values of the five topographic factors were obtained by running the raster analysis function of GIS software on DEM images depicting the entire study site. As GIS software, we used QGIS ver. 3.10.14 (QGIS Development Team) for elevation, slope, and aspect (with a 100-m search radius) and ArcGIS Pro ver. 2.7.0 (Esri, Redlands, CA, USA) for curvature.

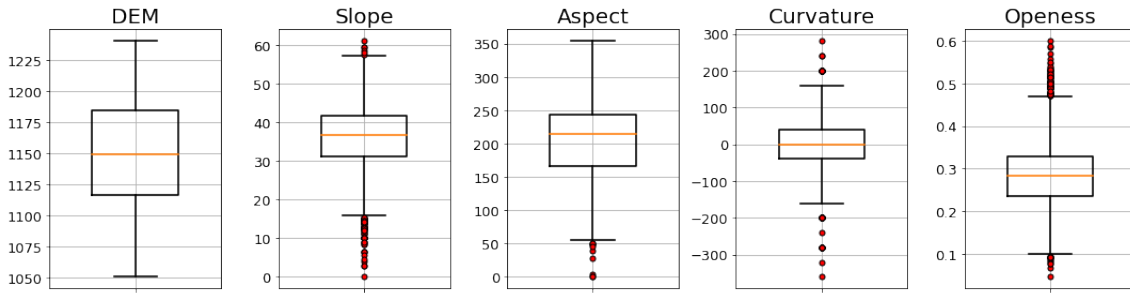


Figure 6. Boxplot of the five topographic factors.

Figure 7 shows the relationships among the input data, output data, and CNN model.

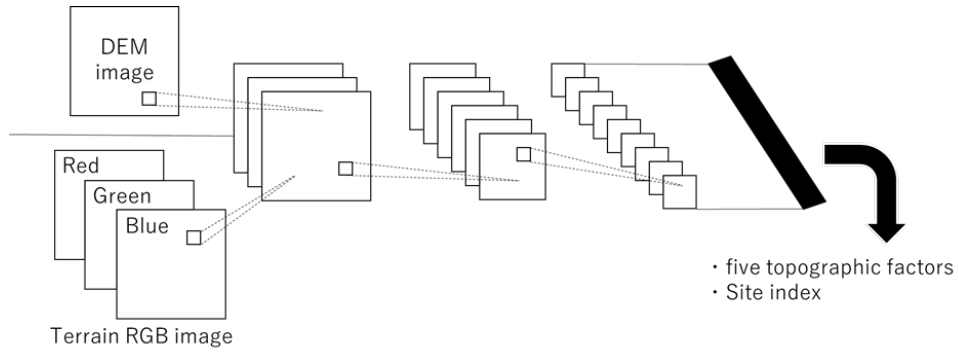


Figure 7. Overview of the input data, output data, and CNN model.

3. Results and Discussion

3.1 Reproducibility of topographic factors

All evaluation indicators indicated that the terrain RGB images, as input data, had higher reproducibility compared with DEM images (Table 1; the architecture of each model with terrain RGB is presented in Table 2, representing 1 of the 10 simulations conducted using AutoKeras). A possible explanation for this result is that the terrain RGB images, which clearly represent topographic changes, are more effective for CNN learning as expected. For the DEM images, R^2 was ≤ 0 , except for elevation, and the accuracy was poor. Conversely, the only factor accurately reproduced by DEM images was elevation, which can be attributed to the center pixel value representing the actual elevation. High reproducibility for elevation was also observed at the same level ($R^2 = 1.00$) in terrain RGB images.

Regarding the topographic factors for terrain RGB images, high reproducibility was observed for elevation, slope, and aspect, with R^2 values of 0.95–1.00; the R^2 values for curvature and openness were lower. The computations for curvature may have been more complex than computations for other topographic factors, reducing its reproducibility. In contrast, openness was calculated with a search radius of 100 m in the GIS analysis, which represents an area much larger than depicted in the input data image (5×5 m); therefore, the information necessary for reproduction was insufficient.

Table 1. Average result of 10 simulations of the topographic factors.

	Input image	R^2	RMSE	relative RMSE
Elevation				
Mean=1151m	DEM	1.00	0.9m	2%
SD=44.1m	Terrain RGB	1.00	0.8m	2%
Slope				
Mean=36.0°	DEM	0.00	8.4°	103%
SD=8.2°	Terrain RGB	0.96	1.7°	21%
Aspect				
Mean=207°	DEM	-3.04	94.7°	170%
SD=55.7°	Terrain RGB	0.97	9.2°	16%
Curvature				
Mean= 0.06	DEM	-0.01	71.4	103%
SD=69.34	Terrain RGB	0.61	43.2	62%
Openness				
Mean=0.29	DEM	-1.91	0.11	138%
SD=0.08	Terrain RGB	0.53	0.05	68%

3.2 Site index prediction

Similar to the topographic factor results, the site accuracy predictions were more accurate when using terrain RGB images than when using DEM images as input data (Table 3). Thus, terrain RGB images, which represent more topographic changes, are effective for site index prediction.

While the CNN model using terrain RGB images demonstrated higher prediction performance than models that used only elevation as the explanatory variable, it showed no advantage over models that used a combination of five topographic factors. Below, we discuss four main reasons for this outcome.

I. Reproducibility and predictive performance of the CNN model

As noted in section 3-1, the CNN model showed low reproducibility for some terrain factors, suggesting that terrain features required for the site index were not fully captured. In this study, AutoML was used to automatically perform model architecture search, hyperparameter tuning, and model selection, but further improvements in accuracy might be achieved by using a CNN architecture (e.g., GoogLeNet; ResNet) not supported by AutoKeras, or by performing detailed hyperparameter searches.

Table 2. Example architectures of CNN models with terrain RGB for 10 simulations.

Elevation	InputLayer (10, 10, 3) ↓ CastToFloat32 ↓ SeparableConv2D (9, 9, 32) ↓ SeparableConv2D (8, 8, 32) ↓ MaxPooling2D ↓ Dropout ↓ SeparableConv2D (7, 7, 32) ↓ SeparableConv2D (6, 6, 256) ↓ MaxPooling2D ↓ Dropout ↓ Flatten (9216) ↓ Dense (32) ↓ ReLU ↓ Dense (1)
Slope	InputLayer (10, 10, 3) ↓ CastToFloat32 ↓ Conv2D (9, 9, 16) ↓ MaxPooling2D ↓ Conv2D (8, 8, 256) ↓ MaxPooling2D ↓ Flatten (16384) ↓ Dense (32) ↓ BatchNormalization ↓ ReLU ↓ Dropout ↓ Dense (16) ↓ BatchNormalization ↓ ReLU ↓ Dropout ↓ Dense (1)
Aspect	InputLayer (10, 10, 3) ↓ CastToFloat32 ↓ Conv2D (9, 9, 32) ↓ Conv2D (8, 8, 64) ↓ Dropout ↓ Conv2D (7, 7, 128) ↓ Conv2D (6, 6, 256) ↓ Dropout ↓ Flatten (9216) ↓ Dense (32) ↓ ReLU ↓ Dense (32) ↓ ReLU ↓ Dense (1)
Curvature	InputLayer (10, 10, 3) ↓ CastToFloat32 ↓ SeparableConv2D (9, 9, 64) ↓ SeparableConv2D (8, 8, 32) ↓ MaxPooling2D ↓ SeparableConv2D (7, 7, 16) ↓ SeparableConv2D (6, 6, 128) ↓ MaxPooling2D ↓ Flatten (4608) ↓ Dense (32) ↓ ReLU ↓ Dense (16) ↓ ReLU ↓ Dense (1)
Openness	InputLayer (10, 10, 3) ↓ CastToFloat32 ↓ Conv2D (9, 9, 32) ↓ Conv2D (8, 8, 256) ↓ MaxPooling2D ↓ Conv2D (7, 7, 32) ↓ Conv2D (6, 6, 512) ↓ MaxPooling2D ↓ Conv2D (5, 5, 32) ↓ Conv2D (4, 4, 32) ↓ MaxPooling2D ↓ Flatten (512) ↓ Dense (128) ↓ ReLU ↓ Dense (128) ↓ ReLU ↓ Dense (1)

Table 3. Average result of 10 simulations of the site index.

Model	Input image	R ²	RMSE	relative RMSE
Deep learning (CNN)	DEM	-4.57	2.37m	145%
	Terrain RGB	0.43	1.19m	73%
Example architecture with Terrain RGB				
Model	Explanatory variable	R ²	RMSE	relative RMSE
Simple linear regression	Elevation	0.21	1.40m	86%
Expression	$Site\ index = 36.349 + (-0.021 \times DEM)$			
Deep learning	Elevation	0.22	1.39m	82%
Example architecture				
Multiple linear regression	Five topographic factors	0.54	1.07m	66%
Expression	$Site\ index = 35.220 + (-0.023 \times DEM) + (-0.027 \times Slope) + (0.011 \times Aspect) + (0.041 \times Curvature) + (9.111 \times Openess)$			
Deep learning	Five topographic factors	0.63	0.96m	59%
Example architecture				

II. Error in the site index

We used site index curves to obtain true site index values from ALS-based tree height data and forest age in the forest register. However, the tree height data accuracy was approximately 80%; the site index data likely contained a degree of error.

III. Input image size

Both types of input data (i.e., DEM and terrain RGB images) were 5×5 m with a 0.5-m resolution; however, this image size may have been insufficient. For example, a study of Japanese hinoki plantations revealed that the optimal resolution for calculating the degree of curvature, which is strongly related to moisture level, was 35 m among several patterns; the researchers concluded that higher resolutions did not necessarily yield better results (Ehara et al. 2009). In the present study, we selected an image size and resolution that would ensure enough samples and pixels were available to construct the CNN prediction model. Future studies should investigate the optimal scale and resolution for site index prediction by comparing multiple image sizes and resolutions from a wider study area.

IV. Number of samples

Typically, when using deep-learning methods such as CNNs, performance tends to improve when more samples are included (Sun et al. 2017). In the present study, we used approximately 4,500 samples. In contrast, Heo et al. (2019) used 25,000 input images for model training. Although the number of required samples depends on the task and model used, considering the relatively small number of samples used in this study, we expect that a greater number of samples would improve the terrain factor reproducibility.

4. Conclusions

We investigated the reproducibility of five representative topographic factors and the site index prediction accuracy based on either DEM or terrain RGB images at a 5×5 m scale using a CNN model for hinoki plantations in Higashi Yoshino Village, Nara Prefecture, Japan. All evaluation indicators suggested that, compared with DEM images, terrain RGB images yielded better results and were effective for CNN model training. High reproducibility ($R^2 > 0.95$) was confirmed for elevation, slope, and aspect. Although the site index prediction accuracy was lower than the accuracy of conventional prediction modeling using five topographic factors as explanatory variables, the use of terrain RGB images alone can provide an accuracy close to existing methods. Moreover, considering that these results were obtained using AutoKeras, which partially automates architecture design and hyperparameter tuning, tasks that have traditionally required advanced technical skills, it has the potential to resolve issues that have hindered the practical application of the site index. As a next step, we will improve the quality and quantity of sample data for the CNN model, especially focusing on curvature and openness where low reproducibility was confirmed, with the goal of improving site index prediction accuracy.

Acknowledgement

The authors are grateful to the referees for their valuable comments and suggestions which improved an earlier version of this manuscript.

Disclosure Statement

The authors declare that there are no competing interest.

Data Availability

Data are available on request for research purposes.

Author Contribution

Conceptualization: HY, MN

Funding acquisition: n/a

Investigation: HY
 Methodology: HY, MN
 Project administration: MN
 Supervision: MN
 Writing - original draft: HY, MN
 Writing - review & editing: HY, MN

ORCID iDs

Naoto Matsumura  <https://orcid.org/0000-0002-4205-6307>

References

- Chollet, F., Jin, H., Song, Q., Hu, X. (2023) AutoKeras: An AutoML Library for Deep Learning, *J. Mach. Learn. Res.* 24 : 1–6.
- Ehara, H., Matsue, K., Shuin, Y., Aizawa, M., Ohkubo, T. (2009) Validity of height estimation of Hinoki cypress using soil moisture indices based on the digital terrain model in Utsunomiya University Forests at Funyu, *Bull. Utsunomiya Univ. For.* 45: 9–15. (in Japanese with English title)
- Farrelly, N., Ní Dhubháin, A., Nieuwenhuis, M. (2011) Site index of Sitka spruce (*Picea sitchensis*) in relation to different measures of site quality in Ireland, *Can. J. Forest Res.* 41: 265–278.
- González-Rodríguez, M.A., Diéguez-Aranda, U. (2020) Exploring the use of learning techniques for relating the site index of radiata pine stands with climate, soil and physiography, *For. Ecol. Manage.* 458, 117803.
- He, K., Zhang, X., Ren, S., Sun, J. (2015) *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*. In: Proc. of 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1026-1034.
- Heo, J., Jung, J., Kim, B., Han, S. (2019) DEM-based Convolutional Neural Network Modeling for Estimation of Solar Irradiation: Comparison of the Effect of DEM Resolutions, In: Proc. of 36th International Symposium on Automation and Robotics in Construction (ISARC 2019), IAARC, 854-861
- Hirose, Y., Matsumura, N. (2018) Estimation of site index categories using deep learning: A case study in Komono Town, Mie, Japan, *J. For. Plann.* 51(2): 81-84. (in Japanese with English title)
- Kodani, E., Matsumura, N., Tarumi, A. (2011) Analyzing the Effects of Environmental Factors on the Sibe Indexes of Sugi (*Cryptomeria japonica*) and Hinoki (*Chamaecyparis obtusa*) Manmade Coniferous Forest Stands in the Shikoku National Forest Using GIS, *J. For. Plann.* 17: 3-8.
- MAFF(Ministry of Agriculture, Forestry and Fisheries, Japan) (2018) A Guide to the Use of Remote Sensing Technology and its Application for the Development and Utilization of High-Precision Forest Information, <https://www.rinya.maff.go.jp/j/keikaku/smartforest/attach/pdf/smart_forestry-20.pdf> (Accessed 31 July, 2023) (in Japanese: Title was translated by the authors of the present study)
- MAFF(Ministry of Agriculture, Forestry and Fisheries, Japan) (2022) Annual Report on Forest and Forestry in Japan Fiscal Year 2021 (Summary), <<https://www.rinya.maff.go.jp/j/kikaku/hakusyo/r3hakusyo/attach/pdf/index-2.pdf>> (Accessed 31 July, 2023).
- Mapbox (2023) Access elevation data <<https://docs.mapbox.com/data/tilesets/reference/mapbox-terrain-rgb-v1/>> (Accessed 31 July 2023).
- Minowa, Y., Suzuki, N., Tanaka, K. (2009) Generalizability and Accuracy of Site Index Estimation Model with Ensemble Learning. Japanese Journal of Forest Planning, *J. For. Plann.* 42: 53–67. (in Japanese with English title and abstract)

- Morisada, K. (2016) Methods for predicting site productivity of sugi plantations, *Kyushu Br. Jpn. For. Soc.* 69: 189-190. (in Japanese with English title)
- Nakao, K., Kabeya, D., Awaya, Y., Yamasaki, S., Tsuyama, I., Yamagawa, H., Miyamoto, K., Araki, M.G. (2022) Assessing the regional-scale distribution of height growth of *Cryptomeria japonica* stands using airborne LiDAR, forest GIS database and machine learning, *For. Ecol. Manag.* 2022, 506, 119953.
- Noordermeer, L., Gobakken, T., Næsset, E., Bollandsås, O.M. (2020) Predicting and mapping site index in operational forest inventories using bitemporal airborne laser scanner data, *Forest Ecol. Manag.* 457(117768): 1-10.
- Sharma, R. P., Brunner, A., Eid, T. (2012) Site index prediction from site and climate variables for Norway spruce and Scots pine in Norway, *Scand. J. Forest Res.* 27: 619-636.
- scikit-learn developers (2023) 3.3. Metrics and scoring: quantifying the quality of predictions — scikit-learn 1.3.2 documentation, <https://scikit-learn.org/stable/modules/model_evaluation.html#r2-score> (Accessed 30 November, 2023)
- Sun, C., Shrivastava, A., Singh, S., Gupta, A. (2017) *Deep Learning Scaling is Predictable, Empirically*. In: Proc. of 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 843-852.
- Taigman, Y., Yang, M., Ranzato, M., Wolf, L. (2014) *DeepFace: Closing the Gap to Human-Level Performance in Face Verification*. In: Proc. of 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 1701-1708.
- Waguchi, Y., Kawai, M., Sakota, K., Yamashita, S. (2019) Creation of Site Index Curves for Hinoki (*Chamaecyparis obtusa*) Plantations in Nara Prefecture in Response to Lengthened Harvest Cycles, *Bull. Nara For. Res. Inst.* 48: 1-4. (in Japanese: Title was translated by the authors of the present study)
- Zeiler, M. D., Fergus, R. (2014) *Visualizing and understanding convolutional networks*. In: European Conference on Computer Vision, pp. 818-833.