Universiteit Leiden

Master Computer Science

Fusing spaceborne radar, lidar and spectrometry data using semi-supervised deep learning to create a high resolution canopy structure dataset

Name: Job Maarten Cornelis Vink
Student ID: s1869205

Date: 03/06/2022

Specialisation: Artificial Intelligence

1st supervisor: Dr. Mitra Baratchi
2nd supervisor: Dr. Suzanne Marselis

Master's Thesis in Computer Science

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

ABSTRACT

High-resolution canopy height maps are essential for environmental studies. Existing methods for obtaining these maps lack the spatial resolution or extent required for applications like studying animal behavior or detecting forest degradation. The Global Ecosystem Dynamics Investigation was launched to the international space station to provide widely available measurements of vertical forest structure. However, the instrument provides sparse measurements: it samples canopy structure with 600 m across-track and 60 m along track spacing. The samples are taken at ~25 m spatial resolution. We propose the use of a semi-supervised deep learning method to create wall-to-wall canopy height maps at 10 m spatial resolution. In this study, we specifically test the use of (1) Synthetic Aperture Radar for estimating canopy height, (2) the use of inverted linear residual blocks in convolutional neural networks to improve current models, and (3) we test the overall accuracy of the applied models. Our results show that SAR data can be used at 10 m resolution to improve model performance. Inverted linear residual blocks can improve model performance at the top of the height range (≈ 55 m) studied. Our experiments show a mean absolute error of 7.03 ± 0.004 and a root mean squared error of 9.47 ± 0.020 when evaluated on our study sites in tropical forests in Costa Rica and Gabon. Future research should investigate whether similar models can be used to estimate other variables of vertical canopy structure, such as Leaf Area Index or vegetation density along the vertical forest axis.

CONTENTS

CHAPTER

ONE

INTRODUCTION

Canopy height and vertical canopy structure are essential indicators for a wide range of applications, such as biodiversity mapping (Marselis et al., 2019), deforestation detection (Saha et al., 2020) and climate change research (Mitchard, 2018). The Global Ecosystem Dynamics Investigation (GEDI) (Dubayah et al., 2020), launched in 2018, gave new insights into the role forests take in the global carbon cycle by collecting information on vertical canopy structure directly. GEDI is attached to the International Space Station (ISS) and is the first spaceborne, full-waveform lidar specifically designed to measure ecosystem structure (Dubayah et al., 2020). The onboard lidar sensor uses laser technology to collect information on the ground's vertical profiles of forest canopies. The instrument scans the earth's surface to obtain measurements with eight beams spaced 600 meters apart. The footprint of one forest structure sample is ~$25\ m$ in diameter, and the footprints are spaced every 60 m along track (Hofton, Blair, Story, & Yi, 2019). The observations made by the GEDI instrument are waveforms of energy as a function of time, containing information about ground elevation, canopy height, and the density of plant material along the vertical forest axis. These observations can be used to obtain maps of canopy height, vertical canopy structure (Figure 1.1b), and above-ground biomass at a $1\ km^2$ spatial resolution.

Unfortunately, for some applications, like studying animal behavior or deforestation detection, the spatial resolution of the GEDI data products is insufficient. Different data fusion approaches have been explored to overcome this challenge, as other spaceborne datasets contain information on canopy structure. For example, Synthetic aperture radar (SAR) and multispectral instruments (MSI) also penetrate the canopy in their ways and therefore include information that can be used to estimate canopy structure (Vaglio Laurin et al., 2018). Such instruments can provide data at a higher spatial resolution than the available spaceborne lidar instruments. The Sentinel-1 and Sentinel-2 missions launched by the European Space Agency (ESA) are equipped with SAR and MSI. The Sentinel-1 mission consists of two satellites currently in polar orbit (Fletcher, 2012). The instruments on these two satellites operate in C-band synthetic aperture radar imaging and have a spatial resolution between $9\ m$ and a little below $96\ m$. The advantage of this instrument is its capability to operate at wavelengths that are not impeded by cloud cover or lack of illumination. These satellites can acquire data during the day and night time. The Sentinel-2 mission also consists of two identical satellites in polar orbit. These satellites carry an MSI that samples 13 spectral bands with a spatial resolution of $10 \times 10\ m$, $20 \times 20\ m$, or $60 \times 60\ m$, depending on the acquisition band used by the instrument. These satellites are especially suitable for capturing vegetation characteristics with the short wave infrared bands.

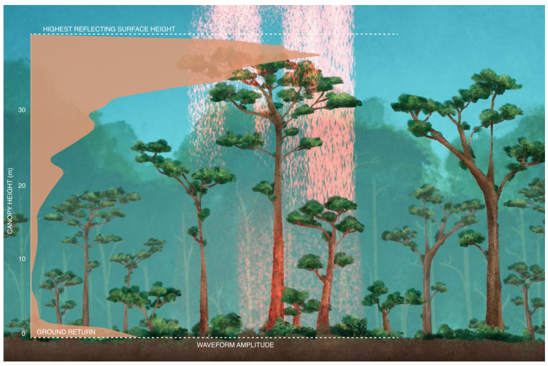
Researchers have proposed methods that use radar and spectral data to estimate canopy structure at higher spatial resolution (Lang, Schindler, & Wegner, 2019; W. Li et al., 2020; Jiang, Zhao, Ma, Li, & Sun, 2021). These methods depend on proxies like shadowing, vegetation type, and vegetation density to estimate canopy structure derived from spectral images or only use data points directly on a location of interest. These techniques can have poor performance at the top of the height range of the study. Lang et al. (2019) find that textural information is not very discriminative above $50\ m$ because species composition and the shape of individual tree crowns do not vary a lot above that range. We will use a convolutional neural network-based approach that combines spectral and radar images to estimate canopy structure. To the best of our knowledge, this configuration for making canopy height estimations has not been proposed before. With these techniques, we hope to improve the performance at the top of the height range.

Our contributions in this thesis are as follows:

- We show that the data from the new GEDI sensor can be used to train a deep learning model.
- We investigate the impact of using SAR data for making canopy height estimations at $10\ m$ resolution.
- We propose the use of inverted linear residual blocks in our model that we call Convolutional Neural Network for Canopy Structure estimations (CNNCS) and compare our model's performance against two baselines.

The remainder of this thesis is structured as follows. Chapter 2 defines the problem for developing and testing a semi-supervised learning method for making canopy height estimations. Chapter 3 discusses related work on interpolation and estimation methods in computer science and remote sensing. In Chapter 4 we describe the study sites in which we conduct our experiments, the dataset, and the various satellites this data is acquired from. Additionally, we describe the preprocessing steps performed on the dataset. Chapter 5 describes our proposed model CNNCS. Chapter 6 describes the research questions addressed in this thesis as well as the baselines and experimental setup. Finally, in Chapter 8, we draw some conclusions and discuss potential future work.

(a) 8 beams over the horizontal swath of the lidar instrument onboard of the ISS.



(b) Visualization of a lidar waveform as collected by the GEDI instrument.

Figure 1.1: Source: `https://svs.gsfc.nasa.gov/13090`

CHAPTER

TWO

PROBLEM DEFINITION

This study aims to develop and test a semi-supervised learning method for training a deep learning model that fuses Sentinel-1 SAR and Sentinel-2 MSI to create canopy structure maps at a higher resolution than possible using GEDI data alone.

We formalize the problem of increasing spatial resolution as a spatial interpolation problem. First, we define the target variable as $y_{h,w} \in \mathbb{R}$ where h is the index of a row and w the index of a column in the matrix $\mathbf{Y}^{H \times W}$, where H and W are the height and width of an area of interest. Then we can define a coordinate inside the area of interest as $c \in \mathbf{c}$ where $\mathbf{c} = [c_1, c_2, ..., c_{H \times W}]$. Every cell in the target matrix $\mathbf{Y}$ is associated with a true value y_c^* that is either known or unknown. Second, we define a feature variable $x_{h,w} \in \mathbb{R}$ associated with the target variable, part of a matrix $\mathbf{X}^{H \times W}$ over the area of interest.

We aim to find a model $\mathcal{M}(c, \mathbf{X})$ to estimate $\hat{y}_c = \mathcal{M}(c, \mathbf{X})$ for all y_c that are unknown in advance. The objective here is to find a model $\mathcal{M}$ that minimizes the residual sum of squares of the model:

$$\mathcal{M}^* \in \arg\min_{\mathcal{M}} \sum_{c \in \mathbf{Y}} (\mathcal{M}(c, \mathbf{X}) - y_c^*)^2 \tag{2.1}$$

CHAPTER

THREE

RELATED WORK

Numerous methods have been developed for spatial interpolation. This section will discuss two general categories of models used for spatial interpolation, namely purely spatial methods and regression with spatial and explanatory variables. In addition, we will discuss approaches proposed by others specifically for spatial interpolation of canopy structure. Some of these proposed methods can also answer the question of parameter importance of the explanatory variables. Here we provide an overview of their findings.

3.1 Interpolation methods

Spatial interpolation methods can be categorized into two basic categories. The first category models the relationship between target variables and their distance. This concept is called spatial autocorrelation and is based on the concept that values of a target variable tend to be more similar to other values closer in proximity. Here we discuss the most popular methods: kriging, nearest neighbor, and splines. Please note that this is not an exhaustive list of applicable algorithms.

Kriging One of the most popular methods used in the field of geo-science is called Kriging (Krige, 1951). This method is also known as a Gaussian process. We can define the basic Kriging model for a coordinate of interest c_i as follows:

$$y_{c_i} = \sum_{c_j \in N(c_i)} (\lambda_{c_i,c_j} \cdot y_{c_j}) + \epsilon_{c_i} \tag{3.1}$$

Here $N(c_i)$ is the neighborhood of the coordinate of interest. λ_{c_i,c_j} is a scalar weight associated with the coordinate of interest c_i and its neighbor c_j. ϵ_{c_i} is the residual of the model. The Kriging method aims to find the weights λ for the coordinates of interest. The Kriging model operates under the stationarity and anisotropy assumption. In addition, the covariance of the target variables is assumed to be constant for all coordinates. A kriging model creates a variogram that models the structure of the measured points.

There are several variants of the Kriging method. Amongst others, ordinary Kriging assumes the unknown mean of the target variables is constant for all coordinates. Universal Kriging allows for modeling a trend in mean.

Nearest neighbor This approach found many applications in statistical learning and was first developed by Fix and Hodges (1989). Nearest neighbor uses the value that is closest in proximity to the coordinate of interest c_i. This approach has the problem of encountering ties between neighbors of equal proximity. As a solution, we can take the average over the neighboring values such that the estimation can be defined as follows:

$$y_{c_i} = \frac{\sum_{c_j \in N(c_i)} y_{c_i}}{|N(c_i)|} + \epsilon_{c_i} \tag{3.2}$$

The neighborhood $N(c_i)$ can be defined in various ways. Often, this is done by selecting k points that are closest in proximity as measured by their Euclidean distance. Other distance measures like hamming distance or Jaccard similarity can also be used. Jiang et al. (2021) use this algorithm to extract contextual features for their estimations model

Splines We can view the target values as a landscape in which the known values are fixed points on a two-dimensional surface. The amplitude of these values can be represented as height on the surface. The spline method tries to interpolate values of a coordinate of interest by estimating a smooth surface between the known points. Formally this can be defined as follows:

$$y_{c_i} = T(c_i) + \sum_{c_j \in N(c_i)} \lambda_{c_j} \cdot R(c_i, c_j) + \epsilon_{c_i} \tag{3.3}$$

$T(c_i)$ is defined to capture the trend for a coordinate of interest. $R(c_i, c_j)$ controls the smoothness of the surface. For the exact mathematics of $T(c_i)$ and $R(c_i, c_j)$ we refer to the work of Harder and Desmarais (1972). Splines are used for example to make estimations of above ground biomass (Filippi, İnci Güneralp, & Randall, 2014)

VPINT Another interpolation method recently proposed by Arp, Baratchi, and Hoos (2022) uses a Markov reward process to interpolate values. At the core of this approach is an iterative element-wise update rule to update an estimation grid. This process can be formalized as follow:

$$y_{c_i} = \frac{1}{|N(c_i)|} \cdot \sum_{c_j \in N(c_i)} \lambda \hat{y}_{c_j} + \epsilon_{c_i} \tag{3.4}$$

In this equation, $N(c_i)$ is the neighborhood of a coordinate of interest. λ are the spatial discount parameters that can be tuned using random search on subsampled data from known values. It could also be replaced by location-specific weights λ_{c_j}.

3.2 Estimation methods

Other methods do use explanatory variables to estimate the interpolation of target variables. We will discuss basic regression models, autoregressive models, and more advanced neural network approaches.

Regression models Numerous models are suitable for regression problems. These models can be applied to spatial interpolation; in the most basic sense, a linear regression model for spatial interpolation can be defined as follows:

$$y_{c_i} = \mathbf{x}_{c_i} \cdot \theta + \beta + \epsilon_{c_i} \tag{3.5}$$

Here θ are the weights as acquired through a learning algorithm. $\mathbf{x}_{c_i}$ is the feature vector for the unknown coordinate of interest. β is the intercept of the model. Other models and algorithms that create a mapping for feature vectors to target variables can also be used in this sense. Note that for the coordinate of interest, that coordinate's feature vector must be known.

Auto regressive models Simple regression models (Haining, 1978) do not use autocorrelation. Autoregressive models consider this autocorrelation by adding a term to the model to capture trends.

$$y_{c_i} = \boldsymbol{x}_{c_i} \cdot \theta + \beta + \lambda \boldsymbol{\phi} \boldsymbol{y} + \epsilon_{c_i} \tag{3.6}$$

The first part of the equation is simply linear regression. Here θ is a vector with the weights, and β is the intercept. The right side of the equation consists of a scalar value λ that acts as a weight between the regression equation and the vector product. $\boldsymbol{\phi}$ is a matrix with weights associated with every value in the neighborhood of the variable of interest. The weights can be defined statically (only considering the values above, left, right, and below the coordinate of interest) or distance based, like a Gaussian. $\boldsymbol{y}$ contains the values in the neighborhood of the area of interest.

Random Forest A random forest model is an ensemble of decision trees (Ho, 1995). Each tree in the random forest is trained on a different training set sample using a bootstrapping method. By creating an ensemble, the variance of the model can be reduced to make estimations more accurate. Random forest models are often used because they make reasonable predictions across a wide range of tasks with little configuration. A majority vote is often taken over all trees in the ensemble to output a label for classification tasks. A simple linear regression model is used at the tree's leaf nodes to make estimations for regression tasks. The output of the random forest is the mean over the output of all trees. Scientists widely use the random forest model to make canopy height estimations (W. Li et al., 2020; Jiang et al., 2021).

Neural Networks In the past decade, many advances have been made in neural network technology. Neural networks and their many variants have been shown to be successful in learning rich contextual features in image and raster data (Krizhevsky, Sutskever, & Hinton, 2012), in recognition of human speech with the use of spectrograms (Abdel-Hamid et al., 2014) and even the evaluation of game positions for *chess* and *go* (Silver et al., 2018). Because of the many successes in various applications, it is not surprising that these approaches have also been used for spatial interpolation. Hashimoto and Suto (2020) have proposed a convolutional neural network designed explicitly for spatial interpolation. They show that their model, called Spatial interpolation with convolutional neural networks (SICNN) can construct wall-to-wall coverage from sparsely measured data points. A similar problem is tackled in the publication by Z. Li, Cao, Wang, and Zhao (2019) who proposed a sparsely self-supervised generative adversarial network. They adopt the standard generative adversarial network (GAN) framework. In this framework a generator is pitted against a discriminator in a zero-sum minimax game. Also, in the field of remote sensing, these neural network approaches find more applications, (Lang et al., 2019) uses a CNN model inspired by the Xception (Chollet, 2016) architecture to make canopy height estimations.

3.3 Estimating canopy structure

As soon as airborne lidar technology became available, it was used to derive vegetation structure (Næsset, 1997). Since then, lidar has been used in numerous forest management and monitoring studies, e.g., (St-Onge, Treitz, & Wulder, 2003; Clark, Clark, & Roberts, 2004). In 2003 the first spaceborne lidar technology was launched with the ICESat mission (Abshire et al., 2005). The lidar sensor on board this satellite provided profiles with a 70 m footprint and a distance of $170\ m$ along track. This mission was followed in 2018 by its successor, ICESat-2, equipped with a photon-counting lidar that samples data with a $13\ m$ footprint. In the same year, GEDI was launched with a full-waveform lidar and provided a $\sim 25\ m$ footprint with a separation of $60\ m$ along track. Operating spaceborne lidar instruments at high altitudes comes at the cost of spatial resolution. To achieve higher resolution, essential for environmental studies, researchers have proposed methods to fuse lidar data with data from other instruments to increase the spatial resolutions of canopy structure maps.

Some initial studies have already shown SAR and MSI's potential to estimate canopy structure. Lang et al. (2019) used Sentinel-2 reflectance images as a predictor for canopy height. They use the Level 1C data product processed with sen2sor for atmospheric correction to create the bottom-of-atmosphere reflectance; level 2A product (Louis et al., 2016). In addition, they use bi-linearly upsampling to convert all 13 channels to square pixels of $10\ m$, which they used to study sites in Switzerland and Gabon. They used a digital surface model (DSM), created by photogrammetric stereo matching, as a reference for their study sites in Switzerland. For the study sight in Gabon, they used data collected from the AfriSAR mission as a reference. This data was used to train a convolutional neural network (CNN). The CNN architecture they used is an adaptation of the Xception architecture (Chollet, 2016). Their approach resulted in a root mean square error of $3.4\ m$ in Switzerland and $5.6\ m$ in Gabon. The mean absolute error in Switzerland was $1.7\ m$ in Switzerland and $4.3\ m$ in Gabon.

W. Li et al. (2020) used SAR and MSI data collected from the Sentinel-1 and -2 satellites and multispectral scans (MSS) from the Landsat missions to predict canopy height acquired using the lidar instrument on the ICESat-2 satellite. From the Sentinel-1 satellites, the ground range detected data product was used. The images were calibrated and ortho-corrected with the Sentinel-1 Toolbox. The Sentinel-2 images were corrected with the sen2cor model (Louis et al., 2016). A median image composition was applied to the Sentinel-1 and -2 images. Images acquired from Landsat were also atmospherically corrected. The authors compared a random forest (RF) model with a Deep Learning (DL) model with five fully connected layers with 100 neurons. The resolution for estimating the canopy height at a spatial scale was $250\ m$. They found a Pearson's correlation coefficient between the observed and predicted canopy height of 0.68 for the RF model and 0.78 for the DL model.

Jiang et al. (2021) proposed the use of a stacking algorithm consisting of a multiple linear regression (MLR), support vector machine (SVM), k-nearest neighbor (kNN), and RF model. They used spectral imaging acquired from the Sentinel-2 satellites as independent variables. Atmospheric correction was applied to the images. In addition, the images were transformed to account for surface reflectance. The Sentinel-2 satellite also has a quality assessment band (QA60), which was used to discard images with cloud cover. The authors use lidar imaging collected from the ATLAS instrument onboard the ICESat-2 satellite to train the models with the addition of a stacking algorithm. The resolution at which the estimations were made is $25\ m$. They found an R^2 of 0.71 between their model's predictions and canopy height collected from the forest management inventory in China.

3.4 Parameter importance

Additionally, some of the studies discussed in the previous section investigate what instruments and what type of measurements are best suited for estimating canopy structure. Researchers have much debate about what instruments are suited for estimating canopy structure. W. Li et al. (2020) showed that the backscattering coefficients acquired from the SAR instrument and the red-edge related variables acquired from the MSI could positively contribute to the prediction of canopy height. Jiang et al. (2021) found that the atmospherically resistant vegetation index is the most important variable, followed by the red-edge chlorophyll index and normalized difference vegetation index (NDVI). Both W. Li et al. (2020) and Jiang et al. (2021) use a random forest model to evaluate this variable importance. Lang et al. (2019) train and test their model with several different band combinations to evaluate the variable importance. They found that the red, green, blue, and near-infrared bands with the high $10\ m$ resolution are most useful for their model. It is worth noting that towards the top of the height range they studied, they found the Red-edge bands with a $20\ m$ resolution and other bands with 60 m resolution to deteriorate the regression rather than improve it. This research shows the potential of including SAR as an independent variable.

3.5 Summary and evaluation of related work

All methods named in this chapter have their strengths and weaknesses. Kriging uses a Gaussian process to make estimations. The stationary assumption that the Kriging method makes can cause problems when applied to different types of forests. Nearest neighbor is simple to implement. The method takes the average value over all its neighbors from a predefined neighborhood. Nearest neighbor can be very sensitive to outliers in this neighborhood. Splines try to interpolate values by estimating a smooth surface between the known points. VPINT interpolates values by using an iterative element-wise update rule. In this way, intermediate values are also taken into account. While these interpolation methods can yield excellent results, data provided by the GEDI instrument is too sparse as, on average, only 0.25% of the total area covered by the sensor contains a value. This sparsity is too high to give reliable results.

Learning methods like linear regression and random forest usually look at single-point estimations on a coordinate of interest. (W. Li et al., 2020) use a Random Forest and simple neural network to make estimations at a $250\ m$ resolution. While this method can yield excellent results when looking at larger areas at a $10\ m$ resolution, you will need to look at more detailed features like shadowing, vegetation type, and vegetation density. Convolutional neural networks excel in these cases because they can have very deep internal representations of these concepts. (Jiang et al., 2021) used a stacking algorithm to combine an MLR, SVM, kNN, and RF for canopy height estimations at a $25\ m$ resolution. Their study only looks at forests with a height range between 4 and $20\ m$. Our work is most similar to that of (Lang et al., 2019). We both use our model's convolutional neural network (CNN) based approach. Our work differs in the choice of independent variables for making estimations and the architecture to base our CNN on. In Chapter 4 and 5 we will introduce these choices in further detail. With these changes in design, we hope to improve the model at the top of the height range.

CHAPTER

FOUR

METHOD

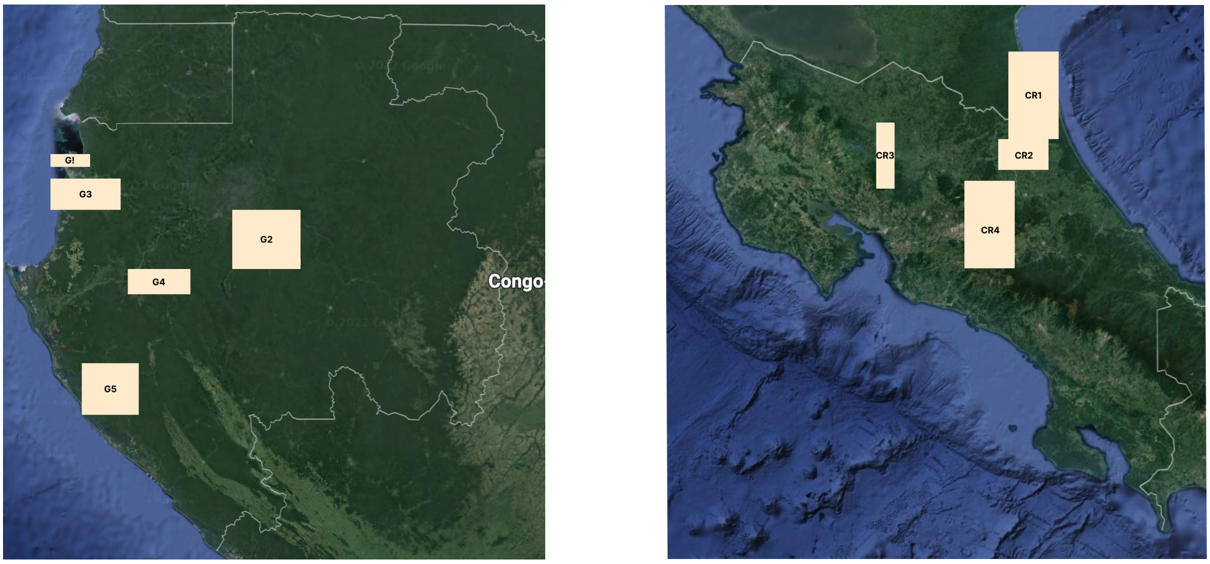
In this chapter, we describe the study sites, datasets, and preprocessing steps required to attain the datasets used to carry out the experiments outlined in Chapter 6. We chose Gabon and Costa Rica because of research applications under development at CML and the availability of validation data. We will first describe these areas, and after the considerations for these areas, we will describe the data collected from various satellites. These datasets require preprocessing to account for atmospheric disturbances. We will describe the techniques we use in Section 4.3. In our approach, we aim to include synthetic aperture radar as an independent variable for its ability to penetrate the canopy. We will describe the collection method for this data in Section 4.4. With the data collected, we aim to use linear bottlenecks and inverse residuals originally proposed by MobileNetV2 (Sandler, Howard, Zhu, Zhmoginov, & Chen, 2019). The motivation behind these specific layers will be described in Chapter 5. The code used in this thesis is available at `https://github.com/jobvink/thesis`.

4.1 Study sites

We chose two study sites for our experiments: Gabon, Africa, and Costa Rica, Central America. In addition to research applications like studying animal behavior or detecting forest degradation and the availability of validation data, we chose Gabon and Costa Rica for their position on and near the equator. Because of this position, there are no significant phenological differences throughout the year – as are seen in temperate forests – and thus, measurements of canopy structure collected throughout the year can be used. In addition, Gabon and Costa Rica give us the unique opportunity to cross-validate our findings with high-resolution canopy data acquired during the AfriSAR missions and the GEDI validation mission. Within Gabon and Costa Rica, we selected nine areas in total. Five are located in Gabon, and the remaining four lie in Costa Rica. High-resolution canopy data was collected during the AfriSAR mission in all selected areas. These areas were selected based on the availability of field measurements, preceding ESA acquisitions, recommendations by experts, and accessibility (Fatoyinbo et al., 2021). The locations of these areas are presented In Figure 4.1.

4.2 Datasets

For this study, we used data acquired from various instruments and missions. All instruments have their unique way of capturing canopy characteristics. Some instruments operate at wavelengths



(a) **(b)**

Figure 4.1: Study sites chosen for our research in Gabon, Africa (a) and Costa Rica, Central America (b).

impeded by cloud cover, rendering some of the captured data useless for our application. We utilized the fact that these satellites revisit the same location at set intervals to create a cloudless composition of spectral images. This section will describe the different characteristics of the datasets used in this study.

4.2.1 GEDI

The full-waveform lidar sensor onboard the international space station (ISS) collects eight ground tracks with three laser sensors. All tracks are spaced 600 m apart. Two lasers are configured to capture four ground tracks operated on full-power mode; the other laser captures four tracks operated on coverage mode. In this study, we use both lasers to measure canopy top height. Measurements taken along track are spaced $60\ m$ apart and have a footprint of ~$25\ m$. GEDI collects data on a global scale between 51.6° N and 51.6° S latitudes. The Level 2A data product of the GEDI mission provides, among others, canopy top height information with other additional observations not relevant to this study. This study will use the relative height (RH98) metric intended for canopy structure measurements from this Level 2A data product (Dubayah et al., 2021). These measurements were taken from January 1st, 2020, until October 1st, 2020. These were the first measurements from the data repository used for this study and closed in timestamp to the data collected during the AfriSAR and GEDI validation missions.

4.2.2 LVIS

During the AfriSAR campaign, the Land, Vegetation and Ice Sensor (LVIS) was flown over several regions in Gabon, Africa (Fatoyinbo et al., 2021). The sensor is designed to measure vegetation structure, sub-canopy ground elevation, and topography of ice sheets and glaciers and is similar to the design of the GEDI lidar sensor. The sensor is mounted on an airplane and collected over $7000\ km^2$ of full-waveform lidar data. The sensor was operated from February 20th to March 8th, 2016, at $7.3\ km$ altitude. The footprint of one sample is ~$22\ m$ with a $9\ m$ separation between tracks, meaning there is some overlap between tracks.

In May-June 2019, an additional mission was launched to sites in the United States and Central America. This mission provided calibration and validation data for the GEDI instrument and operated the same LVIS sensor as during the AfriSAR campaign. Our research used the flight trajectories above Costa Rica as additional cross-validation data.

An overview of the area in km^2, the number of measured data points from the GEDI and LVIS sensor, and the average height are presented in Table 4.1. The LVIS sensor took more measurements in all areas than the GEDI sensor. The difference in the number of measurements is expected since the GEDI sensor takes sparse measurements.

Table 4.1: Overview of the study sites over Gabon and Costa Rica as captured by the LVIS and GEDI sensor. The number of samples is the number of pixels with a ground return value in a location. The study sites in Gabon start with GA, and the study sites in Costa Rica start with CR. An overview of where these study sites are located can be found in Figure 4.1.

		LVIS		GEDI	
	Area (km^2)	nr. samples	mean (m)	nr. samples	mean (m)
GA1	452	2,401,295	12.81	5,559	17.57
GA2	5,128	9,151,636	33.56	212,304	30.76
GA3	2,911	7,462,942	18.57	46,065	24.81
GA4	2,830	6,665,323	28.56	2,659	23.60
GA5	5,317	6,027,759	31.43	23,110	24.47
CR1	2,955	4,856,829	13.21	324,572	18.20
CR2	928	3,229,223	12.07	208,676	16.11
CR3	2,918	9,774,591	17.19	227,585	16.87
CR4	686	514,221	12.31	20,334	15.61

The LVIS sensor does not cover the complete study area as indicated in Figure 4.1 as the flight paths used do not provide full coverage. Additionally, due to cloud cover and other disturbances during the mission, the data provided by the LVIS sensor also contain some data gaps. While the study sites are rectangular, the actual measured surface is slightly less due to the measuring method. This difference in shape introduces a slight variation in the distribution of measured canopy top heights within the study area. The distributions within Gabon and Costa Rica are presented in Figure 4.2.

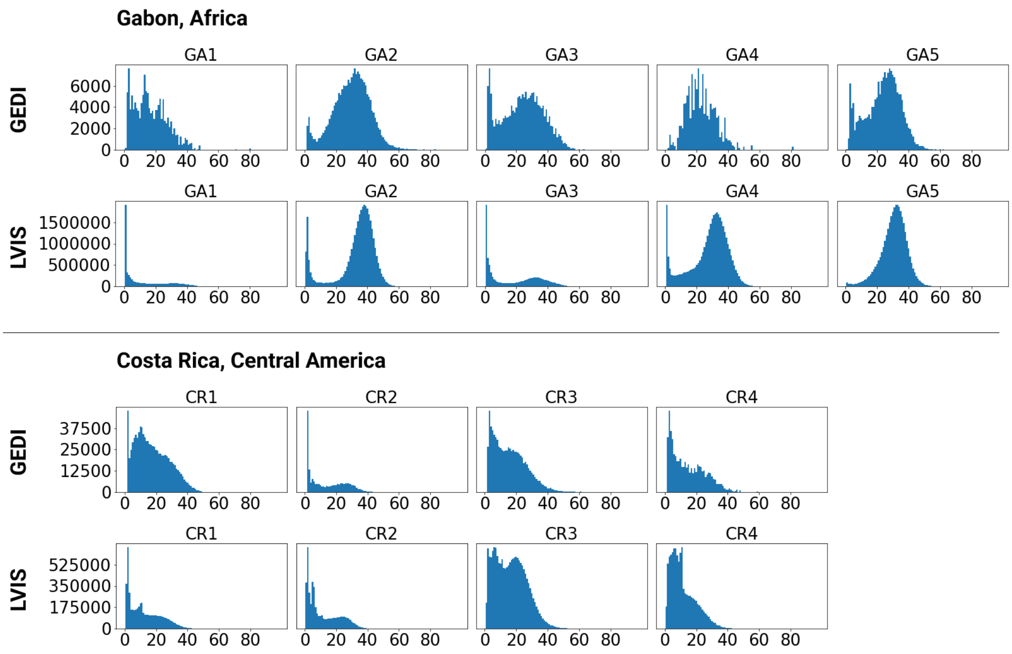


Figure 4.2: Distributions of canopy height in our study sites above Gabon and Costa Rica. The location of the study sites is visualized in Figure 4.1

4.2.3 Sentinel-1

The Sentinel-1 constellation consists of two polar-orbiting satellites that operate in C-band synthetic aperture radar imaging. The advantage of the C-SAR instrument on the satellites is that it operates at wavelengths not impeded by lack of illumination or cloud cover. The Sentinel-1 satellites have a 6-day revisit time. The swath width of these satellites is $250\ km$. The level-1 ground range detected product consists of multi-looked SAR data projected to the ground range using the Earth ellipsoid model. The spatial resolution of the images depends on the mode it operates in. There are three modes: full resolution, high resolution, and medium resolution. The resolution of these modes ranges from $9 \times 9\ m$ per pixel for full resolution mode to $93 \times 87\ m$ per pixel for medium resolution mode. The C-band synthetic aperture radar has a wavelength of 5.404 GHz. Our research uses the Interferometric Wide (IW) swath, the primary acquisition mode over land, and provides 5x20 m resolution.

4.2.4 Sentinel-2

The multispectral instruments onboard the Sentinel-2 satellites capture 13 spectral bands. The mission was launched under the Global Monitoring for Environment and Security program. The mission consists of two satellites to achieve a 5-day revisit time. The swath width of these satellites is $290\ km$ with a spatial resolution of $10 \times 10\ m$ for four visible and near-infrared bands, $20 \times 20\ m$ for six red edge and infrared bands, and $60 \times 60\ m$ for three atmospheric correction bands, the full specifications of these bands are presented in Table 4.2. The wavelengths of these bands range from $443\ nm$ to $2190\ mn$.

Table 4.2: The bands, resolution, and wavelengths that the MSI instrument onboard the Sentinel-2 satellite uses.

Band	Resolution	Central Wavelength	Description
B1	60 m	443 nm	Ultra blue (Coastal and Aerosol)
B2	10 m	490 nm	Blue
B3	10 m	560 nm	Green
B4	10 m	665 nm	Red
B5	20 m	705 nm	Visible and Near Infrared (VNIR)
B6	20 m	740 nm	Visible and Near Infrared (VNIR)
B7	20 m	783 nm	Visible and Near Infrared (VNIR)
B8	10 m	842 nm	Visible and Near Infrared (VNIR)
B8a	20 m	865 nm	Visible and Near Infrared (VNIR)
B9	60 m	940 nm	Short Wave Infrared (SWIR)
B10	60 m	1375 nm	Short Wave Infrared (SWIR)
B11	20 m	1610 nm	Short Wave Infrared (SWIR)
B12	20 m	2190 nm	Short Wave Infrared (SWIR)

4.2.5 Dataset comparison

Measurements collected by spaceborne sensors can suffer from geo-location errors. In addition, because we collected data from different timestamps, temporal decorrelation can occur. Natural forest growth, forest degradation, disturbance, and deforestation are reasons for temporal decorrelation for canopy height metrics. We compared the canopy height metrics collected by the LVIS sensor with the metrics collected by the GEDI sensor at the same geo-location on mean error (ME), mean absolute error (MAE), and root mean squared error (RMSE). The results of these evaluations are presented in Table 4.3. Most notable is that all mean differences in Costa Rica are positive, meaning there was more forest gain than forest loss on average. In Gabon, this was the opposite; there was more forest loss in these areas. The most significant difference in measurements was taken in area CR4. There were no overlapping data points on GA5 between 1st January 2020 and 1st October 2020. The lack of overlapping data points was due to the flight trajectory of the international space station and weather conditions during this period.

Table 4.3: Comparison between LVIS and GEDI over our study sites on mean error (ME), mean absolute error (MAE), and root mean squared error (RMSE).

study site	ME	MAE	RMSE
CR1	0.88	5.99	8.63
CR2	3.63	7.46	15.39
CR3	1.79	5.59	8.21
CR4	3.62	9.70	12.59
GA1	-6.34	8.77	12.31
GA2	-3.76	9.11	13.48
GA3	-2.84	5.25	8.51
GA4	-3.51	8.43	10.98
GA5	—	—	—

4.3 Data preprocessing

Radar imaging acquired from the Sentinel-1 satellites needs atmospheric correction. This preprocessing step can be done with the Sentinel-1 toolbox developed by the European Space Agency (Veci et al., 2014). The following three preprocessing steps for radar imaging were needed on the Sentinel-1 ground range detected data product: Thermal noise removal, Radiometric calibration, and Terrain correction. All three steps were performed with Google Earth Engine.

One main problem that affects the multispectral instrument onboard the Sentinel-2 satellites is cloud cover. Thick cloud cover can completely corrupt the reflectances in all optical frequency bands by obstructing the surface underneath the clouds. This problem can cause considerable data gaps in the spatial and temporal domains. Studies with the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument onboard the Terra and Aqua satellites have shown that clouds cover 67% of the Earth. Overland surface averages 55% (King, Platnick, Menzel, Ackerman, & Hubanks, 2013). Techniques for removing clouds can be categorized into three major categories: multispectral, multi-temporal, and inpainting techniques. Multispectral approaches are usually applied for haze and cirrus clouds. Because clouds do not entirely block the surface of the Earth, but reflectances are only partially absorbed, the signal can be restored using mathematical (Sundar, Sharma, & Naresh, 2013) or physical models (Zhu, Wu, Wu, & Zhao, 2017). Multi-temporal techniques restore scenes covered by clouds by combining information for a scene Captured at different times. These techniques are most popular because they substitute covered signals with actual observations not impeded by clouds. Multi-temporal techniques can also be problematic for rapidly changing surface conditions. The rapid change in conditions can be due to phenological events. Inpainting methods fill data gaps by using surface information from clear parts of the same image. These techniques do not require additional images but only achieve good results on images with small clouds.

The spectral images acquired from the Sentinel-2 satellites also need preprocessing. Scene classification and atmospheric correction were applied to create the level 2A Bottom-Of-Atmosphere reflectances from the level 1C top-of-atmosphere reflectances. Google performed this preprocessing step. The preprocessing of these level 1C data products creates additional data layers we use for cloud filtering. Our filtering method was based on the per-pixel cloud probabilities created with the Sentinel2-cloud-detector library (Braaten, 2020). We used a threshold of 25% to mark pixels with clouds. This threshold was set empirically by visual inspection above our study sites. In addition to filtering pixels with clouds based on these probabilities, we also used these marked pixels to remove the shadows cast by the clouds. We projected the clouds onto the ground using the mean solar azimuth angle captured by the satellite. We intersected these pixels with illumination below 15% on the near-infrared band (B8). The pixels marked as clouds or dark pixels were filtered from the images. After this preprocessing step, the per-pixel average was taken to produce the dataset. In addition, we created a metadata layer for filtering areas with water at evaluation time because estimating a canopy height of 0 above water is a trivial task.

Raw waveforms collected by the GEDI instrument were converted into relative height metrics at 100 intervals of 1 percent, meaning that RH50 provides the relative height (from the identified ground) at which 50 percent of the energy was returned. The USGS LP DAAC performed the prepossessing of the L1A raw waveform. Google and USFS Laboratory did the rasterization of the relative height metrics for Applications of Remote Sensing in Ecology (LARSE). Data points were filtered using the quality and degraded flag.

Data collected using the LVIS sensor was downloaded from the National Snow and Ice Data center (NSIDC). The data points were rasterized using Google Earth Engine and projected on the WGS-84 ellipsoid model. The grid was interpolated using inverse distance weighing with a $25\ m$ influence range. The gamma value for this algorithm was set empirically to 0.3.

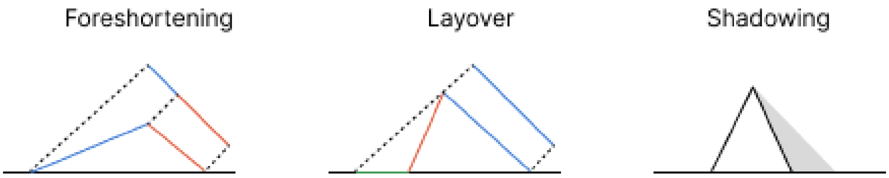


Figure 4.3: Frequent problems with Synthetic Aperture radar. Source: (Flores et al., 2019)

4.4 Synthetic aperture radar

Synthetic aperture radar (SAR) collected by the Sentinel-1 satellite provides dual-polarization C-band radar images. The frequency of this C-band is between 3.4 GHz and 4.2 GHz. The use of synthetic aperture radar for estimating canopy height is based on two principles of SAR (Flores, Herndon, Thapa, & Cherrington, 2019). Rough surfaces increase the number of scatters and thus increase the reflective power. Secondly, an increase in vertical structure results in the randomization of polarisation. Despite these properties of SAR, radar imaging is not yet widely adopted in models for estimating canopy height. The lack of adoption may be due to issues and artifacts with radar imaging, namely, foreshortening, layover, and shadowing. These problems are visualized in Figure 4.3. In this research, we take a closer look at the inclusion of SAR data for canopy height estimations.

CHAPTER

FIVE

MODEL

In order to make canopy height estimates, we need a model configured for creating per pixel regression estimations. An architecture with fully convolutional layers is suitable for such a model because they can have the same input and output shape. In our experiments, we used inverted linear residual blocks. The layout of these blocks is visualized in Figure 5.1. Because these residual blocks do not have an activation function at the end, the normalized linear combinations are used for the next block. The primary motivation behind using these linear residual blocks is that they create an activation in the positive and negative directions (Sandler et al., 2019). We aimed to find out if this technique can benefit regression tasks. Our proposed model is specifically designed for making canopy structure estimations which is why we call it; Convolutional Neural Network for Canopy Structure or CNNCS for short.

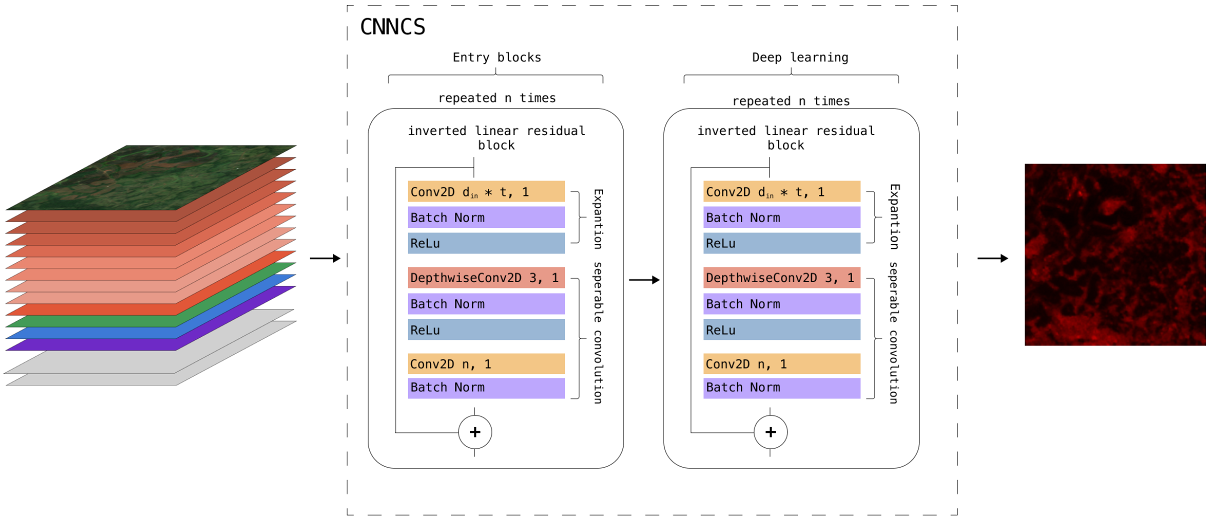


Figure 5.1: The setup for making canopy height estimations. Our proposed model uses twelve Sentinel-2 bands and two Sentinel-1 measurements. Our model uses 12 blocks with 384 filters determined by hyperparameter optimization. Before these deep residual blocks, the channels are gradually up-sampled by two layers with 128 and 256 filters.

5.1 Model architecture

The model is partially defined by hyperparameter optimization described in Section 6.3. Our model consists of an entry block that gradually upsamples the number of filters with an interval of 128 units until the desired number of channels is reached. After the entry block, there are n consecutive inverted linear residual blocks. In residual blocks, the input is also passed through a skip connection to the output of the block. By adding the input, the block learns an additive residual function to the identity map of the block. These residuals enable the learning of very deep networks and prevent gradients from vanishing before propagating back through the network (He, Zhang, Ren, & Sun, 2015). Each block expands the number of channels by a factor t. Each linear residual block start with an expansion layer. This is a 2D convectional layer with a kernel size of 1; the amount of output channels is defined by the expansion ratio t times the number of input channels d_{in}. This expansion layer is followed by batch normalization and a ReLU activation function. These layers are followed by a depthwise convolutional layer with a 3x3 kernel size, followed by batch normalization and a ReLU activation function. At the end of the block are a pointwise convolutional layer and batch normalization. The final layer of the network is a single pointwise convolutional layer that combines the d_{in} output channels to a single prediction for each pixel.

5.2 Linear bottlenecks

Linear bottlenecks are built to prevent the usual non-linear per coordinate transformations from losing too much information. Non-linear activation layers, for example, the ReLU function, inevitably lose information in a channel in which it collapses. If there is some structure in the activation space, this information could still be captured in other channels. Linear bottlenecks can embed the input space into a lower-dimensional subspace of the activation space than with the use of non-linear transformations (Sandler et al., 2019). This results in a more efficient model with layers that prevent non-linear transformations from destroying too much information.

5.3 Inverted residuals

The bottleneck block contains an input layer followed by several bottlenecks and expansion. These bottlenecks contain all necessary information. The expansion layer ensures the input layer has the same dimensionality as the output layer, so the residual can be added back to the output. Introducing shortcuts into a neural network architecture aims to improve the gradients' ability to propagate over multiple layers.

5.4 Training

For our regression problem, we learn the networks parameters by minimizing the l_2-loss function only over the measured values by the GEDI instrument. To exclude pixels with no information, we exclude them with an additional term in the loss function. For each pixel, we define m_i to be 1 at location x_i if the GEDI instrument took a measurement in that location and 0 otherwise. We can define the total loss function then as follows:

$$Loss = \frac{1}{\sum_{i=1}^{N} m_i} \sum_{i=1}^{N} ((f(x_i) - y_i) \cdot m_i)^2 + \lambda \frac{1}{W} \sum_{j=1}^{W} (w_j)^2 \tag{5.1}$$

Here f is the neural network with the weights and constant biases w_i. Each x_i is the normalized input intensity, and each y_i is the ground truth canopy heights measured by the GEDI instrument.

$f(x_i)$ at pixel value i is the estimated canopy height at that location. N denotes the number of pixels in the training sample. W denotes the number of parameters in the model. λ is a hyperparameter that controls the strength of the regularisation.

CHAPTER

SIX

EXPERIMENTS

In this chapter, we will describe our experiments. We designed our experiments to answer the following three questions:

- Q1: Can synthetic aperture radar imaging improve canopy height estimations?
- Q2: Can linear bottlenecks and inverse residuals improve canopy height estimations?
- Q3: What accuracy can be achieved through our proposed model between the locations of the measured GEDI signals?

In the case of Q1, we train our proposed model and two additional baselines with SAR data and without and compare their performance. For answering Q2, we compare our proposed model, which uses linear bottlenecks and inverse residuals, against two baselines that do not use these techniques. Finally, to answer Q3, we evaluate the performance on four performance metrics using full-resolution data collected from the LVIS sensor and the data collected by the GEDI sensor. In the remainder of this chapter, we will describe the hyperparameter optimization process used to form the models' architecture and configuration. Next, we will describe the evaluation protocol used to answer our research question.

We will compare our model against two baselines; namely, the model proposed by Lang et al. (2019), which is based on the Xception architecture (Chollet, 2016) and a random forest model. At the time of writing, the random forest model is used by most researchers (Jiang et al., 2021; W. Li et al., 2020). This research focuses on canopy structure in Gabon, Africa, and Costa Rica, Central America. This region was chosen because of research applications under development at the Institute of Environmental Sciences (CML) at Leiden University and the availability of validation data. Canopy structure information at high spatial resolution ($25\ m$) was collected over a large area ($> 7000\ km^2$) in Gabon during the AfriSAR mission in 2016 using the Land Vegetation and Ice Sensor (LVIS) (Fatoyinbo et al., 2021). An additional mission was launched in 2019 to validate GEDI samples. The LVIS sensor was flown over sites in the Southeastern United States and Central America. This study uses the data collected over Costa Rica ($> 5000\ km^2$).

6.1 Baselines

In order to answer our research questions, we will use two baseline models, namely two learning models:

- **Lang's method**; as described by (Lang et al., 2019) and implemented by the TensorFlow library (Abadi et al., 2015). After hyperparameter optimization, the number of units was 512, regularisation was 0.01, and the learning rate was 0.0001. The number of blocks was 10 for the configuration without SAR and 11 for the configuration with SAR.
- **Random forest**; as implemented by the scikit-learn library (Pedregosa et al., 2011). This model is used by W. Li et al. (2020) and Jiang et al. (2021). After hyperparameter optimization, the number of estimators was 1000, max depth was 100, min sample split was 20, and min sample leaf was 20 for the configuration without SAR. The configuration with SAR was as follows; the number of estimators was 980, max depth was 100, min sample split was 9, and min sample leaf was 20.

6.2 Experimental setup

For the first two research questions, the data will be collected over the entirety of Gabon and Costa Rica separately and split into patches of $150 \times 150\ m$. The set of patches will be split into three separate parts. One set for training the models, one set for testing how well the models can estimate canopy heights, and a final set to validate our findings. The area used for the training and test set is strictly non-overlapping with the area for the validation set. Data collected over the entirety of Gabon and Costa Rica, without the areas for our study sites, will be split up with 80% of the data to form the training set and 20% to form the test set (Géron, 2017). To speed up this process, a random subsample of 20% from the training set is used during hyperparameter optimization. After hyperparameter optimization, the resulting network is trained a final time with the entire dataset, tested on the test set, and evaluated in our study sites. For Q3, we will use data collected over Gabon and Costa Rica with the exclusion of the areas for which data were collected during the AfriSAR mission as training data; we will validate the models' performance in the areas for which full-resolution canopy data is available.

6.3 Hyperparameter optimization

As mentioned in Section 5.1, prior to our experiments, hyperparameter optimization was performed on CNNCS and our baselines. The configuration space used for these models is presented in Table 6.1. The hyperparameters were optimized using a Bayesian optimization algorithm (Snoek, Larochelle, & Adams, 2012). This algorithm models the performance of the network with a surrogate function. In our experiments, we use a Gaussian process. The tractable posterior distribution produced by the Gaussian process is used to search the hyperparameter space efficiently. Hyperparameter optimization was continued until all models converged on performance. This was after 14 trials. All models were given the same amount of trials for hyperparameter optimization. For the implementation of this optimization algorithm, we used Keras Tuner (O'Malley et al., 2019).

6.4 Linear bottlenecks and inverse residuals

To answer Q2 and evaluate the effectiveness of linear bottlenecks and inverse residuals. We will set two baselines. The first baseline will be the model proposed by Lang et al. (2019). This model will be trained on Sentinel-1 and 2 data as independent variables and GEDI data as a dependent

Table 6.1: Configuration space for the models used in this reserach.

CNNCS / Lang's method		random forest	
unit size	(64, 768)	estimators	(100, 1000)
blocks	(10, 18)	max depth	(1, 100)
regularisation	[0.0, 0.001, 0.01, 0.1]	min sample split	(2, 20)
learning rate	[0.0001, 0.001, 0.01]	min sample leaf	(1, 20)

variable. The number of residual blocks and the size of these blocks will be tuned using Bayesian optimization. In addition, we will use a random forest regression model. The number of trees will also be optimized using Bayesian optimization. This optimization uses the Bayes theorem to find optimal candidate parameters that minimize our objective function. It does so by utilizing a surrogate function to estimate the objective function. We can, in turn, use this surrogate function based on our belief about the objective function. Every time an additional sample in the search space is evaluated, it is added to the dataset used to train the surrogate function.

6.5 Evaluation

We will compare the predictions $\hat{y}_{c_i} \in \hat{Y}$ made by the model with the ground truth values $y_{c_i} \in Y$ at a given coordinate for interest on four performance metrics. In the following equations, N is the number of examples on which the model is evaluated. The mean absolute error (MAE) will give an interpretable performance measurement and is defined as follows:

$$MAE = \frac{1}{N} \sum |\hat{y}_{c_i} - y_{c_i}| \tag{6.1}$$

In addition, we will measure the models' performance on the root mean square error (RMSE), which is defined as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum (\hat{y}_{c_i} - y_{c_i})^2} \tag{6.2}$$

These metrics will give us an indication of how well the model is capable of estimating canopy height from radar and spectral data. In addition to these two standard error metrics, we also included two metrics that are commonly used in computer vision, namely the peak to signal noise ratio (PSNR) and the structural similarity index (SSIM)

$$PSNR = 20 \cdot \log_{10} \frac{\max(Y)}{\sqrt{\frac{1}{N} \sum (\hat{y}_{c_i} - y_{c_i})^2}} \tag{6.3}$$

PSNR describes the ratio between the maximum power of the true signal $\max(Y)$ and the power of noise. The primary motivation for using PSNR as an evaluation metric is its property to scale values by the maximal signal. It is, therefore, a better metric to compare the different study sites.

$$SSIM = \frac{(2\mu_{\hat{Y}}\mu_Y + c_1)(2\sigma_{\hat{Y}Y} + c_2)}{(\mu_{\hat{Y}}^2 + \mu_Y^2 + c_1)(\sigma_{\hat{Y}}^2 + \sigma_Y^2 + c_2)} \tag{6.4}$$

SSIM measures the similarity between two images. The metric aims to be consistent with human perception. Using structural information instead of the absolute or squared error, this metric aims to better represent the difference in structure between the true labeled image and the estimations. The SSIM is only measured with the images captured with the LVIS sensor and the estimations from models used in our research. The images produced by GEDI contain too few measurements to accurately estimate the actual variance in a given location.

We will repeat the experiments four times to show the significance of the results. We used bootstrapping to generate ten simulated datasets from our evaluation dataset to evaluate each model on (Efron, 1979). The training and test data for Gabon contained $2,504,750$ images of 15×15 containing $11,362,084$ pixels with measurements and the for Costa Rica this was $762,300$ images containing $5,535,057$ pixels with measurements. The experiments were run on a PNY GeForce RTX 2080TI with 12 GB of video memory. The training time for models trained on the Gabon dataset was around three days, while the training time for the models trained on the Costa Rica dataset was 1.5 days.

CHAPTER

SEVEN

RESULTS AND DISCUSSION

The full results for all study sites of our experiments can be found in Apendix A and B. In this chapter, we will compare models that included SAR as an independent variable against models without and answer Q1. In the second section, we will focus on the comparison of our proposed model CNNCS against the other baseline models and answer Q2. Finally, we will look at the overall best performing configuration and analyze the cross-validated results on LVIS data to answer Q3. The results of our experiments are shown in Table 7.1 and 7.2. We performed a Wilcoxon signed-rank test ($\alpha = 0.05$) to determine the significance of our results.

Table 7.1: The Mean absolute error of our proposed model CNNCS, Lang's method, and RF evaluated on our study sites over Gabon and Costa Rica on GEDI. THE Model that significantly outperformed other models in marked in **bold**.

	Gabon		Costa Rica	
	MSI	MSI + SAR	MSI	MSI + SAR
CNNCS	9.01 ± 0.071	8.30 ± 0.072	7.09 ± 0.026	7.10 ± 0.032
Lang's method	8.44 ± 0.058	8.24 ± 0.071	$\mathbf{6.86 \pm 0.019}$	$\mathbf{6.87 \pm 0.031}$
RF	$\mathbf{8.00 \pm 0.067}$	$\mathbf{7.85 \pm 0.053}$	7.65 ± 0.021	7.38 ± 0.025

Table 7.2: The Structural Similarity Index of our proposed model CNNCS, Lang's method, and RF evaluated on our study sites over Gabon and Costa Rica on LVIS. In this table, the mean and 95th percentile are presented. The model that significantly outperformed other models in marked in **bold**.

	Gabon		Costa Rica	
	MSI	MSI + SAR	MSI	MSI + SAR
CNNCS	$\mathbf{0.41 \pm 0.003}$	$\mathbf{0.41 \pm 0.003}$	$\mathbf{0.45 \pm 0.004}$	$\mathbf{0.45 \pm 0.004}$
Lang's method	0.40 ± 0.003	0.40 ± 0.003	0.43 ± 0.003	0.44 ± 0.003
RF	0.39 ± 0.003	0.40 ± 0.003	0.44 ± 0.003	0.44 ± 0.004

7.1 Q1: Syntetic aperature radar

As mentioned in Section 4.4 the reasoning behind using SAR data is based on the reflective properties of the vertical canopy structure. To measure if synthetic aperture radar can improve canopy height estimations. We trained and evaluated our proposed model CNNCS and two other baseline models with and without SAR data. Table 7.1 shows the MAE in Gabon and Costa Rica. The models improved their performance on average with $0.22\ m$. The highest increase in performance was measured in Gabon. Only the RF model increased its performance in Costa Rica. As shown in Appendix A, CNNCS scored better on MAE in 6 out of 9 study sites using SAR data. If we compare the performance of models on SSIM, we do not observe a significant difference in performance between the models that include SAR and those that do not.

7.2 Q2: Linear layers

To address Q2 and investigate if linear bottlenecks and inverse residuals can improve canopy height estimations, we compare our proposed model that uses inverted linear residual blocks against two other baselines. As can be seen in Table 7.1, in general, our proposed model CNNCS does not outperform our baselines if we look at MAE. However, if we look at Table 7.2, our proposed model outperformed the baselines on SSIM. This improvement can also be seen in the confusion plot visualized in Figure 7.2. The confusion plot was created by discretization of the true and estimated values. The random forest shows a clear bias for estimating the mean value for canopy height. This finding further shows the importance of using models that can use the neighborhood of a coordinate of interest as the area of influence for making estimations. Our proposed model shows the best agreement between the ground truth and the estimated heights. Just as Lang et al. (2019) points out, their model degrades towards the top of the height range ($\approx 55\ m$). A more detailed view of performance for different forest height classes is given in Figure 7.1. Our proposed model scores slightly better at the top of the height range while scoring slightly worse at the lower and middle height ranges.

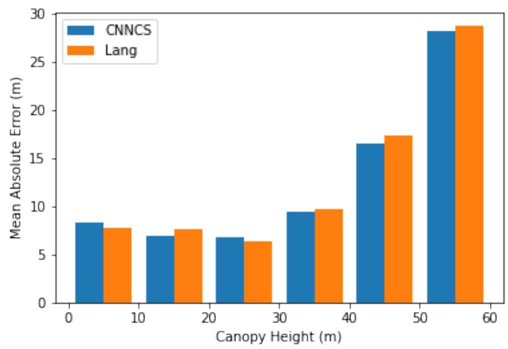


Figure 7.1: Comparison of mean absolute error per $10\ m$ canopy height intervals between two convolutional neural networks.

7.3 Q3: Canopy height estimations

To address Q3 and investigate the accuracy CNNCS can achieve between the measured GEDI points. We evaluate CNNCS on data provided by the LVIS sensors. The results of this experiment can be found in Apendix B. The difference between the acquisition time of data provided by the GEDI sensor and Sentinel-1 and 2 satellites is less than that of the data provided by the LVIS sensor. Because of this difference in acquisition time, we expect the error to be higher when evaluated on LVIS Data. This is, however not the case as can be seen in Apendix A and B. The difference is possibly because in some areas, the distribution of measurements taken by the LVIS sensor is more concentrated on canopy heights below $10\ m$, as can be interpreted from Table 4.1

In Figure 7.3 an evaluation area of $300 \times 300\ m$ above Costa Rica is presented to visualise the comparison between models. At the $10\ m$ resolution, the random forest model distinguishes between

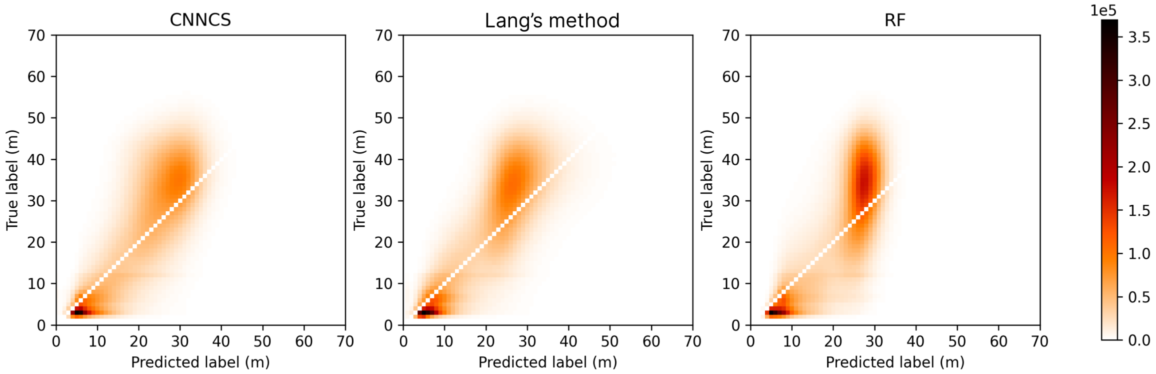


Figure 7.2: Confusion plot for our proposed model CNNCS, Lang's method and RF. The R^2 values of the results are 0.56 for CNNCS, 0.55 for Lang's method and 0.52 for RF.

forest and no-forest at a higher level of detail; however, within forest patches, the random forest model captures the least amount of differences in canopy heights. If we look inside the forest, we can see that the random forest model shows the least difference in height within a forest, while the convolutional models are more capable of showing this difference.

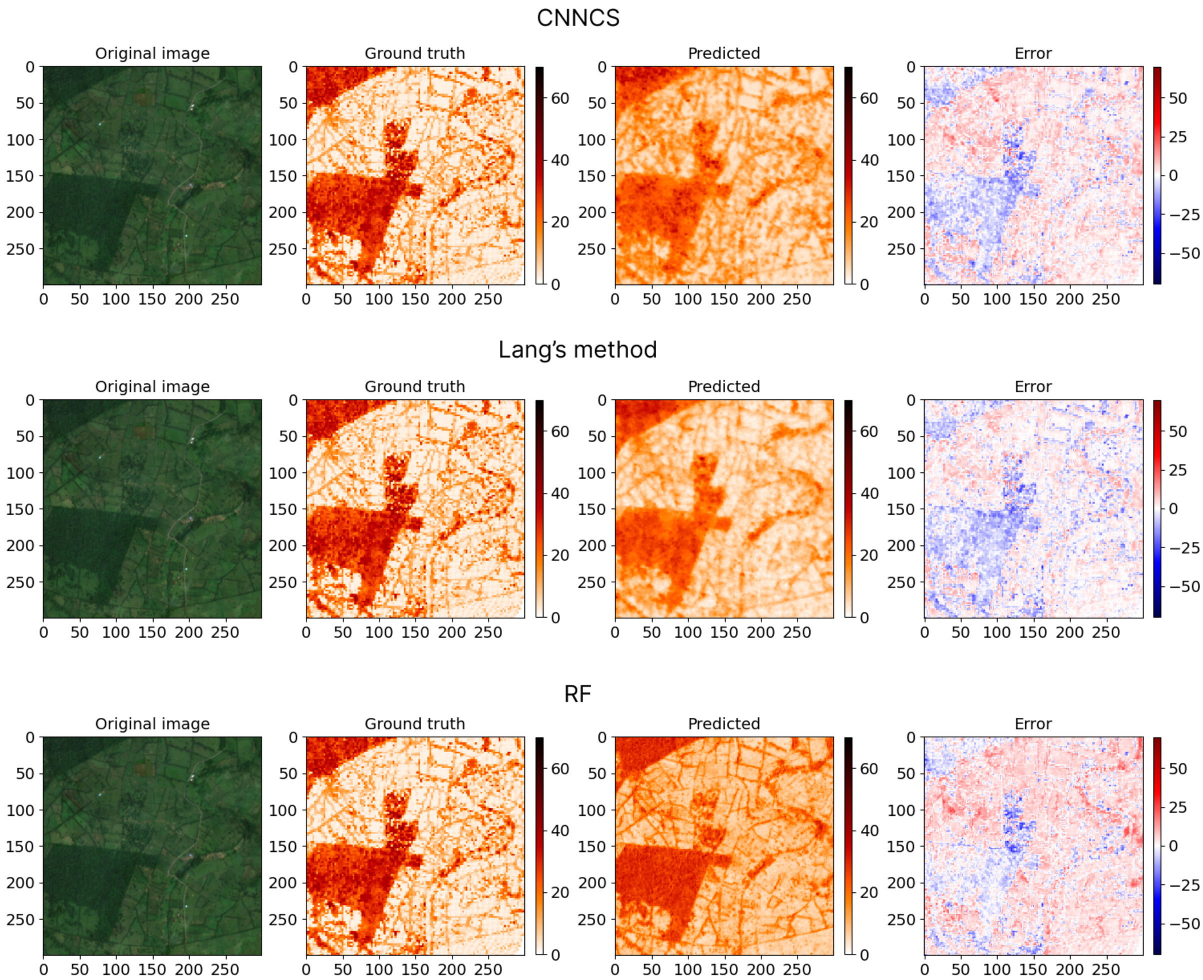


Figure 7.3: Example of estimations made by the models used in this research.

CHAPTER EIGHT

CONCLUSION AND FUTURE WORK

In this thesis, we investigated a semi-supervised learning method for training a deep learning model that fuses Sentinel-1 Synthetic aperture radar (SAR) and Sentinel-2 Multi-spectral imaging (MSI) data to create canopy height maps at a higher resolution than possible using GEDI data alone.

Our experimental results show that SAR data can improve canopy height estimations if the direct neighborhood of an area of interest is taken into account. SAR data is an excellent source to estimate canopy height more precisely but does not improve the structural similarity of the produced estimations. All models used in this research improved their estimations if we looked directly at the height of the estimations when using SAR data.

Additionally, our results show the benefit of using linear residual layers for estimating canopy height metrics. We showed that these linear embedding could benefit the structural similarity between estimated canopy height maps and measurements. This benefit comes at the cost of accuracy. Additionally, our proposed model improves the performance towards the top of the height range in our study sites.

We show that our semi-supervised deep learning method can use GEDI data to learn complex mappings from SAR and MSI data to canopy height maps at $10\ m$ resolutions.

In future work, we want to use our proposed model to estimate the leaf area index (LAI). The LAI is half the total area of green canopy elements per unit of horizontal ground area. In practice, it is often used to quantify the thickness of canopy cover. LAI is an essential climate variable and is used in hundreds of studies on crops (Lobell, Thau, Seifert, Engle, & Little, 2015; Akinseye et al., 2017; Blancon et al., 2019), global water balancing and carbon circulation (Wang, feng Huang, lin Tang, & zhen Wang, 2007) and the study of forests (Xie et al., 2021). Measuring LAI can be labor-intensive, costly, and time-consuming. LAI can also be estimated using lidar sensors like the one used in GEDI (Dubayah et al., 2021). Similarly, with canopy height maps, the resolution of the maps produced by these investigations is not high enough for some applications. Our work demonstrates the power of machine learning and deep learning for these applications.

APPENDIX

A

GEDI EVALUATION

Table A.1: Mean Absolute error for the study sites in Gabon and Costa Rica evaluated on GEDI data. (a) doen not include synthetic aperture radar as an independent variable. (b) does include synthetic aperture radar as independent variable.

	GA1		GA2		GA3		GA4		GA5	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	10.95	8.36	8.43	9.24	8.82	8.27	9.59	8.85	7.29	6.85
Lang's method	7.48	7.28	10.14	9.31	8.34	8.33	8.05	9.76	7.18	7.53
RF	9.06	8.34	8.15	9.05	7.49	7.53	8.15	7.88	7.17	7.18

	CR1		CR2		CR3		CR4	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	6.77	6.81	7.31	7.25	6.28	6.43	8.05	7.94
Lang's method	6.35	6.36	7.01	7.13	6.25	6.07	7.83	7.92
RF	7.37	7.01	8.05	7.82	6.86	6.58	8.33	8.11

Table A.2: Root Mean Square error for the study sites in Gabon and Costa Rica evaluated on GEDI data. (a) doen not include synthetic aperture radar as an independent variable. (b) does include synthetic aperture radar as independent variable.

	GA1		GA2		GA3		GA4		GA5	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	13.35	10.72	10.96	12.01	11.49	10.68	11.95	11.05	9.59	8.77
Lang's method	9.78	9.47	13.04	12.12	10.73	10.68	10.38	12.36	9.01	9.74
RF	11.07	10.36	10.69	11.74	9.49	9.60	10.01	9.88	9.05	8.97

	CR1		CR2		CR3		CR4	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	8.68	8.65	11.69	11.76	9.01	10.57	10.25	10.03
Lang's method	8.24	8.42	11.72	11.76	9.07	8.95	10.12	10.26
RF	9.24	8.92	12.02	11.92	9.35	9.19	10.48	10.19

Table A.3: Peak Signal to Noise Ratio for the study sites in Gabon and Costa Rica evaluated on GEDI data. (a) doen not include synthetic aperture radar as an independent variable. (b) does include synthetic aperture radar as independent variable.

	GA1		GA2		GA3		GA4		GA5	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	0.84	0.91	1.04	1.02	0.93	0.89	0.81	0.88	0.86	0.93
Lang's method	1.01	0.97	0.88	1.05	0.90	0.90	0.90	0.83	0.90	0.81
RF	0.87	0.94	1.09	1.00	0.94	0.90	1.02	0.90	0.87	0.90

	CR1		CR2		CR3		CR4	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	1.06	1.06	1.13	1.09	1.21	1.19	0.86	0.77
Lang's method	1.11	1.04	1.04	1.02	1.19	1.26	0.79	0.77
RF	0.98	1.02	1.09	1.03	1.23	1.29	0.74	0.70

APPENDIX

B

LVIS EVALUATION

Table B.1: Mean Absolute error for the study sites in Gabon and Costa Rica evaluated on LVIS data. (a) doen not include synthetic aperture radar as an independent variable. (b) does include synthetic aperture radar as independent variable.

	GA1		GA2		GA3		GA4		GA5	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	9.83	7.45	8.82	10.93	9.36	8.02	7.17	8.41	5.89	6.34
Lang's method	7.70	7.95	12.80	9.46	8.25	8.30	8.63	8.48	7.98	7.36
RF	7.31	7.15	8.90	10.62	7.14	7.22	8.39	9.23	6.46	7.06

	CR1		CR2		CR3		CR4	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	6.14	6.24	5.22	5.07	5.48	5.62	4.86	5.25
Lang's method	5.91	5.44	4.73	4.88	5.46	5.13	4.87	4.89
RF	7.37	6.79	6.67	6.30	6.32	5.78	6.57	6.36

Table B.2: Root Mean Square error for the study sites in Gabon and Costa Rica evaluated on LVIS data. (a) doen not include synthetic aperture radar as an independent variable. (b) does include synthetic aperture radar as independent variable.

	GA1		GA2		GA3		GA4		GA5	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	12.47	9.87	10.59	13.13	12.52	10.46	9.28	10.56	7.59	8.05
Lang's method	10.18	10.26	14.83	11.56	10.64	10.55	10.58	10.58	9.59	9.07
RF	9.15	9.05	10.71	12.57	9.17	9.34	10.31	11.22	8.11	8.75

	CR1		CR2		CR3		CR4	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	7.83	7.94	6.68	6.64	7.24	11.55	6.41	7.16
Lang's method	7.62	7.18	6.22	6.35	7.20	6.86	6.37	6.55
RF	9.14	8.55	8.44	8.06	8.07	7.47	8.33	8.08

Table B.3: Peak Signal to Noise Ratio for the study sites in Gabon and Costa Rica evaluated on LVIS data. (a) doen not include synthetic aperture radar as an independent variable. (b) does include synthetic aperture radar as independent variable.

	GA1		GA2		GA3		GA4		GA5	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	0.75	0.80	0.92	0.73	0.75	0.85	0.93	0.86	1.00	1.01
Lang's method	0.85	0.77	0.71	0.82	0.79	0.79	0.87	0.84	0.90	0.95
RF	0.82	0.82	0.84	0.73	0.87	0.89	0.87	0.88	0.96	0.94

	CR1		CR2		CR3		CR4	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	1.06	1.00	0.93	0.97	1.10	1.10	1.00	0.96
Lang's method	1.03	1.08	0.98	0.93	1.00	0.98	1.00	1.01
RF	0.93	1.01	0.82	0.89	0.98	1.09	0.86	0.79

Table B.4: Structural Similarity Index for the study sites in Gabon and Costa Rica evaluated on LVIS data. (a) doen not include synthetic aperture radar as an independent variable. (b) does include synthetic aperture radar as independent variable.

	GA1		GA2		GA3		GA4		GA5	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	0.37	0.37	0.54	0.55	0.39	0.38	0.39	0.38	0.37	0.36
Lang's method	0.36	0.36	0.55	0.54	0.37	0.39	0.37	0.38	0.35	0.35
RF	0.36	0.38	0.52	0.54	0.38	0.38	0.37	0.38	0.34	0.34

	CR1		CR2		CR3		CR4	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
CNNCS	0.34	0.33	0.68	0.69	0.64	0.64	0.14	0.12
Lang's method	0.31	0.32	0.66	0.69	0.64	0.63	0.11	0.11
RF	0.33	0.32	0.66	0.66	0.65	0.63	0.13	0.13

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., … Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems.* Retrieved from https://www.tensorflow.org/ (Software available from tensorflow.org)

Abdel-Hamid, O., Mohamed, A.-r., Jiang, H., Deng, L., Penn, G., & Yu, D. (2014). Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *22*(10), 1533-1545. doi: 10.1109/TASLP.2014.2339736

Abshire, J., Sun, X., Riris, H., Sirota, J., Mcgarry, J., Palm, S., … Liiva, P. (2005, 11). Geoscience laser altimeter system (glas) on the icesat mission: On-orbit measurement performance. *Geophys. Res. Lett*, *32*, 21-2. doi: 10.1029/2005GL024028

Akinseye, F. M., Adam, M., Agele, S. O., Hoffmann, M. P., Traore, P., & Whitbread, A. M. (2017). Assessing crop model improvements through comparison of sorghum (sorghum bicolor l. moench) simulation models: a case study of west african varieties. *Field Crops Research*, *201*, 19–31.

Arp, L., Baratchi, M., & Hoos, H. (2022). Vpint: value propagation-based spatial interpolation. *Data Mining and Knowledge Discovery*, 1–32.

Blancon, J., Dutartre, D., Tixier, M.-H., Weiss, M., Comar, A., Praud, S., & Baret, F. (2019). A high-throughput model-assisted method for phenotyping maize green leaf area index dynamics using unmanned aerial vehicle imagery. *Frontiers in plant science*, 685.

Braaten, J. (2020). *Sentinel-2 cloud masking with s2cloudless.* Google. Retrieved from https://developers.google.com/earth-engine/tutorials/community/sentinel-2-s2cloudless

Chollet, F. (2016). Xception: Deep learning with depthwise separable convolutions. *CoRR*, *abs/1610.02357*. Retrieved from http://arxiv.org/abs/1610.02357

Clark, M., Clark, D., & Roberts, D. (2004, 05). Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. *Remote Sensing of Environment*, *91*, 68-89. doi: 10.1016/j.rse.2004.02.008

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., … Silva, C. (2020). The global ecosystem dynamics investigation: High-resolution laser ranging of the earth's forests and topography. *Science of Remote Sensing*, *1*, 100002. Retrieved from https://www.sciencedirect.com/science/article/pii/S2666017220300018 doi: https://doi.org/10.1016/j.srs.2020.100002

Dubayah, R., Hofton, M., Blair, J., Armston, J., Tang, H., & Luthcke, S. (2021, jan). Gedi l2a elevation and height metrics data global footprint level v001 [Computer software manual].

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, *7*(1), 1 – 26. Retrieved from `https://doi.org/10.1214/aos/1176344552` doi: 10.1214/aos/1176344552

Fatoyinbo, T., Armston, J., Simard, M., Saatchi, S., Denbina, M., Lavalle, M., … Hibbard, K. (2021). The nasa afrisar campaign: Airborne sar and lidar measurements of tropical forest structure and biomass in support of current and future space missions. *Remote Sensing of Environment*, *264*, 112533. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0034425721002534` doi: https://doi.org/10.1016/j.rse.2021.112533

Filippi, A. M., İnci Güneralp, & Randall, J. (2014). Hyperspectral remote sensing of aboveground biomass on a river meander bend using multivariate adaptive regression splines and stochastic gradient boosting. *Remote Sensing Letters*, *5*(5), 432-441. Retrieved from `https://doi.org/10.1080/2150704X.2014.915070` doi: 10.1080/2150704X.2014.915070

Fix, E., & Hodges, J. L. (1989). Discriminatory analysis. nonparametric discrimination: Consistency properties. *International Statistical Review / Revue Internationale de Statistique*, *57*(3), 238–247. Retrieved from `http://www.jstor.org/stable/1403797`

Fletcher, K. (2012). *Sentinel-1: Esas radar observatory mission for gmes operational services*. ESA Communications.

Flores, A., Herndon, K., Thapa, R., & Cherrington, E. (2019). *The sar handbook: Comprehensive methodologies for forest monitoring and biomass estimation*. doi: 10.25966/nr2c-s697

Géron, A. (2017). *Hands-on machine learning with scikit-learn and tensorflow : concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O'Reilly Media.

Haining, R. (1978, 01). The moving average model for spatial interaction. *Transactions of the Institute of British Geographers*, *3*. doi: 10.2307/622202

Harder, R. L., & Desmarais, R. N. (1972). Interpolation using surface splines. *Journal of Aircraft*, *9*(2), 189-191. doi: 10.2514/3.44330

Hashimoto, R., & Suto, K. (2020). Sicnn: Spatial interpolation with convolutional neural networks for radio environment mapping. In *2020 international conference on artificial intelligence in information and communication (icaiic)* (p. 167-170). doi: 10.1109/ICAIIC48513.2020.9065029

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. *CoRR*, *abs/1512.03385*. Retrieved from `http://arxiv.org/abs/1512.03385`

Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (Vol. 1, pp. 278–282).

Hofton, M., Blair, B., Story, S., & Yi, D. (2019, Dec). Algorithm theoretical basis document (atbd) for gedi transmit and receive waveform processing for l1 and l2 products. *Technical Report. Goddard Space Flight Center, Greenbelt, MD*.

Jiang, F., Zhao, F., Ma, K., Li, D., & Sun, H. (2021). Mapping the forest canopy height in northern china by synergizing icesat-2 with sentinel-2 using a stacking algorithm. *Remote Sensing*, *13*(8). Retrieved from `https://www.mdpi.com/2072-4292/13/8/1535` doi: 10.3390/rs13081535

King, M. D., Platnick, S., Menzel, W. P., Ackerman, S. A., & Hubanks, P. A. (2013). Spatial and temporal distribution of clouds observed by modis onboard the terra and aqua satellites. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(7), 3826-3852. doi: 10.1109/TGRS.2012.2227333

Krige, D. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, *52*(6), 119-139. doi: 10.10520/AJA0038223X_4792

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*, 1097–1105.

Lang, N., Schindler, K., & Wegner, J. D. (2019, Nov). Country-wide high-resolution vegetation height mapping with sentinel-2. *Remote Sensing of Environment*, *233*, 111347. Retrieved from `http://dx.doi.org/10.1016/j.rse.2019.111347` doi: 10.1016/j.rse.2019.111347

Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., & Chen, H. (2020, 10). High-resolution mapping of forest canopy height using machine learning by coupling icesat-2 lidar with sentinel-1, sentinel-2 and landsat-8 data. *International Journal of Applied Earth Observation and Geoinformation*, *92*, 102163. doi: 10.1016/j.jag.2020.102163

Li, Z., Cao, J., Wang, H., & Zhao, M. (2019). Sparsely self-supervised generative adversarial nets for radio frequency estimation. *IEEE Journal on Selected Areas in Communications*, *37*(11), 2428-2442. doi: 10.1109/JSAC.2019.2933779

Lobell, D. B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment*, *164*, 324–333.

Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., … Gascon, F. (2016, 08). Sentinel-2 sen2cor: L2a processor for users..

Marselis, S. M., Tang, H., Armston, J., Abernethy, K., Alonso, A., Barbier, N., … Dubayah, R. (2019, sep). Exploring the relation between remotely sensed vertical canopy structure and tree species diversity in gabon. *Environmental Research Letters*, *14*(9), 094013. Retrieved from `https://doi.org/10.1088/1748-9326/ab2dcd` doi: 10.1088/1748-9326/ab2dcd

Mitchard, E. T. A. (2018, Jul 01). The tropical forest carbon cycle and climate change. *Nature*, *559*(7715), 527-534. Retrieved from `https://doi.org/10.1038/s41586-018-0300-2` doi: 10.1038/s41586-018-0300-2

Næsset, E. (1997). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *52*(2), 49-56. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0924271697830006` doi: https://doi.org/10.1016/S0924-2716(97)83000-6

O'Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al. (2019). *Keras Tuner.* `https://github.com/keras-team/keras-tuner`.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Saha, S., Saha, M., Mukherjee, K., Arabameri, A., Ngo, P. T. T., & Paul, G. C. (2020). Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest, reptree: A case study at the gumani river basin, india. *Science of The Total Environment*, *730*, 139197. Retrieved from https://www.sciencedirect.com/science/article/pii/S0048969720327145 doi: https://doi.org/10.1016/j.scitotenv.2020.139197

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2019). *Mobilenetv2: Inverted residuals and linear bottlenecks.*

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., … Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, *362*(6419), 1140-1144. Retrieved from https://www.science.org/doi/abs/10.1126/science.aar6404 doi: 10.1126/science.aar6404

Snoek, J., Larochelle, H., & Adams, R. P. (2012). *Practical bayesian optimization of machine learning algorithms.* arXiv. Retrieved from https://arxiv.org/abs/1206.2944 doi: 10.48550/ARXIV.1206.2944

St-Onge, B., Treitz, P., & Wulder, M. (2003, 01). Tree and canopy height estimation with scanning lidar. *Remote sensing of forest environments: concepts and case studies*, 489-509. doi: 10.1007/978-1-4615-0306-4_19

Sundar, S., Sharma, R., & Naresh, R. (2013, 12). Modeling the role of cloud density on the removal of gaseous pollutants and particulate matters from the atmosphere. *Applications and Applied Mathematics:An int.Journal*, *8*, 416-435.

Vaglio Laurin, G., Balling, J., Corona, P., Mattioli, W., Papale, D., Puletti, N., … Urban, M. (2018, 01). Above-ground biomass prediction by sentinel-1 multitemporal data in central italy with integration of alos2 and sentinel-2 data. *Journal of Applied Remote Sensing*, *12*, 1. doi: 10.1117/1.JRS.12.016008

Veci, L., Prats-Iraola, P., Scheiber, R., Collard, F., Fomferra, N., & Engdahl, M. (2014, 07). The sentinel-1 toolbox. In (p. 1-3).

Wang, F., feng Huang, J., lin Tang, Y., & zhen Wang, X. (2007). New vegetation index and its application in estimating leaf area index of rice. *Rice Science*, *14*(3), 195-203. Retrieved from https://www.sciencedirect.com/science/article/pii/S1672630807600274 doi: https://doi.org/10.1016/S1672-6308(07)60027-4

Xie, R., Darvishzadeh, R., Skidmore, A. K., Heurich, M., Holzwarth, S., Gara, T. W., & Reusen, I. (2021). Mapping leaf area index in a mixed temperate forest using fenix airborne hyperspectral data and gaussian processes regression. *International Journal of Applied Earth Observation and Geoinformation*, *95*, 102242. Retrieved from https://www.sciencedirect.com/science/article/pii/S0303243420308850 doi: https://doi.org/10.1016/j.jag.2020.102242

Zhu, X., Wu, F., Wu, T., & Zhao, C. (2017). Remote sensing imaging simulation and cloud removal. *International Journal of Modern Physics B*, *31*(16-19), 1744079. Retrieved from https://doi.org/10.1142/S0217979217440799 doi: 10.1142/S0217979217440799