UNIVERSITÀ DEGLI STUDI
DI TRENTO

DEPARTMENT OF INFORMATION ENGINEERING AND COMPUTER SCIENCE
IECS Doctoral School

Advanced Techniques For Prediction of Forest Above Ground Biomass Using Satellite Remote Sensing Data

Parth Rajubhai Naik

Advisors

Prof. Dr. Lorenzo Bruzzone

Full Professor, Università degli Studi di Trento

Dr. Michele Dalponte

Senior Researcher, Fondazione Edmund Mach

January 2023

Abstract

Large-scale mapping of forest Aboveground Biomass (AGB) is a challenging task and crucial for forest management and planning. The use of Satellite Remote Sensing (SRS) data has recently increased for AGB prediction due to their large footprint and low cost availability. There are various limitations and problems with SRS data that require innovative and effective solutions for large-scale AGB mapping. This thesis provides three main contributions in the context of using SRS for AGB prediction.

The first contribution of the thesis involves the use of Satellite Multispectral (SMS) data characterized by different spectral specifications, spatial resolutions and temporal availability. A systematic framework involving an adaptive regularization method was implemented to observe and quantify the response linked to various characteristics of the SMS data. The second contribution presents a dynamic generative neural network architecture for modelling AGB using multi-sensor satellite RS data. It proposes a method to derive AGB-oriented features and provides a seamless multi-sensor feature fusion method for AGB prediction. The third contribution presents a framework developed from the combination of a hyperparameter optimization procedure and a meta-learning algorithm to set up an end-to-end automated pipeline for modelling AGB. The contribution focuses on automatic development and extraction of features from MS data as well as automatic stacking of algorithms to compose an optimal ensemble model.

The comprehensive analysis for each contribution was based on quantitative and qualitative results from the performed experiments. The first contribution pin-pointed the strengths and shortcomings of SMS data for AGB prediction in terms of effective spectral channels, effect of temporal information and role of spatial resolution for AGB prediction. The second contribution demonstrates the effectiveness of the proposed generative process in producing features that deliver more accurate AGB predictions as compared to the conventional approaches. Lastly, the third contribution generated automated features and stacked ensemble of models that outperformed individual models. The systematic series of experiments and flow of studies confirm that SRS data can be effectively used for accurately modelling AGB using the advanced methods demonstrated in the thesis.

Acknowledgements

I am firstly very grateful to my parents for believing in my capabilities and constantly motivating me to achieve the highest degree of education. Without their love and support, I would be far from pursuing a doctoral degree. All gratitude would be incomplete without mentioning Ganpati Bappa who answers all prayers of my Mother. A special thanks to my very affectionate Grandparents who are proud of my academic career. I am also really blessed to have a loving sister who in spite of her irritating quality keeps a jolly environment around me at all times.

I am lucky to receive the best possible guidance from my supervisors - Prof. Lorenzo Bruzzone and Dr. Michele Dalponte. In spite of a busy schedule, they both provided a constant attention to my PhD activities. Dr. Michele assisted in providing a suitable direction to my work with his exceptional experience in the forestry and remote sensing. I especially appreciate his patience and commitment towards me. Prof. Bruzzone on the other side provided his visionary insights on all work aspects and helped me visualize the bigger picture. His unconventional work standards and achievements are a great source of inspiration to me.

The only person who contributed to both my work and personal growth is my partner - Rupsa Chakraborty. She motivated me when I felt low and kept me going during my difficult times. I also thank my friends and colleagues Aravind, Aniruddha, Abhishek, Sudipan, Sanchari, Akshara and Valentina for being supportive in different ways and making it a smooth journey for me.

Parth Rajubhai Naik

The work compiled in this thesis has been fully supported by the research grants from the University of Trento and Fondazione Edmund Mach, Italy

Dedicated to my Mom & Dad

Contents

1 Introduction **3**

1.1 Background 3

1.2 Motivation 7

1.3 Research Problems 11

1.4 Objectives and Novel Aspects of the thesis 13

1.5 Structure of the Thesis 17

2 State of the Art **19**

2.1 State-of-the-art for AGB prediction using SMS data 20

2.2 State-of-the-art for AGB prediction using satellite SAR data 22

2.3 State-of-the-art for AGB prediction using combined MS and SAR data 24

3 Multitemporal Multispectral Satellite Remote Sensing Data for AGB Prediction **27**

3.1 Introduction 27

3.2 Study Area and Datasets 30

3.2.1 Study Area 30

3.2.2 Field Data 30

3.2.3 Remote Sensing Data 32

3.3 Methodology 34

3.3.1 Data Pre-processing 34

3.3.2 Feature Extraction 35
3.3.3 Adapative Lasso and Generalized Linear Modelling 35
3.3.4 Design of Experiments 38
3.4 Results and Discussion 40
3.4.1 Temporal Analysis 40
3.4.2 Spectral Analysis 41
3.4.3 Spatial Analysis and Mapping AGB 43
3.4.4 Discussion 45
3.5 Conclusion 52

4 Generative Features from Multi-source Satellite Remote Sensing Data for AGB Prediction 55
4.1 Introduction 55
4.2 Study Area and Datasets 58
4.2.1 Study Area Description and Field Data 58
4.2.2 Remote Sensing Data 59
4.3 Proposed Generative Model 60
4.3.1 Generative Model Framework 60
4.3.2 Generative Network Architecture 63
4.4 Proposed Approach for AGB Prediction 65
4.4.1 Data Pre-processing 65
4.4.2 Computed Features and Input Data Preparation . . 67
4.4.3 Model Implementation and Experiments 69
4.5 Results and Discussion 71
4.5.1 Analytical Features Selected Post A-L1 Regularization 71
4.5.2 Predictive Analysis Using Developed Models 72
4.5.3 Two dimensional t-SNE Visualization of Latent Space 76
4.5.4 Discussion 79
4.6 Conclusion 86

5 Automated Machine Learning for Stacked Ensemble Model Development for Forest AGB Prediction **87**
5.1 Introduction 87
5.2 Study Area and Datasets 92
5.2.1 Study Area Description and Field Data 92
5.2.2 Remote Sensing Data 93
5.3 Description of Algorithms 93
5.3.1 Tree Structured Parzen Estimator Algorithm for Automated Feature Extraction 94
5.3.2 Super Learner Algorithm for Automated Ensemble Modelling 96
5.4 Proposed Approach 97
5.4.1 Data Pre-Processing 97
5.4.2 Implementation of Algorithms 99
5.4.3 Model Explanations 102
5.5 Results and Discussion 103
5.5.1 TPE Optimized Features and Feature Importance . 103
5.5.2 AutoML Leaderboard and Predictive Analysis . . . 106
5.5.3 Discussion 109
5.6 Conclusion 112

6 Conclusion and Future Developments **115**

List of Publications **123**

Bibliography **125**

List of Tables

3.1 Field Estimated Mean tree height (m), mean tree DBH (cm), and number of trees per plot for Lavarone and Pellizzano . 32

3.2 Acquisition dates (YYYY-MM-DD) of the multi-scale SMS data used for the study (Sentinel-2 at 10m, RapidEye at 5m and Dove at 3m) 33

3.3 Vegetation indices extracted from SMS data used as features for modelling and their equations 36

3.4 Assessment metrics for assessment of Precision, Agreement and Overfitting of the models computed using Leave-One-Out Cross-Validation (LOOCV) 40

3.5 Selected features using adaptive least absolute shrinkage and selection operator for Group 1, 2 and 3. 45

4.1 Acquisition Dates (YYY-MM-DD) of the four Sentinel-2 and Sentinel-1 Images acquired over different seasons 59

4.2 Analytical Features From Sentinel-1 SAR Data extracted using First and Second Order Scattering Information with their respective Equations 68

4.3 Validation Statistics for Assessment of Prediction Accuracy, Precision, Agreement and Overfitting of Models 70

4.4 List of Selected Analytical Features from MS, SAR and DS data post A-L1 Regularization 73

4.5 Results of Performance Metrics computed for Proposed Model and Conventional Models (GLM, MLP, RF and XGBoost) developed using MS, SAR and DS data 74

5.1 Acquisition dates (YYYY-MM-DD) of the Sentinel-2 images acquired over the four different seasons 94
5.2 List of 33 Empirical Equations of spectral indices categorized depending on the number of spectral bands used for the TPE algorithm 100
5.3 List of Base models considered for the SL algorithm provided by H2O AI Cloud Platform 101
5.4 TPE Optimized Features for each season computed as per Empirical Equations Indexed as I_n stated in Table 5.2 . . . 104
5.5 Model Leaderboard rankings for SEM and Base Models with six Assessment Metrics 107
5.6 Forest plot-type (Broadleaves, Coniferous and Mixed) based Prediction RMSE for developed models 110

6.1 Coefficients of the allometric equations of Scrinzi et al. (2010) and wood densities (WD) from (IPCC 2003). The wood density (WD) is expressed in kg/m3 149

List of Figures

1.1 Cost and Time Effectiveness versus Measurement Uncertainty of various acquistion modes (Field, UAV, Airborne and Satellites) for Global AGB mapping. 4
1.2 Multiscale (10, 5 and 3 m) SRS data as compared to Airborne Lidar Data for Forest Mapping. 5
1.3 Increasing Trend in Number of Earth Observation Satellites launched from 2013 – 2020 (Source: Northern Sky Research). 8
1.4 Market Share Projections (2021-27) for Satellite Based Earth Observation in Various Areas (Source: Euroconsult). . . . 9
1.5 Publication and Citation statistics of studies on use of AI for Biomass Estimation (Source: Web of Science) 10

3.1 Map of the selected study areas - Pellizzano and Lavarone in Trentino Alto-Adige, Italy (Satellite images—Sentinel2) 31
3.2 Methodology Flowchart describing steps for adaptive regularization and modelling using MS satellite data 39
3.3 Precision Box Plots based on assessment metrics - MAD% and RMSD% computed for single-time and multitemporal models 42
3.4 Agreement Box Plots based on assessment metrics - R2cv and R2fit computed for single-time and multitemporal models 43
3.5 Overfitting Box Plots based on assessment metrics - R2R and SSR computed for single-time and multitemporal models 44

3.6 Multitemporal Regression Plots representing Field Estimated Versus RS Predicted AGB using FS1 Feature Subset —Lavarone area 46
3.7 Multitemporal Regression Plots representing Field Estimated Versus RS Predicted AGB using FS1 Feature Subset —Pellizzano area 47
3.8 Scatterplots of AGB Versus Green Leaf Index (GLI) computed at different spatial scales (10, 5 and 3 m) using FS2 subset for Lavarone area 48
3.9 Scatterplots of AGB Versus Visible Index Green (VIgreen) computed at different spatial scales (10, 5 and 3 m) using FS2 subset for Pellizzano area 49
3.10 Scatterplots of AGB Versus Canopy Chlorophyll Content Index (CCCI) computed at different spatial scales (10, 5 and 3 m) using FS2 subset for Lavarone area 50
3.11 Scatterplots of AGB Versus Red Edge NDVI (NDVIre) computed at different spatial scales (10, 5 and 3 m) using FS2 subset for Pellizzano area. 51
3.12 Spatial Map of AGB predicted using the best MT Sentinel-2 model for Lavarone area 52
3.13 Spatial Map of AGB predicted using the best MT Sentinel-2 model for Pellizzano area 53

4.1 Location of the three study sites (Pellizzano, Lavarone and Cembra) and the plot locations of field estimated AGB (yellow dots) on Google Earth images 60
4.2 Standard False color composite of Sentinel-2 image (left) and Polarimetric matrix element (C_{11}) of Sentinel-1 image (right) of the study area acquired in the summer season . . 61

4.3 Framework of the Proposed Generative Model demonstrating the triple loss mechanism to regulate inference and generative parameters of the network 63

4.4 Dynamic architecture of the proposed generative network with variable layer dimensions R^n: Network layer of dimension 'n'; D: Dense network connection; λ: Reparametrization function; (z, r): Sampled latent vectors; (μ_z, σ_z): mean and variance of 'z'; (μ_r, σ_r): mean and variance of 'r' . . . 64

4.5 Flowchart of the Proposed Approach showing steps for Data Pre-processing, Feature computation and Model Implementation BOA: Bottom of Atmosphere; ARD: Analysis Ready Data; GLM: Generalized Linear Modelling; MLP: Multilayer Perceptron Regressor; RF: Random Forest; XGBoost: Extreme Gradient Boosting; M: Proposed Generative Model; i: Input; i': reconstructed; l: latent space; t: target; RL: Reconstruction loss; KL: KL loss; LL: Label 66

4.6 Regression scatterplots of the Field Estimated versus Model Predicted AGB for the proposed model using single source and DS data 75

4.7 Regression scatterplots of the Field Estimated versus Model Predicted AGB for the MLP model using single source and DS data 76

4.8 Regression scatterplots of the Field Estimated versus Model Predicted AGB for the GLM model using single source and DS data 77

4.9 Regression scatterplots of the Field Estimated versus Model Predicted AGB for the XGB model using single source and DS data 78

4.10 Regression scatterplots of the Field Estimated versus Model Predicted AGB for the RF model using single source and DS data 79
4.11 Color-coded t-SNE plots of reduced dimensionality latent space and input feature space of MS, SAR and DS data . . 80
4.12 AGB Map of the study area produced using the Best Proposed Model and location of the independent reference plots 81
4.13 Correlation assessment of the AGB predicted using the Best Proposed Model and Field Estimated AGB for independent reference plots 82
4.14 Feature Importance graph computed using an independent Gradient Boosting Machine algorithm for Generative and Analytical features 83

5.1 Map of the Study Area with locations of the reference AGB plots indicated as red dots 93
5.2 A flowchart of the Proposed Approach demonstrating sequential implementation of the TPE and SL algorithms for AGB prediction 98
5.3 Stacked Bar Chart of Scaled Feature Importance computed for TPE Optimized Features for all Base Models 106
5.4 Regression scatterplots of Field Estimated AGB versus Predicted AGB for SEM and Base Models 108

List of Abbreviations

A-L1	Adaptive L1
AGB	Aboveground Biomass
AI	Aritifial Intelligence
ANN	Artificial Neural Network
ARD	Analysis Ready Data
AutoML	Automated Machine Learning
DBH	Diameter at Breast Height
DNN	Deep Neural Net
DRF	Distributed Random Forest
DS	Dual-source
EO	Earth Observation
GAN	Generative Adversarial Networks
GBM	Gradient Boosting Machine
GLCM	Grey Level Co-occurence Matrix
GLI	Green Leaf Index
GLM	Generalized Linear Model
KL	Kullback-Leibler
lidar	Light Detection and Ranging
LOOCV	Leave-one-out Cross-validation
ML	Machine Learning
MLP	Multi-layer Perceptron
MS	Multispectral

MT	Multitemporal
MT-MS	Multitemporal-Multispectral
NAS	Neural Architecture Search
NIR	Near-Infrared
PSMS	Particle Swarm Model Selection
PSO	Particle Swarm Optimization
RF	Random Forest
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SEM	Stacked Ensemble Model
SL	Super Learner
SMAC	Sequential Model-Based Algorithm Configuration
SMBO	Sequential Model Based Optimization
SMS	Satellite Multispectral
SRS	Satellite Remote Sensing
SVM	Support Vector Machine
SWIR	Short-wave Infrared
t-SNE	t-distributed Stochastic Neighbor Embedding
TPE	Tree structured Parzen Estimator
TPOT	Tree-Based Pipeline Optimization Tool
VRE	Vegetation Red-Edge
XGBoost	eXtreme Gradient Boosting Machine

List of Symbols

$(Y_i - X_i^T)$	Random Sample with Identical Distribution
$(\mathrm{C}_{11}, \mathrm{C}_{12}, \mathrm{C}_{21}, \mathrm{C}_{22})$	Polarimetric Matrix Elements
$EI_{y^*}(x)$	TPE Expected Improvement
$L(O, \psi)$	Super Learner Loss Function
MAE	Mean Absolute Error
P_o	Super Learner Data Distribution
$R2R$	R-squared Ratio
$RMSD$	Root Mean Squared Difference
R^2_{cv}	Coefficient of determination via Cross-Validation
R^2_{fit}	Coefficient of Determination via Residuals
SS^{CV}	Cross-validated Sum of Squared Differences
SS_{tot}	Sum of Squared Differences of each observation from overall mean
β	Regression Coefficients
$\ell(x)$ and $g(x)$	TPE Densities
γ	Weight Vector Adjustment Constant
$\hat{\beta}^{LS}$	Adaptive Lasso Estimator
$\hat{\beta}^{ini}_j$	Ridge Regression Coefficient

$\hat{k}$	Super Learner Cross-validation Selector
λ	Tuning Parameter
λ_1, λ_2	Eigen Values of Covarinace Matrix
$\mathcal{L}$	List of Modelling Algorithms for AutoML
$\mathcal{L}(i)$	Total Loss of Generative Network
C_2	SAR Covariance Matrix
ω_j	Adaptive Weight Vector
$\psi_0(A)$	Optimal Value of Parameter of Interest
σ^0_{VH}	Cross Polarized Backscattering Coefficient
σ^0_{VV}	Co-polarized Backscattering Coefficients
d_n	Super Learner Expected Difference
k	Scattering Vector
obs_i	Observed value of sample 'i'
$p(i^{'} \mid s)$	Decoder network
$p(s \mid t)$	Latent Generator
$p(x \mid y)$	TPE Density Split
pre_i^{CV}	Cross-validated Prediction value of sample 'i'
$q(s \mid i)$	Probabilistic Encoder
$q(t \mid i)$	Probabilistic Regressor
x	TPE Parameters
y	TPE Loss

Chapter 1

Introduction

1.1 Background

The emission of carbon dioxide that leads to climate change is a compelling reason for acquiring accurate and large-scale information on forests. The amount of carbon stored in the trees is directly proportional to the forest AGB. Human activities have continuously accelerated the release of carbon dioxide in the atmosphere. As a result, AGB is identified as an essential climate variable [1, 2]. Moreover, the problem of forest fragmentation that affects biodiversity can be studied using dynamics of forest AGB [3, 4]. Therefore, a consistent, high resolution and large-scale AGB mapping is a critical requirement for ecological conservation. AGB mapping is also important for energy and resource management that directly affects the economy. Accurate AGB predictions are required to regulate harvesting of large forest areas. The most typical methods to estimate forest AGB are in-situ measurements performed by regional or national forest authorities. These field measurements are valuable for providing accurate estimates of forest biomass that are based on allometric equations. However, such measurements are used at the local scale and are costly and difficult to be acquired over large forest areas. Thus, in-situ measurements have a limited coverage and cannot provide spatially explicit information on forest AGB.

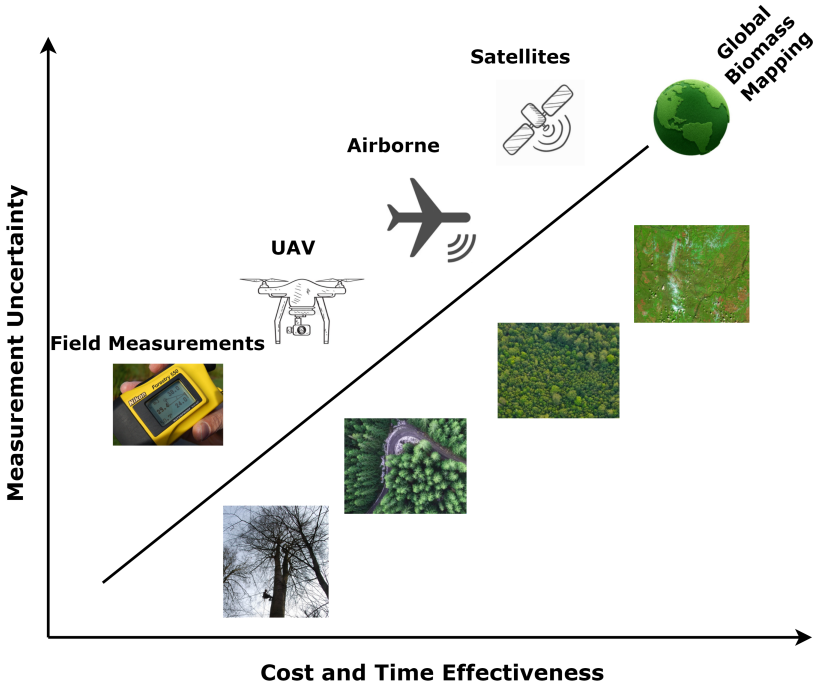

Figure 1.1: Cost and Time Effectiveness versus Measurement Uncertainty of various acquistion modes (Field, UAV, Airborne and Satellites) for Global AGB mapping.

Satellite Remote Sensing (SRS) based Earth Observation (EO) data is an optimal source for AGB estimation in terms of spatial coverage, cost effectiveness and temporal availability at the regional or global scale (Figure 1.1). Airborne Light Detection and Ranging (lidar) data that provides 3D structural information of forest are very reliable for modelling AGB. However, they are expensive and their acquisition for large forest areas is restricted to certain countries. The modelling and mapping of AGB at a regional or global scale requires data that are economically viable with a comprehensive coverage and availability. A surge in the availability of satellite EO data has increased the number of studies conducted to analyze its performance for mapping forest AGB [5, 6, 7, 8]. These attempts focus on increasing reliability of cheap and extensively available SRS data (e.g., MS, SAR) as compared to the expensive and globally limited datasets (e.g.,

lidar). Many studies have utilized the Landsat archive as a time-series resource for monitoring forest carbon stocks [5, 9, 10]. The Sentinel missions – Sentinel-1 (C-band SAR) and Sentinel-2 multispectral images have also been used for prediction and quantification of forest AGB [11, 12, 13]. Recently, the private sector has revolutionized the space industry by small satellite missions (e.g., RapidEye, Dove, Iceye, etc.). The use of small satellite technology can be effective in developing economical solutions for ecological problems. The latest developments suggest that small satellite data can be cost effective and efficient to quantify and map forest AGB at large scale [14].

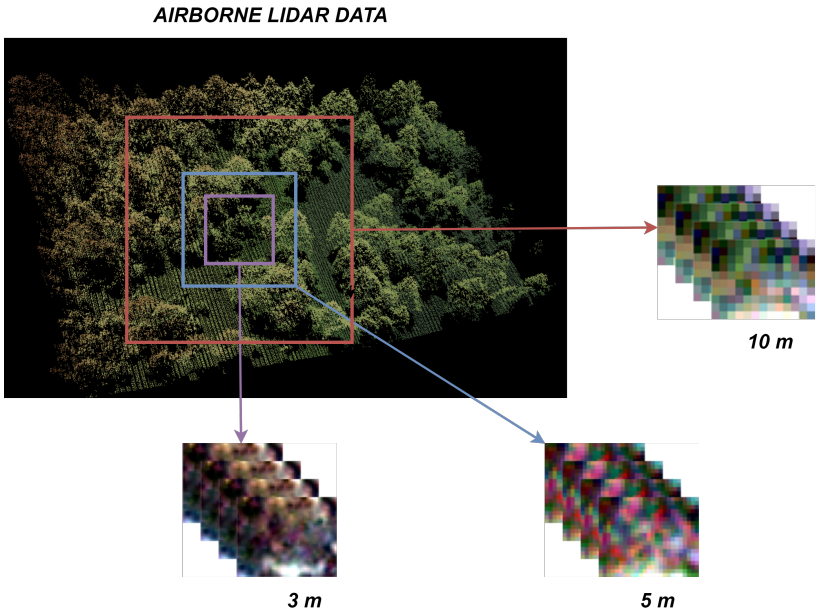

Figure 1.2: Multiscale (10, 5 and 3 m) SRS data as compared to Airborne Lidar Data for Forest Mapping.

Multispectral (MS) data holds a significant share of the recently developed spaceborne EO data capacity. SMS data are available at various spatial scales with a high temporal coverage of the Earth (Figure 1.2). A high temporal coverage has provided a greater possibility for collection of cloud-free images. A long life-span and high repetition rate of satellites enable

long-term forest monitoring. Apart from the spatio-spectral information, temporal information can be derived and used for AGB modelling. In literature, a varying degree of success has been achieved by different studies for modelling AGB using MS data. Landsat derived indices achieved a 25% - 30% accuracy for tropical forests of Brazil and 43% - 72% for commercial forests of Africa [15, 16]. There are consistent attempts being made to develop techniques to improve performance and achieve reliable AGB predictions with SMS data [17, 18, 19, 20]. With regards to Synthetic Aperture Radar (SAR) data, the interaction of radar waves with tree scattering elements establishes a physical relation with AGB. The sensitivity of AGB depends on the radar wavelength. Long wavelength SAR (P, L bands) are more sensitive to AGB as compared to short wavelength SAR (C band). In the current scenario, there is no availability of long-wavelength open-source SAR and the cost of commercial SAR is relatively high. Recent developments attempt to combine short-wavelength SAR (e.g., Sentinel-1 C-band) with MS data for implementing advanced modelling methods to improve AGB predictions [21, 22, 23, 24, 25].

The key elements for deriving accurate AGB predictions from SRS data are the extraction of targeted features and selection of a suitable modelling approach. Spectral bands (especially VNIR and SWIR) hold information on forest structure, texture and pigmentation that are correlated with AGB. Different spectral features (e.g., spectral bands, vegetation indices, transformed images) and spatial features (e.g., textural images) from MS data are used to model and predict forest AGB [12, 26, 27]. It was found that vegetation indices formulated using near-infrared wavelengths are weakly related to AGB as compared to shortwave infrared wavelengths for forest areas with complex stand structures [6]. The Grey Level Co-occurence Matrix (GLCM) based texture measures are typically used to examine the relationships between AGB and textural images [28]. In case

of SAR data, features such as backscatter amplitudes and phases are commonly used for AGB modelling [29, 30]. Many techniques involving SAR data are relatively complex and do not follow the traditional approach of modelling. For example, InSAR technique requires collection of coherence data over a short time increment from two identical instruments and the PolInSAR technique combines varying polarization of radar signals along with analysis of phase differences to produce differential range for height-based biomass estimation [31]. The traditional modelling approaches constitute of using parametric and non-parametric algorithms for modelling AGB. Parametric algorithms require a specific pre-determined set of parameters of fixed size to predict AGB [32]. The most conventional parametric algorithms for AGB modelling are correlation coefficient analysis, stepwise regression and spatial co-simulation algorithms [6, 33, 34]. In contrast, non-parametric algorithms do not explicitly pre-define the model structure and instead follows a data-driven strategy [35]. The most widely used non-parametric algorithms for AGB modelling are k-NN, Regression tree, Random forest, Support vector machines and Artificial neural networks [36].

1.2 Motivation

Mapping AGB at large-scale is a challenging task that requires to solve problems related to data, features and process of modelling. The foremost requirement is data availability as forests cover a very large area on the Earth. The recent trends indicate a steady growth in number of launched EO satellites and promising prospects for the future (Figure 1.3). The recent participation of private sector and advances in small satellite technology has created a huge scope for developing frameworks to handle big satellite data and retain relevant information. It is observed that a signif-

icant proportion of EO satellites in space are small satellites. According to a recent survey (Source: pixalytics.com), more than 350 of the 857 EO satellites are optical or MS imaging satellites. These satellites provide a daily acquisition capacity at a high spatial and spectral resolution. Thus, an increased availability, fine specifications and ease of access to the satellite imaging data are prime motivations for using them to map AGB at a large scale.

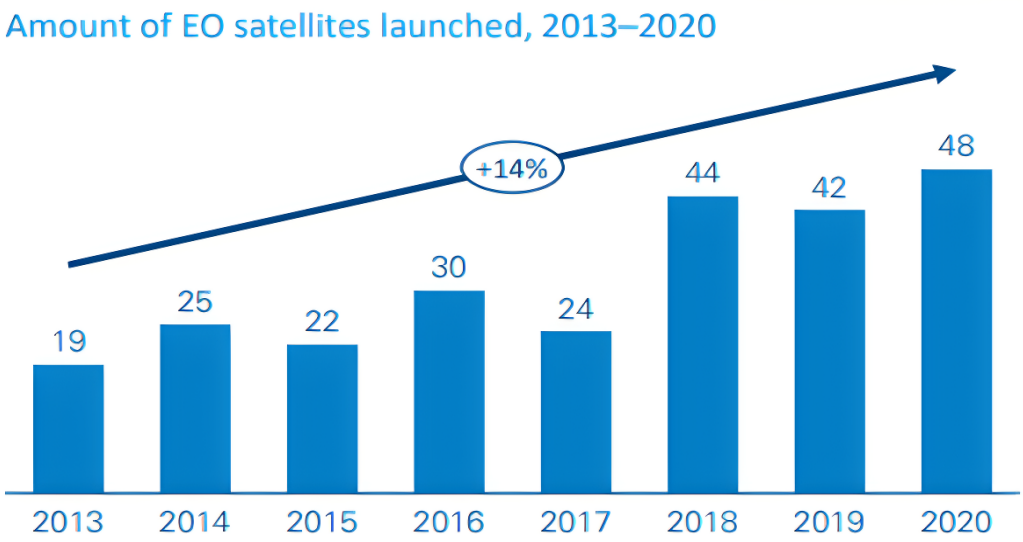

Figure 1.3: Increasing Trend in Number of Earth Observation Satellites launched from 2013 – 2020 (Source: Northern Sky Research).

The economic prospects related to use of satellite EO data are very promising for natural resources and environmental monitoring. Natural resource management and environmental monitoring together are expected to contribute about 18% of the total $98 billion satellite EO economy by 2027 (Figure 1.4). Moreover, the net-zero transition goals present immense opportunities and hold a very huge economic potential. The market size related to carbon management are expanding to support low emission by means of green construction and manufacturing. This presents a great motivation from an economic perspective for satellite based AGB monitor-

ing and mapping. This also creates a requirement for automation, development and maintenance of models for efficient performance over multi-source satellite data. The feature extraction and modelling techniques need to be more accessible and reproducible for large scale AGB mapping. The traditional approaches should shift from isolated individual experiments to collaborative scaled up frameworks that are accessible to the scientific community.

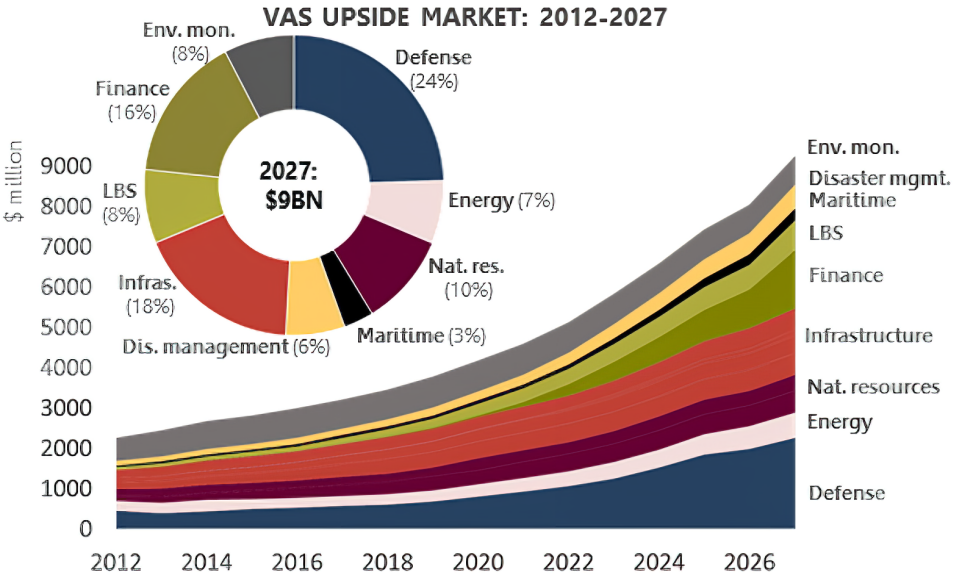

Figure 1.4: Market Share Projections (2021-27) for Satellite Based Earth Observation in Various Areas (Source: Euroconsult).

At present, there are multiple approaches and methods proposed in literature for AGB prediction. Some of them are exclusively designed for certain types of forest, tree distribution or type of tree species. Hence, there is a growing need to develop methods that are more generalized and adaptable to wider AGB modelling scenarios. Moreover, instead of isolated training and evaluation of different models, a common unified framework can provide better basis for comparison and generalization of models. Domain adaptation and similar Aritifial Intelligence (AI) based techniques are

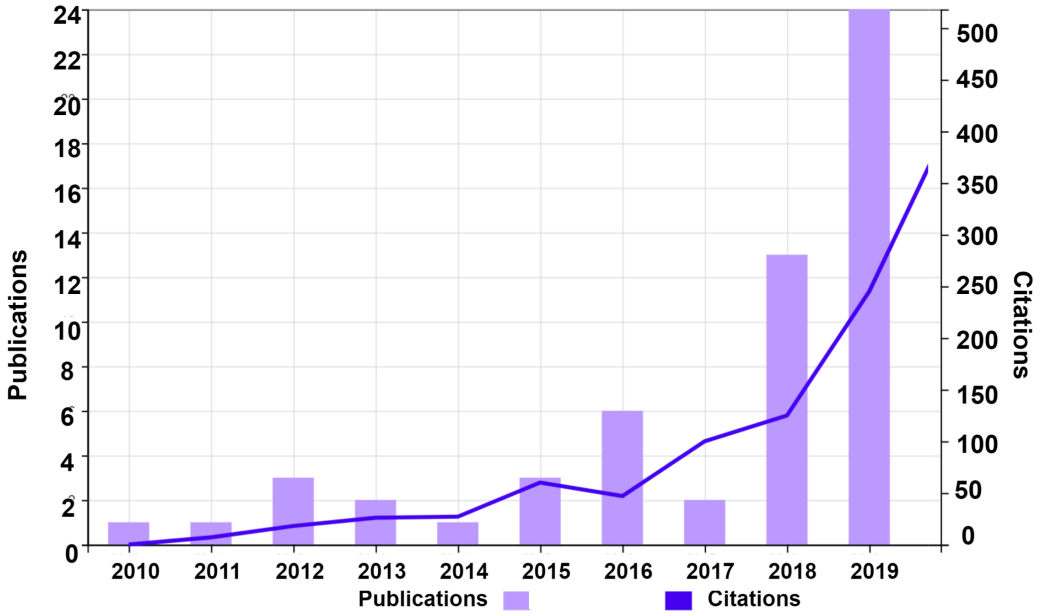

Figure 1.5: Publication and Citation statistics of studies on use of AI for Biomass Estimation (Source: Web of Science)

becoming more relevant to resolve AGB modelling problems. Advances in generative and graph neural networks have provided solutions to many complicated problems including those in RS of forest. The number of peer-reviewed scientific publications on application of AI for AGB modelling have steadily increased in the recent years (Figure 1.5). This shows an exceptional interest of the scientific community to capitalize on the potential of AI and overcome problems in development and deployment of AI based solutions for AGB mapping. This provides motivation to build custom AI models and modelling pipelines for delivering advanced solutions to the AGB modelling problem. There are some notable challenges for using AI in RS applications that requires efficient solutions. The most prominent among them are data scarcity, high computing power requirements and complexity. These challenges however provide excellent research avenues and encouragement to provide elegant solution to research problems

in using AI for forest AGB prediction.

1.3 Research Problems

Absence of structural information in satellite multispectral data

SMS remote sensing uses a passive mechanism and provides two-dimensional information of forest canopy. Despite the widespread application of SMS data, the data finds limited usage for AGB estimation due to restricted penetration capacity in forest canopy. Three-dimensional data such as lidar can capture forest structural features more precisely unlike two-dimensional SMS data that require empirical solutions for AGB estimation. Although, the extensive availability and lower cost of SMS data provides creates requirement for development of methods that can resolve the problems related to the usage of the data. SMS data derived canopy reflectance and vegetation indices are used to predict forest AGB using empirical models (regression or machine learning algorithms). Moreover, SMS data (e.g. Landsat-8, Sentinel-2, MODIS) are sensitive to vegetation density which relates to AGB. The problems identified in literature are with regards to the moderate accuracy of the empirical models that vary with the spectral variables and the local environment. There are observable research gaps that require rigorous inspection of SMS data to develop robust models using spectral variables. There are seasonal impacts causing spectral variation that affect model parameters and thus cause change in response of the model. Therefore, it is important to identify the spectral channels that remain stable and consistently correlated with the target AGB. Also, the radiometric specifications of different satellites may cause certain variation in response of identical spectral variables from different satellite data. To sum up, all properties associated with different SMS data requires examination that can make the empirical modelling strategy more reliable and

compensate for the lack of structural information in the data.

Data Saturation and Low Prediction Precision

Data saturation is the common problem encountered using SRS data such as MS and SAR for AGB prediction using empirical modelling strategies. The AGB prediction precision of the model is compromised due to data saturation problem with spectral variables. Also, the cloud cover or aerosol interference obscures MS remote sensing signal creating data unavailability or noisy data. SAR data are not affected by clouds and atmospheric aerosol and penetrate vegetation depending on the radar frequency. However, as per current scenario, open-source SAR data are only available in C-band, which is problematic due to high sensitivity to soil conditions, surface roughness and soil moisture in low vegetation coverage. Also, estimating AGB using its relationship with C-band SAR backscatter results to saturation at high biomass levels. This creates a requirement of additional source of information and extraction of target oriented features from the data. SRS has the advantage of delivering MT data for an area of interest. The role MT data as additional information source to deal with saturation needs to be investigated and quantified. Moreover, the features that are extracted from SMS or SAR data are required to be more sensitive to AGB at all levels. Therefore, there is a need to build mechanisms to manipulate these extracted features that can make them more sensitive and AGB oriented. Resolving these problems can lead to increase in the modelling accuracy and deliver precise AGB predictions from the model.

Lack of Robust Modelling Pipeline and Evaluation Strategy

There are multiple features that can be extracted from SRS data and multiple models that can be experimented to model forest AGB. The task of

determining suitable features and suitable models from a large collection is tedious and requires statistically robust strategy. There are a few conventional feature regularization and model selection techniques that can be used to deal with part of these issues. However, it is challenging to effectively link the feature extraction and modelling process with an automated framework to deal with the problem of systematic errors. Also, there are diverse algorithms based on different mechanisms such as bagging, boosting, neural nets, etc., which are required to be selected, tuned and parameterized in a robust manner. It is difficult to execute such tasks using conventional selection and regularization techniques. Manual process of understanding internal logic of the algorithm and its relation to AGB is time-consuming and resource-intensive. Thus, there is a need to replace manual analytics with automated algorithms for regulating elements of the modelling pipeline such as identifying relevant part of the data or features and choosing the type of algorithm to make the process more reliable. Fine-tuning the end-to-end machine learning process or the machine learning pipeline automatically can increase efficiency at scale. This can increase the standardization and consistency into the generation of AI models and provide a way to clearly define when an AI model is good enough to be deployed.

1.4 Objectives and Novel Aspects of the thesis

The overall goal of the thesis is to develop approaches and methods for solving the major issues related to the use of SRS data for AGB prediction such as assessment of feature response from SRS data , targeted feature generation and automation of AGB modelling pipelines. In particular, the specific objectives of thesis are defined as follows:

1. Develop a procedure for assessing the effect of temporal, spectral and

spatial properties of SMS data on spectral features for prediction of forest AGB and overcoming absence of structural information.

2. Develop a generative deep neural network for target-oriented feature generation and multi-source feature fusion to counter data saturation and improve AGB prediction precision.

3. Develop an automated end-to-end pipeline for robust feature generation, model ensembling and model ranking strategy for prediction of forest AGB.

The first objective requires deploying an adaptive regularization procedure for performing response assessment of spectral features to the target parameter (AGB). This provides crucial information on response of data and saturation of spectral features that are preliminary to extraction and engineering of target-oriented features for AGB modelling (second objective). The qualitative and quantitative results provides details regarding the requirements to refine features and append source of information for improving the model performance. The second objective aims at the development of a deep generative neural network to engineer target oriented features and feature fusion benefits from the results obtained from the first objective. The response assessment facilitates better implementation of the network architecture and fusion mechanism. The third objective attempts to automate multiple processes such as feature extraction, hyperparameter optimization, model selection and ensembling that are involved in modelling AGB. This facilitates the modelling process in overcoming drawbacks observed for previous two objectives with respect to feature generation and model development to deliver an integrated framework for modelling. The developed method aims to deliver a more robust ensemble model as compared to single base models developed for the previous two objectives. The thesis collectively aims at providing contributions for solving the AGB pre-

diction problem from an "Information Engineering and Computer Science" perspective. The constituent contributions of the thesis are organized and addressed from the standpoint of predictive analytics and machine learning applications. Therefore, the thesis follows a red thread limiting discussions regarding the advancements provided in the contributions with respect to statistical tools and data processing pipelines.

The novel aspects that are incorporated to achieve the objectives of the thesis are briefly stated as follows:

1. *Novel aspects in the procedure developed for effect assessment of various properties (temporal, spatial and spectral) of SMS data on spectral features considered for modelling AGB*

   In order to achieve the first objective, a generalized linear model is deployed with adaptive lasso regularization for evaluating the response of spectral features with respect to different capacities of SMS data. The procedure is implemented for its 'oracle' property. The adaptive lasso algorithm has properties to identify the optimal subset of variables as if the true underlying model were given in advance. The response of spectral features extracted from SMS data are analyzed and compared for the spatial, spectral and temporal capacities of MS data. The adaptive lasso provides consistency in this regard as ordinary lasso often includes too many features and the true response is usually provided by a subset of these features. A secondary stage regularization of adaptive lasso controls the bias and leads to consistent selection of features (oracle property). Thus, a systematic and efficient framework was put in place for detailed investigation of each MS data capacity with regard to spectral features that can be followed for development of effective tools and techniques to model AGB using SMS data.

2. *Novel dynamic architecture developed for the extraction and fusion of target-oriented features from MS and SAR data*

   In order to achieve the second objective, a novel, dynamic and generative neural network architecture was proposed for extracting target oriented features to model forest AGB. The proposed architecture is an attempt to combine generative model with a variational autoencoder neural network. The main difference between an ordinary and the proposed variational autoencoder is that conventionally the latent space is generated from a prior Gaussian distribution. Whereas, in this contribution fine blend of inference and generative parameters are used to impose a conditional distrubtion of latent representation based on the target AGB balues . The combination of three loss functions enables the generative features to represent an ordered variation in the target AGB parameter. These loss function enables the component networks of the proposed model to regularize each other and deliver more accurate AGB predictions. The proposed architecture is dynamic in the sense that it shrinks and expands depending on the input dimension to retain the relative ratio with number of model parameters. Moreover, it provides a platform for seamless fusion of multi-senor remote sensing data.

3. *Novel approach developed based on the concept of automated machine learning for deploying end-to-end pipeline for AGB prediction*

   In order to achieve the third objective, the concept of automated machine learning is employed to automate all sections of the modelling pipeline for prediction of forest AGB. An end-to-end pipeline is developed for automatic feature extraction using a hyperparameter optimization procedure. Particularly, a Tree-Parzen estimator (TPE) is used to optimize the features in form of empirical equations and fit the

best spectral band to compose the feature. This kind of automation produces target specific features and rejects the notion of predefined features such as standard spectral indices. Moreover, a meta-learning algorithm is implemented for automating a large selection of diverse regression algorithms to model AGB. Particularly, a super-Learner (SL) algorithm optimizes the machine learning hyperparameters and performs comparative ranking to point out the best base models. The algorithm also enables to analyze combinations of different base models to deduce an optimal combination suitable to best fit the data and deliver best AGB prediction. This novel procedure based on combination of TPE and SL algorithm enables deployment of an automated end-to-end AGB prediction pipeline.

1.5 Structure of the Thesis

The current chapter (Chapter 1) is a brief overview of the background, motivation and research problems for the work presented in the thesis. The chapter also briefly discusses the objectives and novel aspects of the thesis. The rest of the thesis is organized as five additional chapters.

Chapter 2 presents the state-of-the-art for AGB prediction using SRS data. The analysis of the state-of-the-art of both active (SAR), passive (MS) and combined (MS and SAR) is covered.

Chapter 3 provides details regarding the methodology proposed to perform an extensive analysis of SMS data with diverse specifications and the effects of each property of the MS data for the prediction of forest AGB.

Chapter 4 provides details regarding the proposed framework and architecture based on a generative neural network based model for AGB-oriented feature generation and fusion. Additionally, it provides a quantitative assessment of generated features and a comparison with contemporary fea-

tures and algorithms to model AGB.

Chapter 5 illustrates the hyperparameter optimization and meta-learning algorithms used for achieving the third objective of the thesis. The proposed method further discusses the process of combining these algorithms to establish an end-to-end pipeline for AGB prediction using SMS data.

Chapter 6 presents the collective conclusions drawn from the results of the three main contributions of the thesis and provides a scope regarding the possible future work.

Chapter 2

State of the Art

The modelling of forest AGB using remote sensing data can be performed based on individual tree based methods and area based methods. Individual tree based methods operate on individual tree statistics extracted from a very high resolution or three-dimensional remote sensing data. The cost of data acquisition for implementing individual tree based techniques is higher but the extracted statistical attributes carry more detailed and accurate information. Whereas, the area based methods operate on statistics extracted over an area or a plot consisting multiple trees that are preferable for coarse resolution remote sensing data. Area based methods are usually recommended for SRS data and are efficient in terms of scale of estimation. Area based methods require a considerably larger number of field samples for model calibration in order to achieve acceptable modelling accuracy. This thesis concentrates on using SRS data and therefore, on developing area based methods for modelling AGB. Satellite based MS (in passive remote sensing) and SAR (in active remote sensing) data are comparatively more accessible than other modes and thus used for the studies conducted in this thesis. The following sub-sections discusses the state-of-the-art approaches using MS, SAR and combined MS-SAR data for AGB prediction.

2.1 State-of-the-art for AGB prediction using SMS data

Spectral indices extracted from SMS data are sensitive to the photosynthetic properties of forest vegetation. These spectral indices are indirectly related to forest AGB by means of an empirical relationship between foliage and total AGB. Thus, the signals captured by MS sensors are sensitive to variations in canopy structure and several methods use this relationship to model AGB [6, 9, 10]. A recent surge in the capacity of MS satellites has increased availability of MS data and led to the development of novel methods for AGB prediction [19, 69, 37]. Studies indicated that a few spectral predictors were observed to saturate at higher AGB values typically greater than 150 Mg/ha [38]. The variation of saturation levels are also dependent on the forest types. Landsat data evaluated for different forest types observed saturation at 123 Mg/ha for broadleaf forest, 143 Mg/ha for Fir and 152 Mg/ha for mixed forest. Since then, the capability of spectral predictors has been gradually enhanced with new freely available MS products such as Sentinel-2 MSI (at 10 m against 30 m of Landsat-8). The three Sentinel-2 red-edge bands are potentially effective in increasing sensitivity of spectral features to map AGB with a good accuracy. The assessment of Sentinel-2 data has accounted for improved AGB prediction accuracy as compared to Landsat data in multiple studies [113, 78]. The RMSE of predicted AGB for Sentinel-2 ranged from 65 - 95 Mg/ha, which is significantly better compared to Landsat with RMSE = 85-100 Mg/ha [113]. Similar studies have attempted to assess the saturation limits of different MS satellite data for estimating forest AGB [40, 25, 41]. These studies collectively indicate better results achieved with recent MS data with enhanced spectral and radiometric properties. However, the assessment of potential for Sentinel-2 and other recent MS satellite missions may

require a more thorough investigation. Also, the implementation of specific advanced modelling algorithms can further minimize the retained levels of saturation for predicting AGB.

There are several algorithms that have been tested for modelling forest AGB using SMS data. Most popular among them are multiple linear regression, Random Forest (RF), Support Vector Machine (SVM) and Artificial Neural Network (ANN) [32, 37]. The multiple linear regression models are used in identifying appropriate spectral variables and analysis of various properties of MS data to predict AGB [32, 114, 116]. Multiple linear regression also allows to determine the relative contribution of each of the independent variables to the total explained variance of the model. Alternately, machine learning algorithms such as RF, SVM and ANN are more flexible and create complex models for capturing non-linearity in the AGB. These models are non-structured and data-driven, thus deemed more reliable for predicting AGB using MS data. RF, SVM and ANN models were compared in a recent study [37] to model AGB using MODIS data. RF model delivered accurate AGB predictions (R^2 = 0.938 and RMSE = 19.88%) due to less sensitivity to noise in the training samples yet prone to overfitting. SVM models produced better generalization but saturated for large AGB values (R^2 = 0.629 and RMSE = 40.56%). The ANN model performed least accurately (R^2 = 0.334 and RMSE = 53.01%) among the three models as the network quickly plunges into a local minimum. In spite of achieving good performance, RF models are less reliable for unbalanced datasets and cause overfitting. eXtreme Gradient Boosting Machine (XGBoost) is a better replacement for RF in the category of decision tree models as its builds one tree at a time by considering gradient of the data. The study in [25] showed that XGBoost (R^2 = 0.66 and RMSE = 42.29%) model outperformed RF model (R^2 = 0.21 and RMSE = 60.68%) on Landsat-8 data for AGB modelling. However, gradient boost-

ing is sensitive to outliers and unable to extrapolate target values beyond limits for regression problems. In summary, all these individual algorithms have certain drawbacks for accurately modelling AGB depending on the specifications of training data, underlying non-linearity or distribution of samples. To overcome the drawbacks of individual algorithms, the most recent methods have been focused on developing stacked ensemble models for accurate AGB modelling [161]. The results proved that stacked ensemble models reduced the bias and achieved better accuracy in estimating AGB as compared to other individual base models.

2.2 State-of-the-art for AGB prediction using satellite SAR data

SAR backscatter is sensitive to forest AGB and its correlation with AGB has been studied for a few decades theoretically and experimentally [42, 43, 44]. SAR based AGB measurements are influenced by the structure of vegetation, the surface topography and the environmental conditions. The sensitivity of SAR signal to forest AGB changes depending on the wavelength and geometry of radar measurements. Space-based SAR data acquisition is currently limited to X-, C- and L-band frequencies for observation of forest ecosystem. Among the past and current SAR missions, ALOS-1 (L-band, non-operational) and Sentinel-1 (C-band, operational) provide free and open-source data. Few studies that used Sentinel-1 data reported varying degree of AGB prediction accuracy ($R^2 = 0.3$ - 0.7) [45, 46, 47, 48]. These results indicated that singular polarization components (VV or VH) saturated in accurately capturing the underlying relationship with AGB. A better accuracy was achieved with Dual Polarization SAR Vegetation Index (DPSVI) - a combination of polarization components [45] as compared to singular polarization components of Sentinel-1 data. The addition

of grey-level co-occurence matrix (GLCM) textural features and Principal components to polarization components and their derivatives (such as DPSVI) delivered a significant improvement and the best accuracy among the other cited approaches [46].

The future SAR missions such as NISAR (L-band) and BIOMASS (P-band) will operate at frequencies that can penetrate deep into forest canopy. Both these missions will provide free and open-source data that presents immense opportunities for modelling forest AGB. There are several AGB modelling studies using ALOS-2 PALSAR-2 (L-band) commercial data that can serve as precursor for using NISAR data [49, 50, 51, 52]. The models developed in these studies using PALSAR-2 data performed better as compared to the Sentinel-1 based models. The model accuracy varied depending on the extracted features and modelling algorithm used for the prediction of AGB. A recent study [51] that used only PALSAR-2 backscatter coefficients (HV, HH) to predict AGB reported low model accuracy and rapid model saturation (RMSE = 89.9 -97.9 Mg/ha, Saturation point : 160 Mg/ha) on single-time images. Although, multitemporal PALSAR-2 images produced better prediction accuracy with an RMSE = 62.8 - 66.4 Mg/ha and raised the saturation point to 240 - 280 Mg/ha. However, the best results were achieved for the studies that used terrain factors, decomposition parameters and covariance matrix elements in addition to backscatter coefficients. Particularly in [49], the model provided the best results (adjusted $R^2 = 0.90$, RMSE = 14.24 t/ha) with backscattering coefficients, polarization decomposition variables, and terrain factors as predictors using a RF model. The achievement of such accurate results were attributed to the synergy among L-band predictors. P-band SAR that penetrates through the canopy to interact with the ground presents possibilities for the development of novel AGB modelling approaches. Currently, there are no operational P-band spaceborne SAR missions but a few

studies were conducted using airborne acquired data [53, 54, 55]. These studies are focused on tomographic processing of multiple P-band SAR images (TomoSAR) that extracts information on vertical backscatter distribution. The studies conducted with airborne P-band data acquired as part of the BioSAR campaign [53, 54] produced high AGB prediction accuracy with TomoSAR approach with a power law AGB model (RMSE = 27 - 33 t/ha) and linear/logarithmic models (RMSE = 20 t/ha). Similar results were also obtained using P-band data from TropiSAR campaign [55] that collectively present a very high potential for accurately mapping global biomass with future spaceborne P-band SAR missions using TomoSAR techniques.

2.3 State-of-the-art for AGB prediction using combined MS and SAR data

A vast majority of studies seek to derive information from a single type of spaceborne data at a time to model forest AGB. A single type of spaceborne data refers to a specific window of the electromagnetic spectrum or a single acquisition method. Single data type based AGB modelling is more common because co-located and nearly simultaneous spaceborne multi-mode data acquisitions are often not available. The recent developments in small satellite technology created a surge in the number of MS and SAR satellite constellation (e.g. PlanetScope, Iceye). This facilitates the availability of multi-mode coherent data both in terms of time and position. MS and SAR frequencies contains independent information for modelling forest AGB and multiple studies indicate that their combination improved the AGB estimates by reducing uncertainty of single sensor data [21, 56, 57, 58, 25, 59, 60]. Sentinel-1 and 2 are the most frequently used combination of SAR and MS data for prediction of forest

AGB [21, 56, 57, 58]. They have coherent revisit time and spatial resolution that render them more suitable for synergistic use for modelling AGB. A fusion model developed with a simple predictor based on the combination of backscatter intensity ratio from Sentinel-1 and NDVI from Sentinel-2 yielded superior AGB prediction results (R^2 = 0.86, RMSE = 0.109) as compared to exclusive Sentinel-2 data (R^2 = 0.72, RMSE = 0.158) [21]. A few other studies that used machine learning approaches with the identical predictors yielded similar results for SVM and RF models [56, 57]. The RF model elevated the AGB saturation point to 200 t/ha with an RMSE of 60 t/ha. However, the best performance was secured with SVM with a linear kernel that obtained highest correlation (r = 0.85) for the Sentinel-1 (VV+VH) and Sentinel-2 (NDVI) fusion model. Apart from Sentinel-1 and 2, there are few studies that considered combination of other sensors for modelling AGB [25, 59, 60]. Landsat-8 (texture and vegetation indices) and Sentinel-1 (texture and backscatter) were combined to model AGB using linear, RF and XGBoost machine learning algorithms [25]. A better model agreement score was achieved for Landsat-8 models (R^2 = 0.21 - 0.66) as compared to Sentinel-1 models (R^2 = 0.03 - 0.38) . Overall, the best agreement score was achieved for a combination model (Landsat-8 + Sentinel-1) using the XGBoost algorithm (R^2 = 0.75). Among the commercial and restrained sources, few studies combined ALOS-2 data with Sentinel-2 and Worldview-2 with ALOS PALSAR for AGB estimation [59, 60]. The best model performance was obtained for the XGBoost algorithm on ALOS-2 and Sentinel-2 (R^2 = 0.8, RMSE = 28.1) whereas multilinear regression obtained the best performance on Worldview-2 and ALOS PALSAR (R^2 = 0.89, RMSE = 24.4). A comprehensive analysis of all these studies suggests that combined MS-SAR data reduces uncertainty of single data and significantly improves the AGB modelling accuracy especially using advanced machine learning algorithms such as XGBoost.

Chapter 3

Multitemporal Multispectral Satellite Remote Sensing Data for AGB Prediction

3.1 Introduction

SMS remote sensing is frequently noted as a potential alternative for replacing tedious field sampling based AGB estimation over large areas due to its coverage, repetitiveness and cost-effectiveness [61, 62]. MS data are available with different spectral, spatial and temporal specifications. Quickbird, IKONOS, WorldView, GeoEye, PlanetScope etc. are some of the spaceborne missions that provide high-spatial resolution ($< 5m$) MS data. These data enhance the vegetation assessment capability with greater spatial details of the forest.The medium and coarse resolution products from missions such as Sentinel-2, Landsat-8 and MODIS provide low spatial resolution data but its enhanced radiometric and spectral capacities are critical for modelling forest AGB.

Part of the chapter appears in:

P. Naik, M. Dalponte, and L. Bruzzone, "Prediction of Forest Aboveground Biomass Using Multitemporal Multispectral Remote Sensing Data," Remote Sensing, vol. 13, no. 7. MDPI AG, p. 1282, Mar. 27, 2021. DOI: 10.3390/rs13071282

The effect of different characteristics (spatial, spectral, temporal and radiometric) for modelling forest AGB requires careful evaluation. The different spectral bands of the MS data have varying degree of correlation with the range of AGB values [63]. Moreover, higher number of spectral bands i.e. greater spectral information can reduce the data saturation effects for AGB prediction. Data saturation was previously studied for MS data from Landsat-8 and MODIS but less studied for recent MS data from Sentinel-2, PlanetScope and other small-satellites [64, 65, 66]. The temporal information can also potentially reduce the problem of data saturation that restricts the prediction of large AGB values. The response of spectral indices derived from MS data vary with seasons. Thus Multitemporal (MT) data can be an additional source of information for modelling forest AGB [9, 67]. A high temporal resolution can also increase the probability of capturing cloud-free images of the forest. Also, the spatial details of the forest are dependent upon the spatial resolution of the MS images. High spatial resolution MS image pixels cover smaller areas of the forest with greater spatial details. However, high spatial resolution MS data need large data storage space and more processing time. These factors influence the application of high spatial resolution images for AGB estimation over broad areas. Overall, it is important to investigate all capabilities of the recent spaceborne MS sensors for accurate prediction of AGB. Such investigative studies can serve as a feedback mechanism for the ongoing missions for improving the quality of the MS data.

There are various image processing techniques in literature that have been applied for AGB estimation with MS remote sensing data [68, 15, 69, 70]. The most frequently used approaches are k-nearest neighbor (kNN), multiple linear regression, machine learning regression and neural networks [71, 72, 73, 25]. These methods use MS remote sensing data to determine tree canopy parameters (e.g. crown diameter) or develop canopy reflectance

models. The methods for AGB prediction based on high resolution MS data involve modelling of tree parameters or forest canopy structures. The study in [74] estimated tree crown size from Quickbird images and [75] used IKONOS for estimating parameters such as crown area and DBH for AGB modelling. The approaches for medium and coarse spatial resolution are relatively different as compared to that of high spatial resolution MS data. For example [76] used canopy reflectance model inversion to estimate forest AGB from SPOT 5 data. Recently, multiple studies used Sentinel-2 data with linear regression and other machine learning algorithms (RF, XGBoost etc.) for modelling forest AGB with derived vegetation indices [77, 78, 79]. Several studies have also used Landsat TM using k-nearest neighbor method and direct radiometric relationships to map AGB [80]. A few other studies have used coarser data such as MODIS using similar methods and vegetation indices for modelling forest AGB [81, 82, 83]. Overall, these approaches involving SMS data are found suitable with certain limitations such as data saturation and variance in response of features and requires proper quantification.

The objective of this contribution is to quantify and analyze the response of spectral features extracted from different MS satellites with different properties (temporal, spectral and spatial) for the prediction of forest AGB. Spectral features from MT and single-time data are considered to evaluate quantitative improvements in model predictions delivered by multi-seasonal data as compared to individual season data. The regularized and selected spectral features are split into three groups depending on the considered of bands. The response of key spectral features are analyzed at three different spatial resolutions (10, 5 and 3 m) to observe their correlation at various spatial scales. Experiments are conducted for two different study areas with different forest compositions and densities.

3.2 Study Area and Datasets

3.2.1 Study Area

The study area includes two sites - Lavarone and Pellizzano, both located in the Southern Italian Alps. The forest in Lavarone is an uneven aged forest located at an altitude ranging from 1200 to 1600 m with species such as Norway spruce (*Picea abies* (L.) Karst.) and silver fir (*Abies alba* Mill.) accounting for 80% of the tree population (in terms of the number of trees). European beech (*Fagus sylvatica* L.), European larch (*Larix decidua* Mill.), and Scots pine (*Pinus sylvestris* L.) accounts for the remaining 20% of the population. The forest of Pellizzano is spread over an altitude from 900 to 2200 m and has a dominant presence of Norway spruce and subdominant presence of some coniferous species (e.g., *Abies alba* Mill., *Larix decidua* Mill., *Pinus cembra* L., *Pinus sylvestris* L., and *Pinus nigra* J., Arnold) and broadleaves species (*Populus tremula* L., *Betula* spp.). The forest is multilayered, more mixed, and structurally complex at lower altitudes and sparse at higher altitudes. The study was conducted in two areas to consider different species and the density of the forest. The geographical locations of both the study areas are shown using detailed maps in Figure 3.1.

3.2.2 Field Data

The field data for the Lavarone area were collected on multiple days from summer of 2016 for 41 circular plots of 30 m diameter and from summer of 2014 for 47 circular plots of 30 m diameter for the Pellizzano area. The plot coordinates were recorded using a survey grade Global Positioning System unit. The field plots were distributed randomly across the forest area to be statistically significant. The positions were measured with a Laser criterion 400, while the DBH was measured with a caliper in two orthogonal

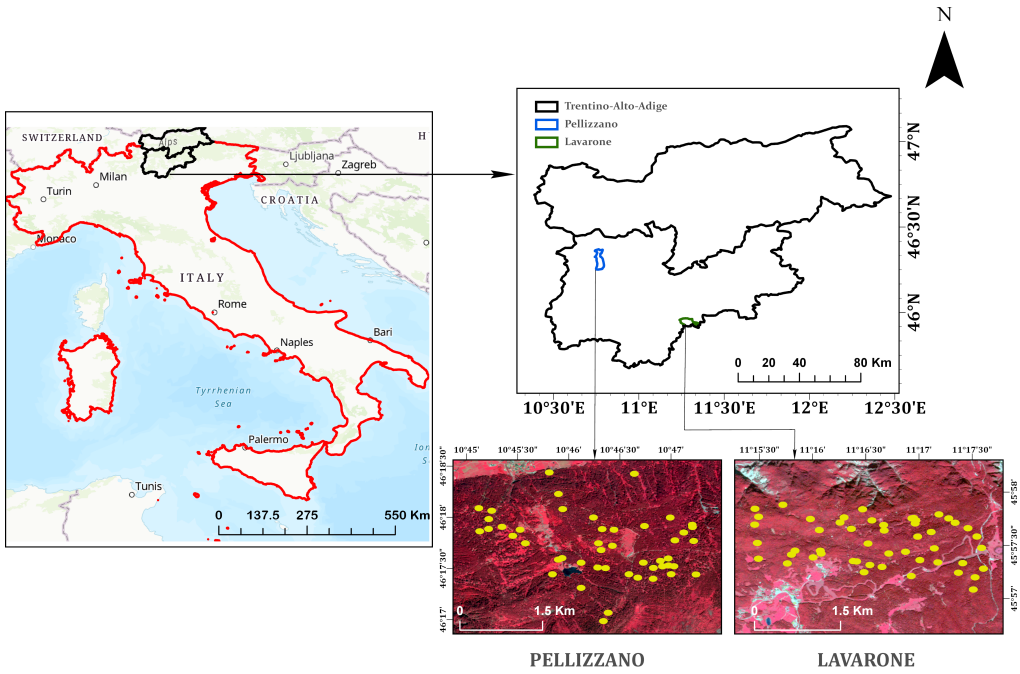

Figure 3.1: Map of the selected study areas - Pellizzano and Lavarone in Trentino Alto-Adige, Italy (Satellite images—Sentinel2)

directions and the average of the two measurements was considered as DBH value. The height was measured with a vertex hypsometer and for the trees where the height was difficult to measure, it was estimated using species specific allometric equations [158]. The DBH threshold was kept at 3 cm (i.e., all trees above 3 cm were measured). The dead trees were included in the plot measurements but their presence was scarce. The AGB of each tree was computed using the allometric equations stated in [84] (see Appendix) and then aggregated for each plot to obtain plot level AGB per hectare. The field plot AGB of the Lavarone area ranged from 69.42 to 412.34 Mg ha-1 and the of the Pellizzano area ranged from 0.9 to 545.3 Mg ha-1. The mean tree height, mean tree DBH, and number of trees per plot for both the study areas are stated in Table 3.1.

Table 3.1: Field Estimated Mean tree height (m), mean tree DBH (cm), and number of trees per plot for Lavarone and Pellizzano

		Lavarone	Pellizzano
Mean tree height per plot (m)	Minimum	7.43	5.94
	Maximum	27.29	37.9
	Mean	16.85	20.58
Mean tree DBH per plot (cm)	Minimum	10.71	8.78
	Maximum	37.83	57.17
	Mean	21.24	31.68
Number of trees	Minimum	17	2
	Maximum	192	147
	Mean	76.1	40.15

3.2.3 Remote Sensing Data

The Sentinel-2 data are acquired by twin satellites under ESA's Copernicus program with a revisit capacity of 5 days. Sentinel-2 images are characterized by 13 spectral bands with four bands at 10 m, six bands at 20 m, and three bands at 60 m spatial resolution. In this study, 4 spectral bands at 10 m and 6 spectral bands at 20 m resampled to 10 m were used in the Blue, Green, Red, Red-edge, Near-Infrared (NIR) and Short-wave Infrared (SWIR) region.

The RapidEye data are acquired by a constellation of five satellites operated by the Planet Labs Inc., with a MS push-broom imager capturing five spectral bands: Red, Green, Blue, NIR, and Red-Edge. The RapidEye satellites are identically calibrated so that the images acquired by the five satellites are indistinguishable from each other. All the acquired RapidEye data have a view angle smaller than 20 degrees. The ground sampling distance (nadir) is 6.5 m, and the orthorectified pixel size is of 5 m. The constellation of satellites acquires data with a daily revisit capacity (off-nadir) and 5.5 days at nadir.

The Dove data are acquired by a constellation of 120 Nano-satellites operated by Planet Labs Inc. with a daily revisit capacity. The satellites are deployed at the International space station orbit at an inclination of 52 degrees and Sun synchronous orbit with an inclination of 98 degrees. The MS data are acquired using a Bayer-masked Charge-Coupled Device camera with ground sampling distance (nadir) of 2.7–3.2 m in the ISS orbit and 3.7–4.9 m in the SS orbit. The orthorectified pixel size of the acquired images is 3 m with 4 spectral bands: Red, Green, Blue, and NIR.

The dates of acquisition of MS data considered for the study are stated in Table 3.2. The time gap between the field sampling and remotely sensed data exists for maintaining consistency of temporal acquisitions of different satellite data. The field estimated annual increment in the DBH was 2.8% [85]. Therefore, the increment of AGB sampled forest plots during the time gap can be considered as negligible and hence, admissible for the conducted experiments.

Table 3.2: Acquisition dates (YYYY-MM-DD) of the multi-scale SMS data used for the study (Sentinel-2 at 10m, RapidEye at 5m and Dove at 3m)

Seasons	Dates of Acquisition (Lavarone)			Dates of Acquisition (Pellizzano)		
	Sentinel-2	RapidEye	Dove	Sentinel-2	RapidEye	Dove
Summer	2016-07-18	2016-07-29	2016-07-11	2016-07-18	2016-07-17	2016-08-22
Autumn	2016-10-16	2016-10-16	2016-10-20	2016-10-16	2016-10-21	2016-10-07
Winter	2017-01-07	2017-01-16	2017-01-23	2017-01-24	2017-01-16	2017-01-24
Spring	2017-04-07	2017-04-30	2017-04-23	2017-04-14	2017-04-22	2017-04-10

3.3 Methodology

The methodological flowchart for this contribution of the thesis is shown in Figure 3.2. In the following sections, each step of the methodology is described in detail.

The general idea supporting the development of this methodology is based on the fact that spectral features are mathematical transformations of original spectral reflectance that can be used for interpreting forest AGB. The main interpretations are derived from the divergent responses produced by spectral features that are capable of explaining the physiological properties of forest vegetation. A high near-infrared reflectance can be observed due to scattering by leaf mesophyll cells and a low red reflectance due to absorption by chlorophyll pigments. Thus, the adopted methodology focuses on developing models that leverage the difference between the strong absorption of red electromagnetic radiation and the strong scatter of near infrared radiation.

3.3.1 Data Pre-processing

Level-1C data for Sentinel-2 data were acquired from the Copernicus open hub platform and converted to Level-2A (Bottom-of-atmosphere) that were atmospheric and terrain corrected using Sen2cor plugin. The spectral bands at 20 m spatial resolution were resampled at 10 m using nearest neighbor sampling to adjust the pixel size to have a uniform band stack and simplified computation. Level-3A (orthorectified, geometrically corrected, and radiometrically calibrated) RapidEye and Dove images were gathered from the data access portal of Planet Labs (`https://www.planet.com/explorer/`). The individual data bands for each dataset were mosaicked and stacked for producing a single data frame for both satellite images for all seasons.

3.3.2 Feature Extraction

Two set of spectral features were considered for each MS dataset: the original band reflectances and the derived vegetation indices. The band reflectances of Sentinel-2 from B2 to B8 represent Blue, Green, Red, Vegetation red edge – 1, 2, 3, and NIR, respectively; B8A represents narrow NIR, and B11 to B12 represent SWIR-1, 2 channels respectively. The spectral reflectances of RapidEye from B1 to B5 represent Blue, Green, Red, Red edge and NIR channels respectively. The spectral reflectances of Dove from B1 to B4 represent Blue, Green, Red, and NIR channels respectively. Numerous vegetation indices have been defined for specific applications in the literature [90, 91, 92]. To limit the number of predictor features from the large number of existing vegetation indices, 25 robust indices identified in literature for forestry applications were selected [86, 87, 26]. These 25 indices are stated in Table 3.3 along with their respective equations.

All these features were extracted for each circular ground plot of 30 m diameter. In order to match the ground sample plots with remote sensing data at different spatial resolutions, mean aggregation of the pixels (at 3, 5, and 10 m according to the satellite considered) belonging to each plot was performed. The "gBuffer" function of the "rgeos" library in R was particularly used for the mean aggregation operation. A round cap style with 15 m radial size was used for the inclusion of pixels and the overlapping pixels with the plot size on the ground were considered.

3.3.3 Adapative Lasso and Generalized Linear Modelling

Feature regularization and Model selection are important to statistically exclude features consequently reducing model complexity and overfitting. A total of 25 vegetation indices and available band reflectances were considered as predictor features for modelling AGB. The objectives of this

Table 3.3: Vegetation indices extracted from SMS data used as features for modelling and their equations

Vegetation Indices	Equations
Canopy Chlorophyll Content Index	$\text{CCCI} = \frac{\frac{\text{NIR} - \text{RedEdge}}{\text{NIR} + \text{RedEdge}}}{\frac{\text{NIR}-\text{R}}{\text{NIR}+\text{R}}}$
Chlorophyll Index Red Edge	$\text{CIRE} = \frac{\text{NIR}}{\text{RedEdge}} - 1$
Chlorophyll Vegetation Index	$\text{CVI} = \text{NIR}\frac{\text{R}}{\text{G}^2}$
Green Atmospherically Resistant Vegetation Index	$\text{GARI} = \frac{\text{NIR}-(G-(B-R))}{\text{NIR}-(G+(B-R))}$
Green Leaf Index	$\text{GLI} = \frac{2G-R-B}{2G+R+B}$
Log Ratio	$\text{logR} = \log\frac{NIR}{R}$
Normalized Difference Vegetation Index	$\text{NDVI} = \frac{\text{NIR}-\text{R}}{\text{NIR}+\text{R}}$
Normalized Burn Ratio	$\text{NBR} = \frac{\text{NIR}-\text{SWIR}}{\text{NIR}+\text{SWIR}}$
Green Blue NDVI	$\text{GBNDVI} = \frac{\text{NIR}-(\text{G}+\text{B})}{\text{NIR}+\text{G}+\text{B}}$
Green Red NDVI	$\text{GRNDVI} = \frac{\text{NIR}-(\text{G}+\text{R})}{\text{NIR}+\text{G}+\text{R}}$
Red Blue NDVI	$\text{RBNDVI} = \frac{\text{NIR}-(\text{R}+\text{B})}{\text{NIR}+\text{R}+\text{B}}$
Green NDVI	$\text{GNDVI} = \frac{\text{NIR}-\text{G}}{\text{NIR}+\text{G}}$
Red Edge NDVI	$\text{NDVIre} = \frac{\text{NIR} - \text{RedEdge}}{\text{NIR} + \text{RedEdge}}$
Pan NDVI	$\text{PNDVI} = \frac{NIR-(R+G+B)}{NIR+R+G+B}$
Visible Index Green	$\text{VIgreen} = \frac{\text{G}-\text{R}}{\text{G}+\text{R}}$
Norm of X (X = R, G, NIR)	$\text{NormX} = \frac{X}{\sum X}$
Blue-Wide Dynamic Range Vegetation Index	$\text{BWDRVI} = \frac{0.1*\text{NIR}-\text{B}}{0.1*\text{NIR}+\text{B}}$
Chlrophyll Index Green	$\text{CIgreen} = \frac{\text{NIR}}{\text{G}} - 1$
Green Difference Vegetation Index	$\text{GDVI} = \text{NIR} - \text{G}$
Blue Normalized Vegetation Index	$\text{BNDVI} = \frac{NIR-B}{NIR+B}$
Redness Index	$\text{RI} = \frac{R-G}{R+G}$
Vegetation Index Number	$\text{VIN} = \frac{\text{NIR}}{\text{R}}$
Modified Simple Ratio	$\text{MSR} = \frac{\frac{\text{NIR}}{R}}{\sqrt{\frac{\text{NIR}}{\text{R}}+1}}$
Specific Leaf Area Vegetation Index	$\text{SLAVI} = \frac{\text{NIR}}{R+\text{ SWIR}}$

R = Red; G = Green; B = Blue; NIR = Near Infrared; RedEdge = Red Edge; SWIR = Short-wave infrared; NDVI = Normalized Difference Vegetation Index.

contribution requires the method to produce an unbiased solution. The adaptive lasso technique is characterized with "oracle" properties so that the features are selected as if the true underlying model are given in advance.

The adaptive lasso estimator $\hat{\beta}^{LS}$ [88] can shrink some coefficients to exactly zero performing selection of features along with regularization. It is an evolution of the lasso algorithm with an adaptive weight vector ω_j that adjusts the penalty differently for each coefficient. The objective function of adaptive lasso can be stated as equation (3.1) and the adaptive weight vector can be stated as equation (3.2):

$$\hat{\beta}^{LS} = \underset{\beta \in R^p}{\arg\min} \sum_{i=1}^{n} \left(Y_i - X_i^T \beta\right)^2 + \lambda \sum_{j=1}^{p} \omega_j \left|\beta_j\right| \tag{3.1}$$

$$\omega_j = \frac{1}{\left(\left|\hat{\beta}_j^{ini}\right|\right)^{\gamma}} \tag{3.2}$$

Here, $\lambda > 0$ is the tuning parameter (penalty parameter) and $\hat{\beta}_j^{ini}$ is the initial estimate of the coefficients obtained by ridge regression and γ represents a positive constant for adjustment of the adaptive weights vector. The term $(Y_i - X_i^T)$ represents a random sample with identical independent distribution associated with regression coefficients β. Thus, the adaptive lasso penalizes coefficients with lower initial estimates to a greater proportion.

The process of implementing generalized linear model with adaptive lasso regularization is initiated with ridge regression used for computing the best ridge coefficients at a minimum value of $'\lambda'$ parameter. The penalty factor is calculated using the best ridge coefficients. The generalized linear model is implemented with lasso penalty accounting for the computed penalty factor. The penalty factor was used to tune the $'\lambda'$ parameter for a differential shrinkage. The penalty factor can be '0' for some predictor

variables with the lasso penalty that implies no shrinkage and inclusion of that particular predictor variable in the model. The adaptive lasso uses lasso penalty to push the coefficient of predictor variables to absolute '0' which improves the model by selecting the most consistent variables and automates the process of variable selection.

A generalized linear modelling [89] approach was used for associating selected variables with the target AGB values. The modeling algorithm uses cyclical coordinate descent to optimize the objective function over each parameter successively for convergence. Moreover, the generalized linear model uses a "gamma" link function (power = 0.5) to map the non-linear relationship of dependent variable to a linear relation. To achieve the objective of this contribution, six metrics were computed for assessing the precision, the agreement and the overfitting of the models. The cross validation metrics, their equations and the related assessment aspect are stated in Table 2.4. A Leave-one-out Cross-validation (LOOCV) cross validation approach was used for the modelling process to determine stable parameter and assess model overfitting. The R-packages – 'tidyverse', 'magrittr', 'glmnet', 'pROC', 'mgcv', and 'ggplot2' were used for implementing the entire process of regularization, selection, modelling, validation and visualization.

3.3.4 Design of Experiments

The designed experiments were aimed at analyzing the effect of temporal, spectral, and spatial capacities of MS satellite data for the prediction of AGB. The temporal analysis was carried out using the proposed approach on feature subset of common spectral channels (FS2) of the three sensors i.e., features computed using only using R, G, B and NIR bands of each satellite data. The (FS2) subset was considered for temporal analysis to remove the effects of additional spectral bands. The process was imple-

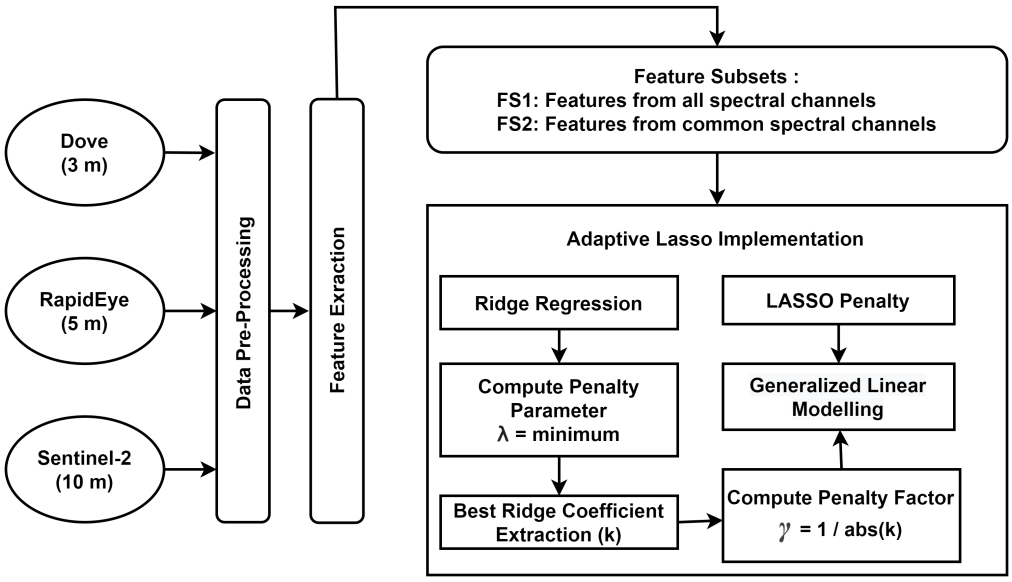

Figure 3.2: Methodology Flowchart describing steps for adaptive regularization and modelling using MS satellite data

mented for each single-time image to produce one model per season. In addition, a model for MT data (i.e., considering all season images) was developed to analyze the effect of temporal information on the AGB prediction. The effects of temporal information were assessed based on the metrics defined in the Table 3.4.

The spectral analysis was carried out using the proposed approach on feature subset including all spectral channels (FS1) i.e., features computed using all available spectral bands with each sensor. The effect of spectral channels was assessed for the features selected by the adaptive lasso operator and the agreement of selected variables with AGB using the coefficient of determination. Features were divided into three groups according to the part of the spectrum i.e., Group 1 that consists features from the common (R, G, B and NIR) bands, Group 2 that consists features Group 1 and Red-edge bands, and Group 3 that consists features from Group 1 and SWIR bands.

The spatial analysis was carried out by analyzing AGB correlation scat-

Table 3.4: Assessment metrics for assessment of Precision, Agreement and Overfitting of the models computed using Leave-One-Out Cross-Validation (LOOCV)

Assessment Metrics	Equations	Assessment Role
Mean absolute difference	$MAD = \sum_{i=1}^{n} \left\lvert pre_i^{CV} - obs_i \right\rvert / n$	Precision
Root mean squared differences	$RMSD = \sqrt{\frac{SS^{CV}}{n}}$	Precision
Coefficient of determination (residuals)	$R^2_{fit} = 1 - SS_{fit}/SS_{tot}$	Agreement
Coefficient of determination (cross validation)	$R^2_{CV} = 1 - SS^{CV}/SS_{tot}$	Agreement
R-squared Ratio	$R2R = R^2_{fit}/R^2_{CV}$	Overfitting
Sum of Squares Ratio	$SSR = \frac{\sqrt{SS^{CV}}}{\sqrt{SS^{fit}}}$	Overfitting

n: total number of samples; pre_i^{CV}: prediction value of sample 'i' obtained by cross validation; obs_i: observed value of sample 'i'; SS^{CV}: sum of squared differences between observed and predicted values by cross validation; SS_{fit}: sum of squares of model residuals; SS_{tot}: sum of squared differences of each observation from the overall mean

ter plots of the most recurrent selected features computed computed at 3, 5, and 10 m spatial resolution. A Pearson correlation coefficient (R) was computed to observe the change in the correlation between the key selected variables at different spatial scales with the AGB.

3.4 Results and Discussion

3.4.1 Temporal Analysis

The effect of temporal information was analyzed using box plot visualization depicting the variation of precision, agreement, and overfitting on FS2 subset models using six metrics stated in Table 3.4.

The precision box plots ($MAD\%$,$RMSD\%$) shown in Figure 3.3 indicates that MT models have higher prediction precision compared to all single-time models. It further shows improved prediction precision by using MT information for both the study areas and for all sensors. The addition of MT information showed more significant improvement for Dove and RapidEye data characterized with higher spatial resolution. Overall, the

best precision was observed for Dove data in Lavarone area and RapidEye for Pellizzano area. A high variance was observed in the prediction precision of all single-time models. This indicates a significant influence of seasons (i.e. time of data acquisition) on the AGB modelling results.

The agreement box plots (R^2_{fit}, R^2_{CV}) shown in Figure 3.4 indicates that all MT models have a better agreement compared to single-time models. MT information improved agreement of all models for both the study areas. A significant improvement was observed for Dove and RapidEye data for the Pellizzano area with a greater sample size. The highest agreement (R^2_{CV}) was observed for MT RapidEye models for both the study areas. The MT models provides a significant advantage over the single-time models that are subject to high variance in the model agreement.

The overfitting box plots ($R2R$, SSR) are shown in Figure 3.5. The desirable values of $R2R$ and SSR should be less than 1.1 wherein the real (cross-validated) precision exceeded 10% of the model residual variance. Most MT models have a magnitude of overfitting close to 1.1 and belonged to the lower quartile of single-time models. Single-time models showed greater variance in the overfitting. This indicates that the level of model overfitting is dependent on the time of data acquisition. The highest level of overfitting was observed for single-time Sentinel-2 models for both the study areas. Overall, the overfitting metrics for Pellizzano were lower compared to Lavarone. It shows that overfitting can be influenced by the distribution and number of reference plots used for modelling.

3.4.2 Spectral Analysis

The spectral analysis was carried out for MT models using adaptive lasso feature regularization and selection applied to FS1 feature subset. Table 3.5 represents the features selected for all the three satellite data for both the study areas. The corresponding values in the bracket show the

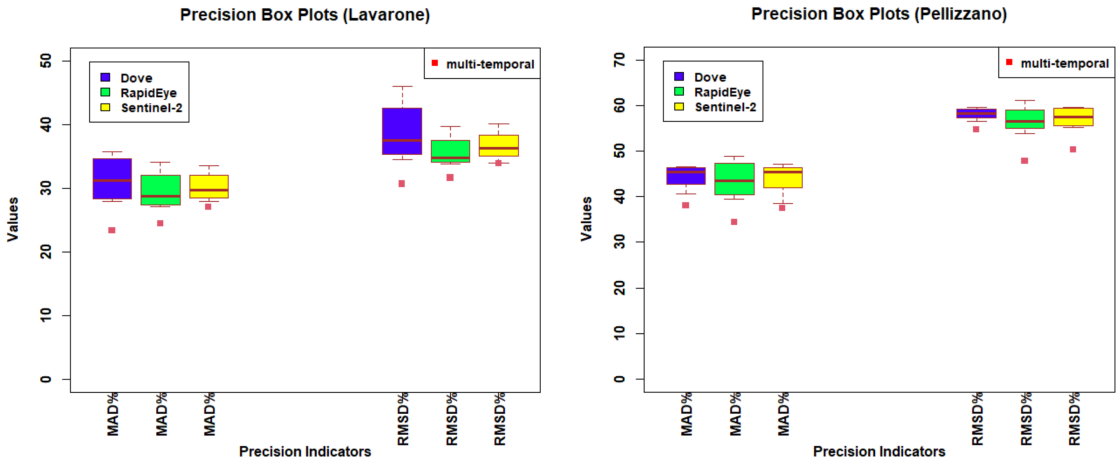

Figure 3.3: Precision Box Plots based on assessment metrics - MAD% and RMSD% computed for single-time and multitemporal models

frequency of occurrence. Table 3.5 offers a clear visualization of the selected features divided in three groups depending on the involved spectral bands. Additionally, Figures 3.6 and 3.7 show the regression plots (Field estimated AGB vs. Predicted AGB) with a cross validated coefficient of determination for Lavarone and Pellizzano area respectively.

Four features were primarily selected to develop the Dove MT models for both the study areas. The GLI feature was the most frequently (four times) selected among all features. The second most frequently selected features were VIgreen (two times) and RI (two times). Note that, GLI, VIgreen and RI - the three most frequently selected features for the Dove data were formulated without the NIR band. The agreement in terms of coefficient of determination were 0.37 and 0.33 for Lavarone and Pellizzano respectively. A total of eight features were selected for the RapidEye MT models. GLI, NDVIre and CCCI were the three most frequently selected features, two among which belonged to Group 2 that involve red-edge bands. The Group 2 significantly contributed to the better agreement of RapidEye models as compared to Dove models. The agreement for RapidEye MT models were

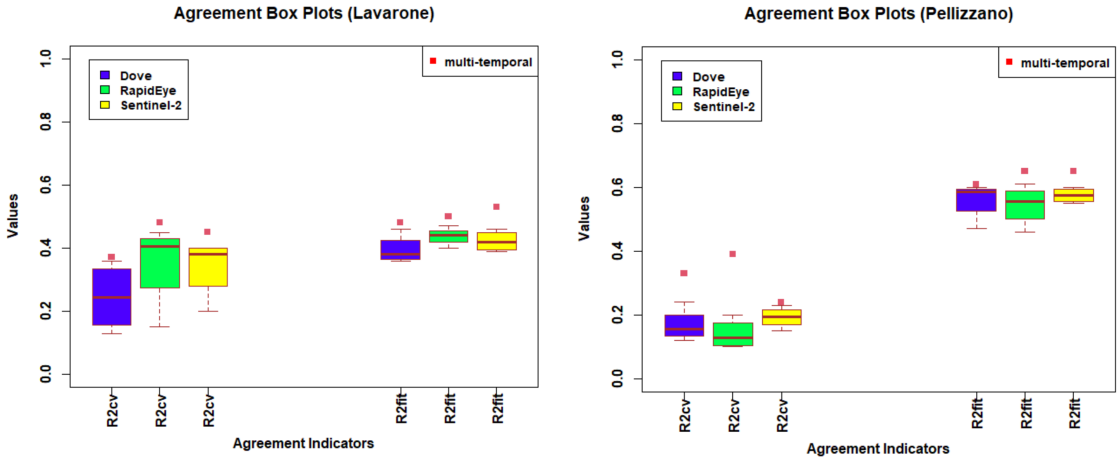

Figure 3.4: Agreement Box Plots based on assessment metrics - R2cv and R2fit computed for single-time and multitemporal models

0.51 and 0.44 for Lavarone and Pellizzano respectively. NDVIre (selected thrice) was the most frequently selected feature for Sentinel-2 MT model. CCCI and CIRE (selected twice) were the other two features frequently selected from Group 2 and GLI from Group 1. NBR was the only feature from Group 3 selected twice that involve SWIR bands. Group 2 and Group 3 features contributed for more than 50% of the total selected features. This delivered more significant improvement in the agreement for RapidEye and Dove models. The coefficient of determination for Sentinel-2 MT models were 0.53 and 0.51 for Lavarone and Pellizzano respectively.

To sum up, GLI was consistently selected feature from all the groups. Sentinel-2 MT models achieved highest modelling agreement as compared to RapidEye and Dove. Moreover, Red-edge and SWIR bands significantly improved the modeling accuracy for the prediction of AGB.

3.4.3 Spatial Analysis and Mapping AGB

The effects of spatial resolution on AGB prediction were analyzed by observing the change in the correlation of key spectral features at different

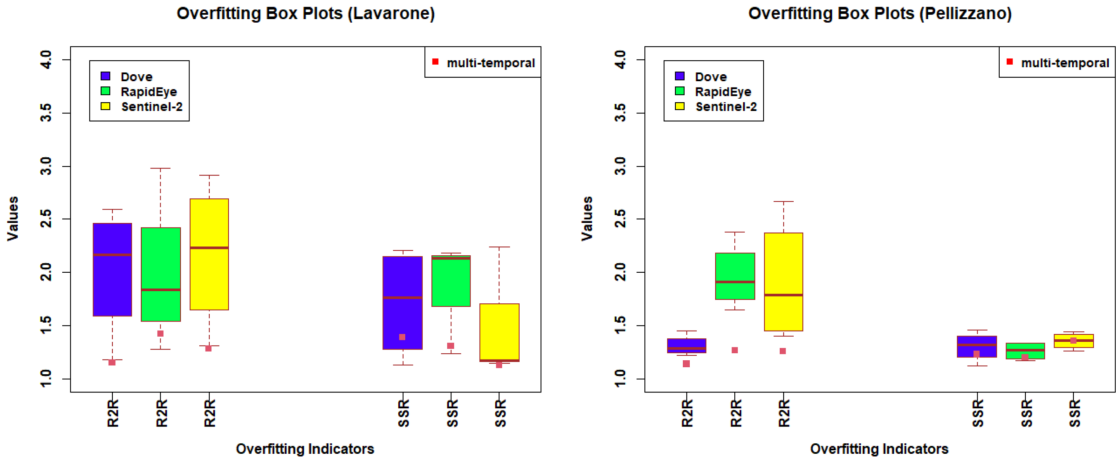

Figure 3.5: Overfitting Box Plots based on assessment metrics - R2R and SSR computed for single-time and multitemporal models

spatial scales. GLI and VIgreen from Group 1 and CCCI and NDVIre from Group 2 were used as key spectral features for the analysis. Figure 3.8 shows the scatter plots of GLI vs. AGB for the Lavarone area and Figure 3.9 shows the scatter plots of VIgreen vs. AGB for the Pellizzano area at 3, 5, and 10 m. Figure 3.10 shows scatterplots of CCCI vs. AGB and Figure 3.11 shows scatterplots of NDVIre vs. AGB at 5 and 10 m for Lavarone and Pellizzano respectively. An identical trend was observed for key spectral features in the Group 1 and Group 2 with the highest correlation observed at 10 m spatial resolution. The highest Group 1 correlation (Pearson coefficient) for GLI and VIgreen are 0.44 and -0.43 respectively at 10 m. The highest Group 2 correlation observed for CCCI was 0.58 and for NDVIre was -0.48 at 10 m. There was no significant difference in correlation for CCCI at 5 m and 10 m but a significant difference was observed for NDVIre.

The spatial distribution of the predicted AGB over the selected areas was mapped to perform the qualitative analysis. The best performing MT Sentinel-2 models for both study areas were used for predictive mapping.

Table 3.5: Selected features using adaptive least absolute shrinkage and selection operator for Group 1, 2 and 3.

	DOVE	RAPID-EYE	SENTINEL-2
Group 1	GLI (4) VIgreen (2) RI (2) normR (1)	GLI (2) VIgreen (1) RI (1) normR (1) normG (1) BWDRVI (1)	GLI (2) normR (1) normG (1) VIgreen (1) RI (1)
Group 2	-	NDVIre (2) CCCI (2)	NDVIre (3) CCCI (2) CIRE (2)
Group 3	-	-	NBR (2)

The AGB maps for Lavarone and Pellizzano are shown in Figure 3.12 and Figure 3.13 respectively. The non-forested areas were masked from the maps.

3.4.4 Discussion

In this thesis contribution, spectral features extracted from MS satellite data were used to predict AGB at various spatial scales from medium to high spatial resolution data. The results were analyzed with respect to temporal, spectral, and spatial capacities of different MS satellite data to assess their potential for AGB prediction. The results presented from the experiments can be used as guidelines to frame a viable strategy for the use of SMS data for the prediction of AGB.

Spectral features were used to predict AGB in several studies with varying degree of effectiveness. The study presented in [11] demonstrated single-time spectral bands and spectral indices from the Sentinel-2 data used to predict forest stock volume. A higher prediction error was observed for forest stock volume greater than 300 m^3 ha^-1 for all single-time models. This thesis contribution observed similar limitation of single-time models

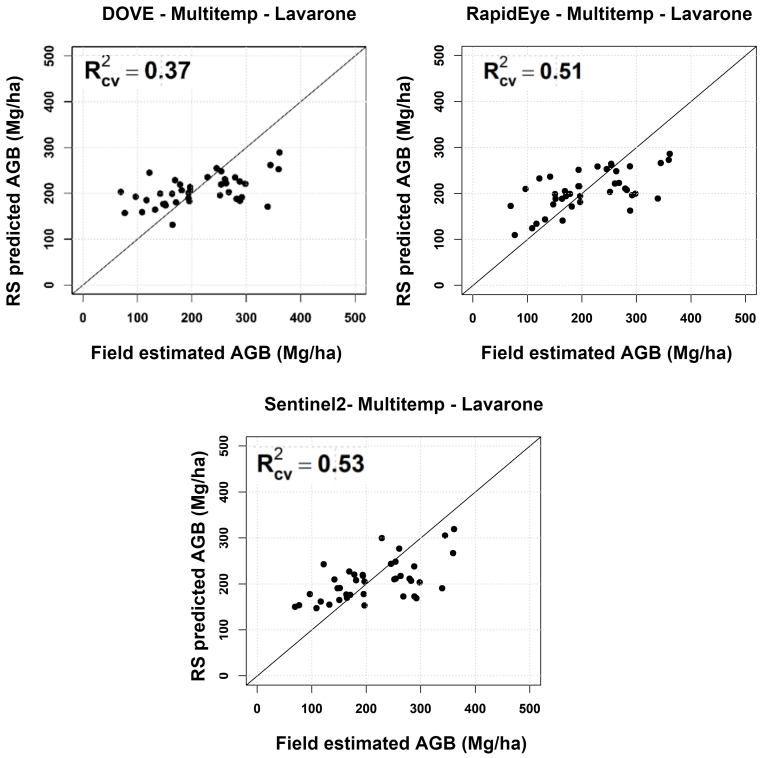

Figure 3.6: Multitemporal Regression Plots representing Field Estimated Versus RS Predicted AGB using FS1 Feature Subset —Lavarone area

with greater prediction errors and high variance in precision, agreement, and overfitting. The experiments with MT data served as a solution to overcome these limitations. MT data provided supplementary information to the models on seasonal variations for all spectral features. MT models for all datasets (Dove, RapidEye, and Sentinel-2) achieved better modelling accuracy than their respective single-time models. Thus, the MT information complemented the spectral features at all spatial scales for both the study areas characterized by different forest compositions. Despite the heterogeneous and complex forest ecosystem, the MT models delivered better modelling accuracy as compared to single-time models. Also, the annual growth of the stem diameter (hence the AGB) is dependent on factors such as rainfall, temperature, humidity, day length, species, etc. [93] directly af-

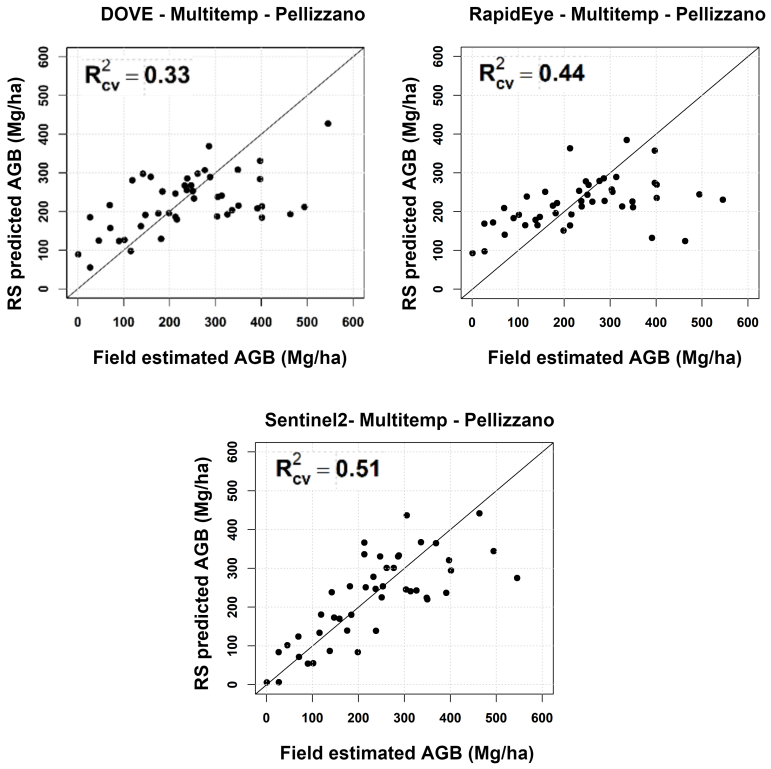

Figure 3.7: Multitemporal Regression Plots representing Field Estimated Versus RS Predicted AGB using FS1 Feature Subset —Pellizzano area

fects the seasonal patterns of vegetation. Therefore, the seasonal variation in spectral reflectance of the forest plots captured using MT spectral features can be related to the structural measures of the forest, and hence the AGB. Based on such assumptions, multiple studies such as [94, 9, 5] were performed using Landsat MT data for the prediction of forest AGB. This thesis contribution provided a few improvements in methodology and analytical approach to quantify the key aspects (i.e., precision, agreement, and overfitting observed in time) for better analysis and prediction of AGB. In addition, the analysis was carried out by experimenting at various spatial scales (3, 5, and 10 m) of MS data that provided greater understanding of the effect of spatio-temporal information on AGB prediction.

The feature selection process is crucial for accurate AGB modelling with

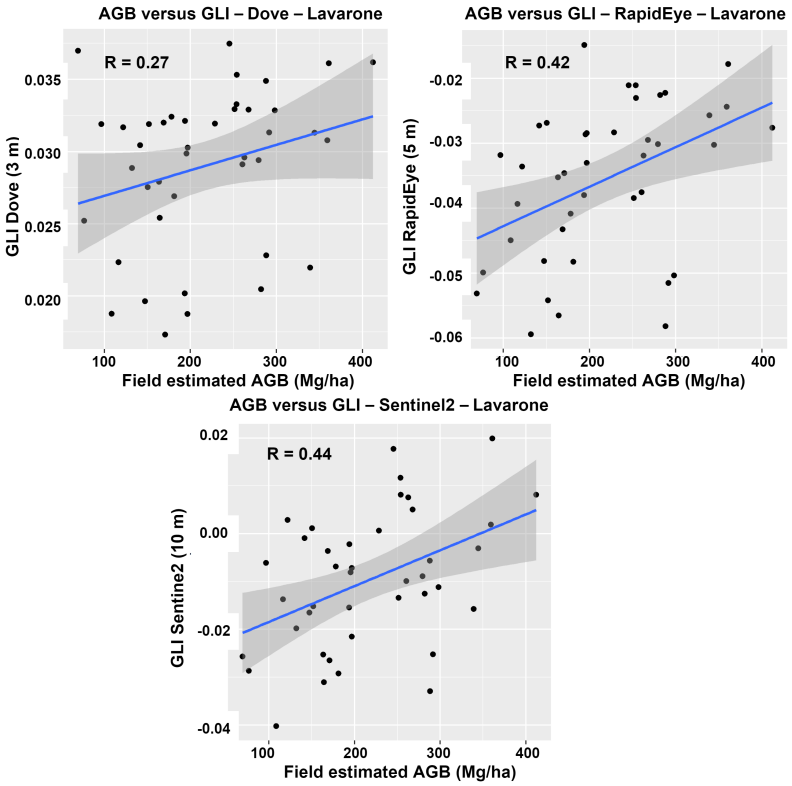

Figure 3.8: Scatterplots of AGB Versus Green Leaf Index (GLI) computed at different spatial scales (10, 5 and 3 m) using FS2 subset for Lavarone area

spectral features from SMS data. A few studies [95, 11, 96, 12, 97] bypass feature selection and produce models with overfitting and prediction errors due to less important or redundant features. For instance, the study in [96] predicted stand volume by using Landsat data and evaluated the effectiveness of additional feature variables. Spectral features were added to a basic model one at a time to determine an optimum combination without any appropriate feature selection resulting to only minor improvement in the estimation accuracy. Similarly, [95] used the individual band reflectances and vegetation indices as features for estimating AGB without feature selection or elimination procedure. In this thesis contribution, a generalized linear model was developed using an adaptive lasso feature regularization

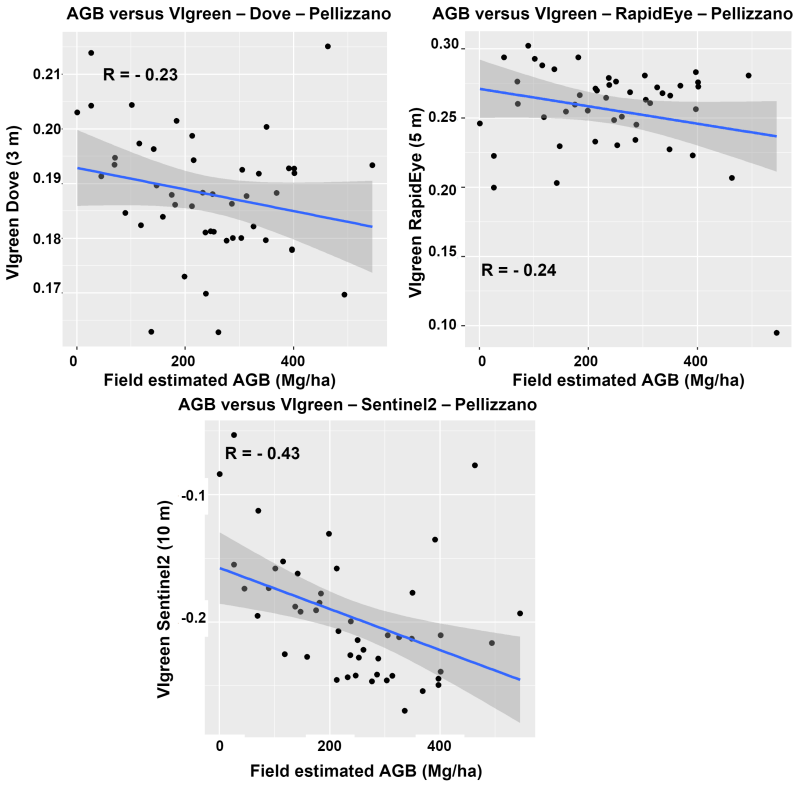

Figure 3.9: Scatterplots of AGB Versus Visible Index Green (VIgreen) computed at different spatial scales (10, 5 and 3 m) using FS2 subset for Pellizzano area

and selection method that possessed oracle properties. Different capacities of SMS data were individually and systematically analyzed providing a better examination of relationship between features and AGB. The study conducted in [11] supported the findings of this thesis contribution and confirmed the importance of SWIR and red-edge bands of Sentinel-2. This thesis contribution provided an additional correlation analysis of such key features quantified by Pearson correlation coefficient across various spatial scales. Moreover, all experiments and analysis were conducted for two different study areas in order to account for the diversity and distribution of forests that can affect the spectral features.

A few studies have used Sentinel-2 data with different AGB modelling

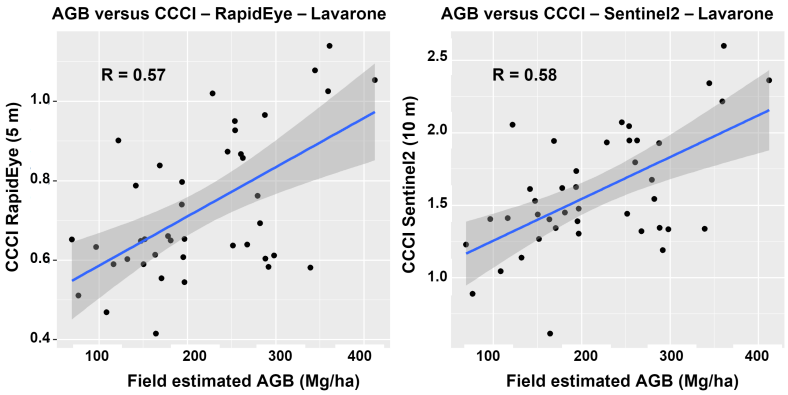

Figure 3.10: Scatterplots of AGB Versus Canopy Chlorophyll Content Index (CCCI) computed at different spatial scales (10, 5 and 3 m) using FS2 subset for Lavarone area

approaches [12, 97, 98]. The study in [98] extracted spectral features from Sentinel-2 and Landsat-8 for the prediction of forest variables. The study concluded that there was no significant difference in the systematic error between the two sensors. Although, it was noted that there was no improvement in the prediction accuracy for the combination spectral features from Sentinel-2 and Landsat-8. This outcome remained unexplained from the analysis conducted in the study and was deemed "surprising". However, it can be explained from analysis performed in this thesis contribution. The correlation of selected key features were analyzed for different spectral ranges at different spatial resolutions. The significant variation in a correlation of the same key variables at different spatial scales pointed out an instrumental role of spatial resolution in the prediction of AGB. But the spatial analysis showed that there was no significant increase in the correlation of most key spectral features with the spatial resolution (especially between 5 m and 10 m data). This caused no significant improvement in prediction accuracy on combination of spectral features of different spatial resolution (Sentinel-2 and Landsat-8) as observed in [98]. Therefore, a higher spatial resolution of MS data may be less relevant for

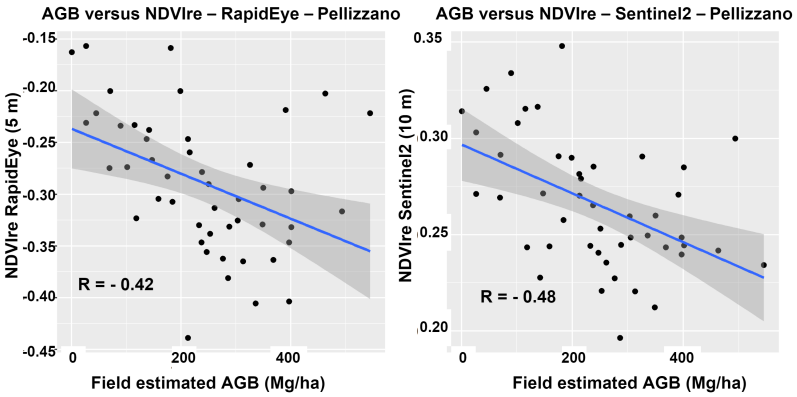

Figure 3.11: Scatterplots of AGB Versus Red Edge NDVI (NDVIre) computed at different spatial scales (10, 5 and 3 m) using FS2 subset for Pellizzano area.

a prediction method based on spectral features. This contribution considered spectral features exclusively (and not the other derivable features) to suit the objectives and design of the experiments. The temporal properties of MS data over different seasons, the spectral properties over different spectral bands and the spatial properties over different resolution were investigated using the extracted spectral features. With regard to alternate feature extraction mechanisms (For e.g. deep features, textural features, principal components etc.) or modelling procedures (e.g. object oriented or patch-level modelling), the deductions from the results can be considered limited to this particular experimental setup that was deployed for strategically achieving objectives of this contribution. It is also worth noting that the role of high spatial resolution of the MS images may be more relevant for textural features or object-oriented modelling.

This thesis contribution involved use of Copernicus and Planet MS data and can benefit public space agencies such as ESA and private sector enterprises such as Planet Labs for upgrading instrumentation for forest applications. In addition, this contribution can prove relevant and useful for the development of more efficient tools and techniques based on auto-

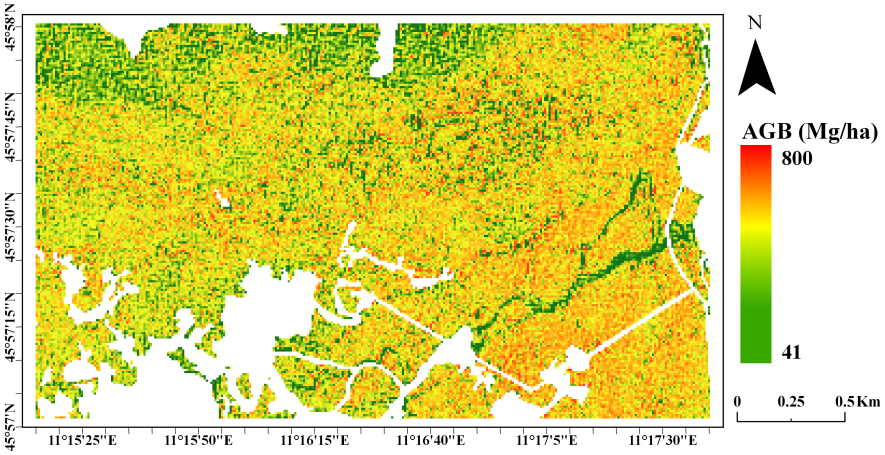

Figure 3.12: Spatial Map of AGB predicted using the best MT Sentinel-2 model for Lavarone area

matic feature extraction and feature selection with more advanced modeling methods using MS data. However, the data saturation problem is not completely resolved and requires more explanatory features from alternate source of data for the prediction of AGB. Also, the limited field plot samples restricted model performance up to some extent and could be further improved with more samples. The spectral features are required to be engineered so that they capture AGB-specific variation. This could increase the effectiveness of the spectral features and improve the modelling accuracy. The results from the conducted experiments and analysis indicated a good potential of MS data for AGB prediction.

3.5 Conclusion

In this thesis contribution, different MS satellite data were used to characterize the effects of temporal, spectral and spatial capacities of the data for predicting forest AGB. The results showed that each characteristic of the MS data distinctly affects the process of prediction of AGB. The MS data with red-edge and SWIR bands improved the modelling accuracy and

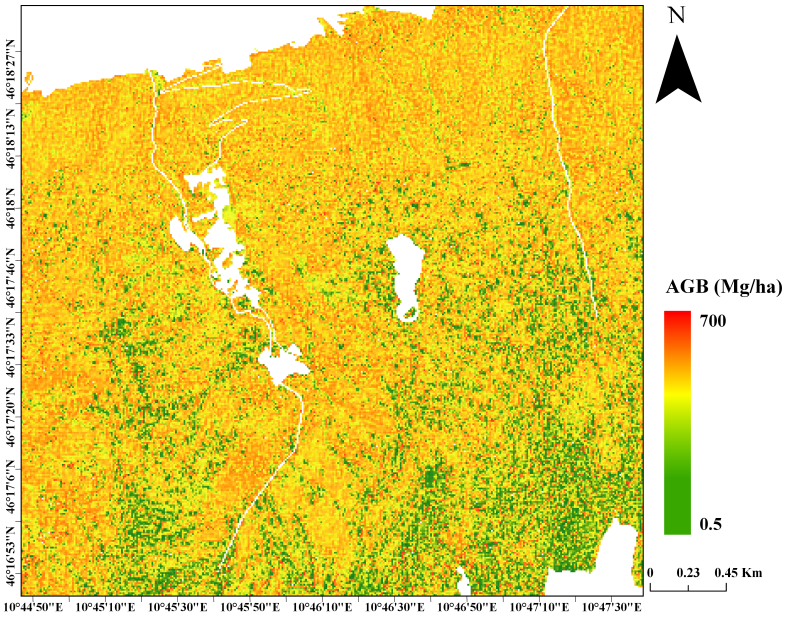

Figure 3.13: Spatial Map of AGB predicted using the best MT Sentinel-2 model for Pellizzano area

decreased the saturation of models. The large AGB values were predicted with greater accuracy as compared to models based on MS data without red-edge and SWIR bands. The seasonal spectral features generated variation in the predictions in terms of precision, agreement and overfitting. The MT data models delivered better accuracy results as compared to the single-time models. The correlation of key features at different spatial scales show better correlation at coarser spatial resolution and no significant difference for RapidEye and Sentinel-2. This indicates that spectral features are less affected by spatial resolution of the MS data for modelling AGB.

Chapter 4

Generative Features from Multi-source Satellite Remote Sensing Data for AGB Prediction

4.1 Introduction

The extraction of effective features and producing precise training data from Remote Sensing (RS) images is crucial for accurately quantifying and mapping various ecological parameters including the carbon stored in forests [99, 100, 101, 102, 103]. The quality of these extracted features directly affects the response of the modelling algorithm to predict the target parameter [104, 105, 106]. Conventional features from RS data are primarily based on specific indices derived from analytical formulations (i.e. arithmetic operations performed on spectral bands, radar backscatters or elevation statistics of lidar point clouds) [107, 18, 108]. In the literature, these analytical features are frequently used for different applications like

Part of the chapter appears in:

P. Naik, M. Dalponte, and L. Bruzzone, "Generative Feature Extraction From Sentinel 1 and 2 Data for Prediction of Forest Aboveground Biomass in the Italian Alps," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15. Institute of Electrical and Electronics Engineers (IEEE), pp. 4755–4771, 2022. DOI: 10.1109/JSTARS.2022.3179027

analysis of vegetation, crops and soil properties [26]. Their effectiveness is found to be highly data dependent [23, 109]. 3D airborne lidar data are very effective in modelling forest parameters. Features such as height and intensity percentiles, height bins and density metrics are used for the prediction of forest AGB with lidar [110]. However, the effectiveness of such analytical features depends on the point density and the footprint of the laser pulse [111, 112]. Moreover, the limited availability of lidar data and the high cost for finer specifications discourages its use in many operational scenarios. In this context, SRS data proves to be a more feasible choice for the forest applications.

Studies performed with SMS and SAR data used vegetation indices (e.g., NDVI, SAVI), vegetation biophysical features (e.g., leaf area index, chlorophyll concentration), textural features (e.g., entropy, variance) and SAR polarimetric indices for the prediction of forest AGB [113, 114, 115]. A prominent limitation identified with MS data was a high variation in response of the extracted features that depended on the time of acquisition and the spatio-spectral specifications [114, 116]. The main causes of low AGB prediction accuracy are linked to the seasonal variation of the spectral responses from the forest. Moreover, the radiometric characteristics of different satellite sensors also affects the performance of the models. In [114], it was observed that MS sensors with different radiometric specifications for an identical spectral range produce non-identical responses. A few other studies that used Sentinel-2, RapidEye and PlanetScope data on different forest types also indicated such prediction precision problems that can be traced back to variations in spectral responses and radiometric specifications of sensors [20, 63]. Despite these identified challenges, a dedicated attention on extracting target-oriented features from SMS and SAR data can greatly increase the prediction precision and provide an economical source for forest parameter estimation.

Different Machine Learning (ML) algorithms (such as RF, ANN, Gaussian processes, SVM etc.) have been used in the literature to model forest AGB using various multi-sensor RS data [128, 129, 69]. Some of these ML algorithms (RF, ANN and SVM) have also been used in combination via ensemble learning for further improving AGB prediction accuracy as compared to single ML algorithms [130, 19]. The recent advances in ML and AI introduced generative and reinforcement learning based methods to extract and engineer target-oriented features [125, 126, 127]. Deep neural networks are capable of extracting high level abstract features from complex distribution underlying various RS data [131, 132] by generating an optimal feature space for modelling a given problem. Such abstract features have delivered improved performances in various RS applications [103, 133, 134, 135]. However, the use of such abstract features for regression problems has been less studied as compared to other applications in the field of RS [136, 137, 138]. In recent papers, several effective generative approaches using Generative Adversarial Networks (GAN) and Conditional GAN have been developed to learn and productively engineer features from the data [139, 140, 141, 142, 143, 144]. This has led to the use of generative neural networks to effectively disentangle the latent space of the network to explain specific factors of variation in a target variable, thus showing a high potential to solve regression problems [145, 146, 147, 148, 149]. Unlike traditional approaches that used a conventional multi-layer perceptron or stacked sparse autoencoder [150, 151, 152], a generative neural network can be used to conditionally distribute the feature space on a target variable, consequently improving the accuracy and generalization of a regression model.

The objective of this contribution is to develop a dynamic generative neural network architecture to model forest AGB using multi-source (MS and SAR) SRS data. The dynamic framework of the generative neural

network is based on variational autoencoding that can extract abstract target-oriented features from the data. The network aims to generate a high-ordered feature space that can potentially minimize the problem of variability of features for AGB prediction. The network also aims to reduce the input dimensionality to produce a low dimensional feature space for reducing the model complexity. This contribution also aims to design the network architecture such that it can adjust the number of network parameters as per the dimension of the input data to quickly optimize the network and learn effectively. The objective to develop such a network is aimed at performing target-oriented feature engineering and modelling AGB using a single dynamic neural network architecture.

4.2 Study Area and Datasets

4.2.1 Study Area Description and Field Data

The study area consists of three sites – Lavarone, Pellizzano and Cembra. These sites are diverse in terms of species, distribution and density of forest located in the Province of Trento, Italy. The 32 km^2 Pellizzano area is located in the municipality of Pellizzano (46°17'22"N,10°46'05"E), the 26 km^2 Lavarone area is located in the municipality of Lavarone (45∘57'30.09"N, 11∘16'25.17"E) and 16 km^2 Cembra area is located in the municipality of Cembra Lisignago (46°10'29" N, 11°13'18" E) in the Italian Alps. The detailed description of forest species and population for Lavarone and Pellizzano sites are described in Chapter 3 (Study area sub-section). The forest in Cembra mainly consists of broadleaves species such as European beech (*Fagus sylvatica* L.) and few coniferous species. The geographical locations of the study area sites are shown using detailed maps in Figure 4.1.

The field data consist of 115 circular plots with fixed radius of 15 m

Table 4.1: Acquisition Dates (YYY-MM-DD) of the four Sentinel-2 and Sentinel-1 Images acquired over different seasons

Seasons	Dates of Acquisition (YYYY-MM-DD)	
	Sentinel-2	Sentinel-1
Spring	$2016-03-10$	$2016-03-04$
Summer	$2016-06-28$	$2016-06-28$
Autumn	$2016-09-06$	$2016-09-08$
Winter	$2016-12-28$	$2016-12-29$

(Figure 4.1). A total of 47 plots in Pellizzano, 48 plots in Lavarone and 20 plots in Cembra were measured using a random sampling design. The plot samples were measured on field in summer of 2014 for Pellizzano area and summer of 2016 for Lavarone and Cembra areas. The details of survey devices, measurement parameters and species specific allometric equations is given in Chapter 3 (Field data sub-section). The field estimated plot level AGB values ranged from 1.07 $Mg\ ha^{-}1$ to 655.14 $Mg\ ha^{-}1$. The independent validation of the proposed model was performed using 55 additional plots surveyed over different sites in the Autonomous Province of Trento. The AGB values of these plots used for independent validation ranged from 11.30 $Mg\ ha^{-}1$ – 711.41 $Mg\ ha^{-}1$.

4.2.2 Remote Sensing Data

The study was performed using MT images acquired by ESA's Sentinel-2 and Sentinel-1 satellite constellations. Four images per satellite constellation were considered, one for each season (Table 4.1). The specifications of Sentinel-2 images are stated in Chapter 3 (Remote sensing data subsection). The Sentinel-1 data was acquired from the NASA's Earth Data portal (`https://urs.earthdata.nasa.gov/`). Level-1 Single Look Complex (SLC) data comprising complex imagery with amplitude and phase captured in the IW mode were used. The resolution of the data (range X

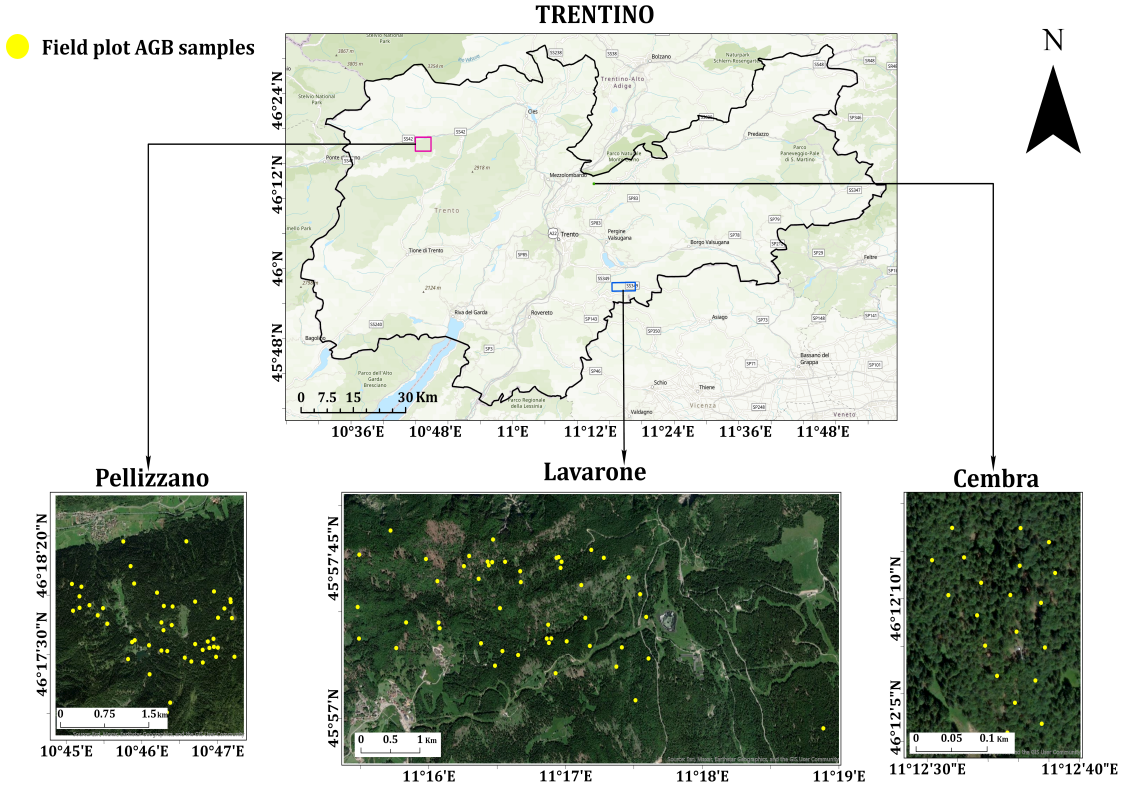

Figure 4.1: Location of the three study sites (Pellizzano, Lavarone and Cembra) and the plot locations of field estimated AGB (yellow dots) on Google Earth images

azimuth) ranges from 2.7 X 22 m to 3.5 X 22 m and the pixel spacing (range X azimuth) is 2.4 X 14.1 m with 1 X 1 number of looks. The acquired SLC product has all bursts in all sub-swaths and resampled to a common pixel spacing grid in range and azimuth. The Sentinel-2 and Sentinel-1 images of the study area for the summer season are shown in Figure 4.2.

4.3 Proposed Generative Model

4.3.1 Generative Model Framework

The proposed generative model adopts a variational autoencoder based modelling framework with a dynamic architecture. A traditional varia-

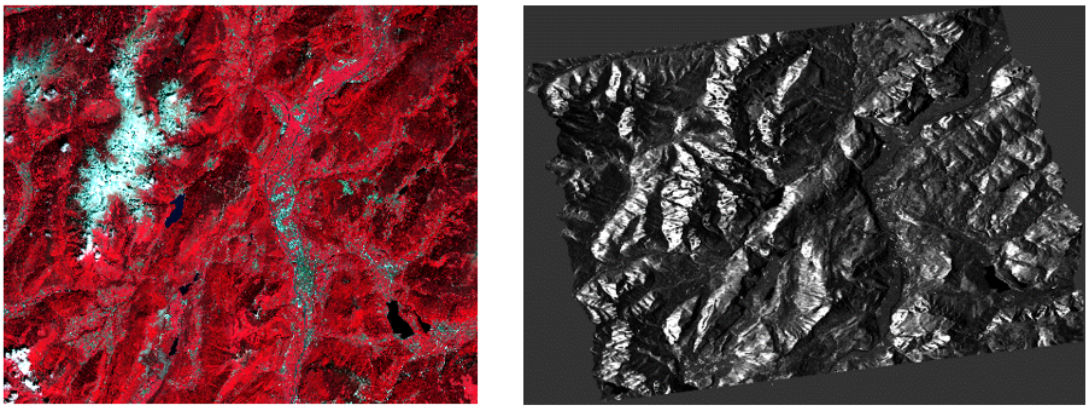

Figure 4.2: Standard False color composite of Sentinel-2 image (left) and Polarimetric matrix element (C_{11}) of Sentinel-1 image (right) of the study area acquired in the summer season

tional autoencoder consists of two main component networks – an encoder and a decoder. The additional two networks added to this traditional architecture are - a latent generator and a regressor network. Unlike traditional approaches that independently trains the regressor, the proposed approach has an integrated regressor network connected to the reparametrized latent space via a latent generator network. All networks are inter-associated with a combination of three loss functions as shown in Figure 4.3.

The probabilistic encoder network ($q(s \mid i)$) learns the distribution from the input data (i) and generates a reparametrized latent space (s) using a conditional prior of the target parameter (t) instead of a traditional Gaussian prior. The dimension of the input data is compressed by the encoder network to the dimension of the latent space and the latent generator ($p(s \mid t)$) captures the target specific prior of the latent representations. In this process, the decoder network ($p(i^{'} \mid s)$) captures the non-linearity of the latent generator via the generated latent representations and attempts to reconstruct the input. The reconstructed form of the input data is given as ($i^{'}$). The conjoined feed-forward neural network ($q(t \mid i)$) is a probabilistic regressor which predicts outputs along with the standard deviation

in the predicted values.

All the four networks are regularized by a loss mechanism that optimizes the model to learn targeted representations from the data and accurately predicts the target parameter. The latent generator with the decoder network accounts for the 'generative parameters' and the encoder network with the regressor network accounts for the 'inference parameters'. This is because the reconstructed data is assumed to be generated from its latent representation that depends on the target parameter. Therefore, the latent generator and the decoder network function as a 'generative model'. The 'inference model' consists of the probabilistic encoder that determines the latent representation from the data and a probabilistic regressor that predicts the target parameter using the latent features. The total loss $\mathcal{L}(i)$ of the proposed generative neural network can be stated as shown in equation (4.1),

$$\begin{aligned} \mathcal{L}(i) = & - D_{KL}(q(t \mid i)||p(t)) \\ & + E_{q(s|i)}[\log p(i \mid s)] \\ & - E_{q(t|i)}\left[D_{KL}(q(s \mid i)||p(s \mid t))\right] \end{aligned} \tag{4.1}$$

In the given equation (4.1), the total loss $\mathcal{L}(i)$ is a sum of three different loss terms associated with regularizing the inference and generative parameters of the proposed network. Particularly, $q(t \mid i)$ is a conventional feed forward regression network that produces bias (standard deviation) as an additional output (therefore a probabilistic regressor). The first term represents the Kullback-Leibler (KL) loss that regularizes the prediction of target with a prior. The second term represents reconstruction loss that emphasizes the reconstructed data to resemble to the input data and the third term represents the label loss that emphasizes the encoder (posterior - $q(s \mid i)$ to resemble to the target-specific prior $p(s \mid t)$.

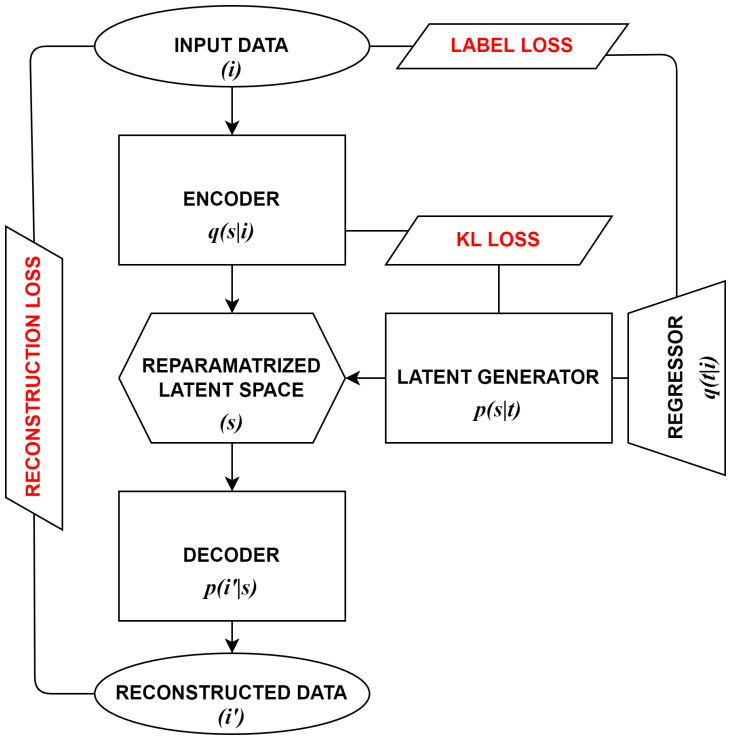

Figure 4.3: Framework of the Proposed Generative Model demonstrating the triple loss mechanism to regulate inference and generative parameters of the network

4.3.2 Generative Network Architecture

The dynamic architecture of the generative neural network shown in Figure 4.4 consists of an encoder network with an input layer of dimension 'n' connected to two intermediate hidden layers of dimensions n-2 and n-4 with 'tanh' activation function where 'n' is the number of input features. The resultant features from the hidden layers were independently connected to two other dense layers of same dimensions (n-6, n-6) that characterize the mean and diagonal covariance of the latent space. The probabilistic regressor network shared the two hidden layers of the encoder network and consisted of a simple dense layer to produce predicted mean and standard deviation of the target parameter. A standard reparametrization trick was applied to use the mean and variance as arguments to return sampled

vectors (z and r) enabling backpropagation through the network. A latent generator of dimension 'n-6' was used to condition 'z' on 'r' by the means of KL divergence loss function. Finally, the decoder network was built with an exact inverse structure of the encoder network for the reconstruction of data from the reparametrized latent space.

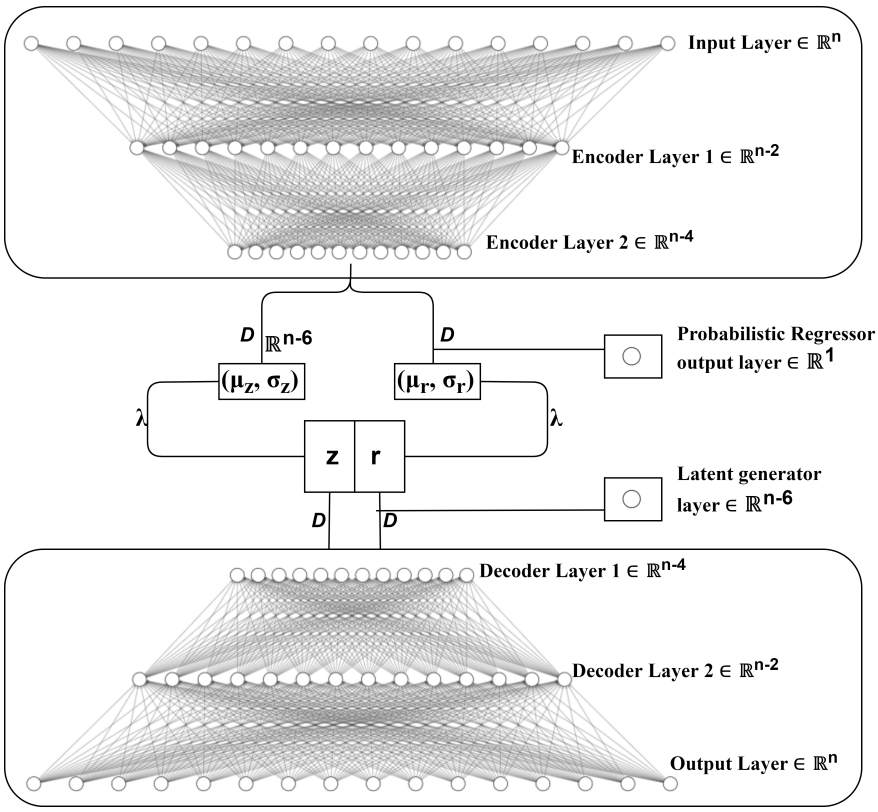

Figure 4.4: Dynamic architecture of the proposed generative network with variable layer dimensions
R^n: Network layer of dimension 'n'; D: Dense network connection; λ: Reparametrization function; (z, r): Sampled latent vectors; (μ_z, σ_z): mean and variance of 'z'; (μ_r, σ_r): mean and variance of 'r'

4.4 Proposed Approach for AGB Prediction

The flowchart of the AGB prediction approach is shown in Figure 4.5. In the following sub-sections each part of the flowchart is described in detail. The data pre-processing of MS and SAR data is described in sub-section 4.4.1. The process of computing analytical features from both pre-processed data and the steps for suitable data preparation as model input are stated in sub-section 4.4.2. Lastly, the implementation of the developed generative model and performed experiments are described in sub-section 4.4.3.

The general idea supporting the development of this methodology from an application (AGB) perspective is based on the results achieved in Chapter 3. The spectral features used for interpretation of biophysical property of forests such as AGB are based on passive systems. The sun view angles, soil background and atmospheric conditions may affect the reflectance measurements from forests. Thus, while dealing with spectral features, it is necessary to ensure that the developed models are more responsive to actual data rather than noise. Although, spectral features such as vegetation indices remove such variations in spectral reflectance, they are unable to sufficiently capture the variance of forest AGB. The proposed method was developed with an intent of using a generative process to produce features that can sufficiently capture the variance and exhibit properties of target parameter i.e. AGB. The SAR data was included to compensate for the drawbacks of passive system and providing additional features to explain AGB.

4.4.1 Data Pre-processing

The Sentinel-2 images were acquired in Level-1C (Top of the atmosphere reflectance) format and converted to Level-2A (Bottom of the atmosphere

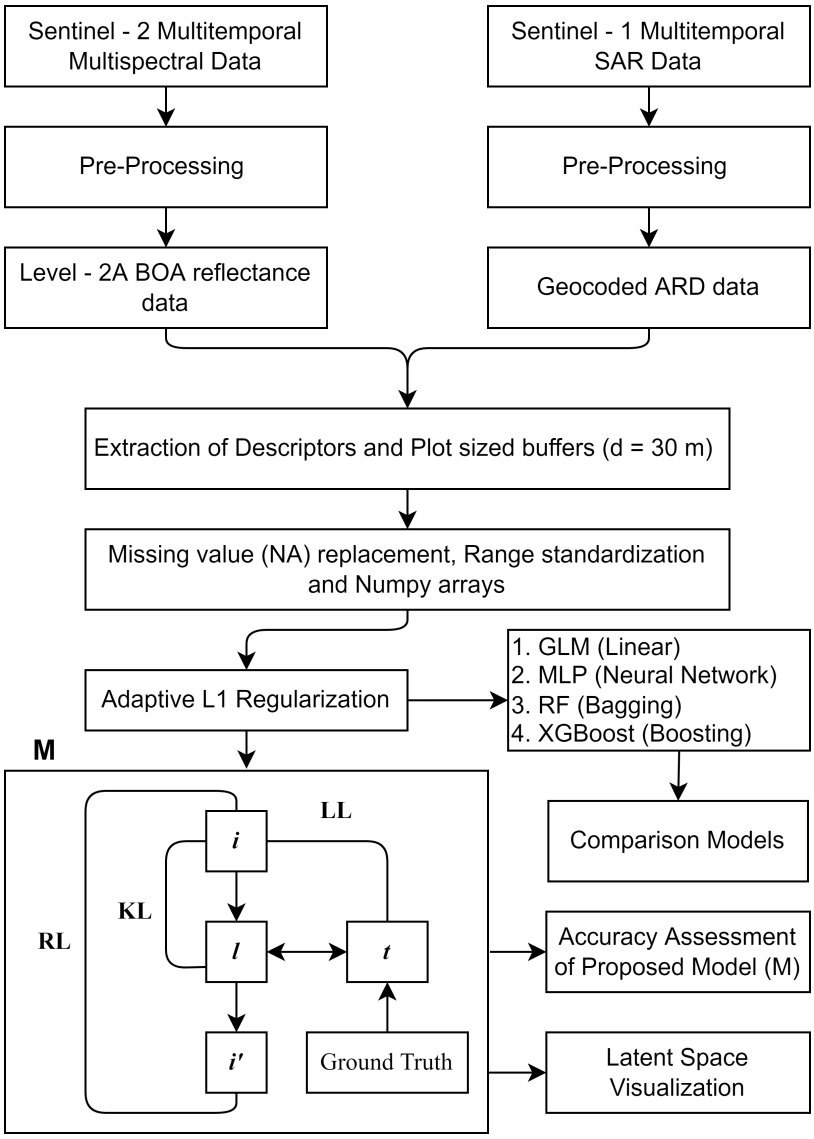

Figure 4.5: Flowchart of the Proposed Approach showing steps for Data Pre-processing, Feature computation and Model Implementation
BOA: Bottom of Atmosphere; ARD: Analysis Ready Data; GLM: Generalized Linear Modelling; MLP: Multilayer Perceptron Regressor; RF: Random Forest; XGBoost: Extreme Gradient Boosting; M: Proposed Generative Model; i: Input; i': reconstructed; l: latent space; t: target; RL: Reconstruction loss; KL: KL loss; LL: Label

reflectance) format including atmospheric and terrain correction using the Sen2cor processor [153]. The spectral bands at spatial resolution 10 m

and 20 m were used in this study and those at 20 m spatial resolution were resampled at 10 m for spatial consistency and for performing computations.

The acquired Sentinel-1 SLC product swath was split with selected bursts into a separate product. The orbit file was acquired and precise orbits are applied to the split Sentinel-1 product. The resultant product was radiometrically corrected and the calibrated SAR images were produced with pixel values that represent the radar backscatter of the reflecting surface. The bursts of the product were merged in the azimuth direction for a seamless image and the debursted split products of different sub-swaths were merged to form a single image. A subset was clipped from the single merged image of the area of interest (study area). In order to produce a polarimetric Analysis Ready Data (ARD), a C2 polarimetric matrix was generated that is an incoherent polarimetric representation of second order partial polarimetric scattering matrix elements. Multilooking and speckle filtering were performed on the resultant ARD product to reduce speckle noise and generate ground range square pixels. Finally, the range doppler terrain correction was performed to geocode and produce a final product of 10 X 10 m spatial resolution. The entire Sentinel-1 data pre-processing was performed using Sentinel Application Platform version 8.0 software (`https://step.esa.int/`).

4.4.2 Computed Features and Input Data Preparation

The pre-processed data for each season were used to compute analytical features, i.e. 25 vegetation indices from Sentinel-2 (Table 3.3) data and four radar vegetation indices from Sentinel-1 data (Table 4.2). The numerous vegetation indices consist of discrete information from the spectral and backscatter data. These vegetation indices reduce the effect of environmental conditions and compensate for atmospheric distortions. Moreover, they maximize the sensitivity to biophysical properties and minimizes to-

Table 4.2: Analytical Features From Sentinel-1 SAR Data extracted using First and Second Order Scattering Information with their respective Equations

Analytical SAR Features	Equations
Radar Vegetation Index	$\mathrm{RVI} = \frac{4\sigma^0_{\mathrm{VH}}}{\sigma^0_{\mathrm{VV}}+\sigma^0_{\mathrm{VH}}}$
Degree of Polarization	$\mathrm{DOP} = \sqrt{1 - \frac{4\lvert \mathrm{C}_2 \rvert}{(\mathrm{Tr}(\mathrm{C}_2))^2}}$
Dual-pol Radar Vegetation Index	$\mathrm{DpRVI} = (1 - \mathrm{DOP}) * \frac{\lambda_1}{\lambda_1+\lambda_2}$
Polarimetric Radar Vegetation Index	$\mathrm{PRVI} = (1 - \mathrm{DOP}) * \sigma^0_{\mathrm{VH}}$

σ^0_{VH}: Cross-polarized backscattering coefficient; σ^0_{VV}: Co-polarized backscattering coefficients; C_2: Covariance matrix; Tr: Matrix Trace Operator; λ_1, λ_2: Eigen values of Covariance Matrix C_2

pographical effects. Therefore, they can be used as input for engineering target-oriented generative features with the proposed framework.

The analytical features from Sentinel-1 data were computed using a QGIS plugin - PolSAR Tools [154]. In addition to the stated analytical features, ten reflectance bands of Sentinel-2 data and four polarimetric matrix elements ($\mathrm{C}_{11}, \mathrm{C}_{12}, \mathrm{C}_{21}, \mathrm{C}_{22}$) of the Sentinel-1 data were also considered for the input dataset. These polarimetric matrix elements are second order scattering information generated from the spatial averaging of the scattering vector "k " as stated in equation (4.2). All these computed analytical features, reflectance bands and polarimetric matrix elements acted as a pre-cursor for engineering abstract and target specific features using the proposed generative neural network.

$$\mathrm{C}_2 = \begin{bmatrix} \mathrm{C}_{11} & \mathrm{C}_{12} \\ \mathrm{C}_{21} & \mathrm{C}_{22} \end{bmatrix} = \begin{bmatrix} \left\langle |S_{VV}|^2 \right\rangle & \langle S_{VV} S^*_{VH} \rangle \\ \langle S_{VH} S^*_{VV} \rangle & \left\langle |S_{VH}|^2 \right\rangle \end{bmatrix} \tag{4.2}$$

Reference plot sized buffers of radius 15 m from the center of the plot were applied at each plot location to extract mean value from the computed analytical features and a plot level training dataset is prepared. The

missing values (if any) in the dataset were replaced using the median operator and the range of observations in each dataset were standardized using a standard scaler that subtracts the mean and scales to unit variance. The prepared dataset was exported to standard binary file format that stores the shape and information necessary for performing operations with the proposed generative neural network model. In order to reduce the model complexity, Adaptive L1 (A-L1) regularization was applied to the prepared dataset to limit the number of input variables. The entire process was carried out to prepare three separate sets of data based on the data source i.e. MS dataset, SAR dataset and Dual-source (DS) dataset. Finally, a clean, standardized and regularized input dataset was fed to the developed generative neural network.

4.4.3 Model Implementation and Experiments

The developed architecture of the proposed generative neural network is dynamic and depends on the number of analytical features selected post A-L1 regularization. Therefore, the number of network filters for each layer change for each of the three datasets based on the schema of the dynamic architecture shown in Figure 4.4. The developed generative neural network was trained with a batch size of 10 and a total of 500 epochs. The dense layers of the network were L2 regularized to keep the weights and biases small and reduce the likelihood of overfitting. The network was trained by using a K-fold stratified cross-validation method with five folds. The model was implemented using open-source software library "TensorFlow" on a Python API. The training of the network was performed on a 128 GB RAM NVIDIA GeForce RTX 3090 GPU on a Linux based OS.

The experiments were designed to quantify the improvement in AGB predictions and the refinement in the quality of feature space achieved using the proposed model. Multiple models were developed based on dif-

Table 4.3: Validation Statistics for Assessment of Prediction Accuracy, Precision, Agreement and Overfitting of Models

Validation statistics	**Equations**	**Validation Aspect**
Mean Absolute Error	$MAE = \sum_{i=1}^{n} \left\lvert pre_i^{CV} - obs_i \right\rvert / n$	Prediction Accuracy
Root Mean Squared Differences	$RMSD = \sqrt{\frac{SS^{CV}}{n}}$	Prediction Precision
Coefficient of determination (cross-validation)	$R^2_{CV} = 1 - SS^{CV}/SS_{tot}$	Prediction Agreement
R^2 Ratio	$R2R = R^2_{fit}/R^2_{CV}$	Overfitting

n: total number of samples; pre_i^{CV}: prediction value of sample 'i' obtained by cross validation; obs_i: observed value of sample 'i'; SS^{CV}: sum of squared differences between observed and predicted values by cross validation; SS_{tot} : sum of squared differences of each observation from overall mean; R^2_{CV}: Coefficient of determination via Cross-Validation; $R2R$: Ratio of coefficient of determination via residual to cross-validation; R^2_{fit}: Coefficient of Determination via Residuals (without cross validation)

ferent techniques such as linear, bagging, boosting and neural networks to compare the results with the proposed model. In particular, Generalized Linear Model (GLM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and a Multi-layer Perceptron (MLP) are used for comparing model performance. The structure of the MLP model was same as structure of the regressor unit used with the proposed model to quantify the improvement delivered by the generative process in model performance (i.e. improvement achieved by the probabilistic regressor network as compared to the MLP regressor network with same structure). All the comparative models were fine-tuned for optimal performance according to respective requirements using random hyperparameter search (for RF), early stopping (for XGBoost, MLP) and adaptive moment optimization (for MLP). All the comparative models were 5-fold cross validated and trained using A-L1 regularized features.

The model agreement was evaluated using cross validated R-squared score (R^2_{cv}), prediction precision using the Root Mean Squared Percentage Difference ($RMSD$), model overfitting using the R-squared Ratio ($R2R$) and scale-depended prediction accuracy using the Mean Absolute Error

(MAE). The equations of these statistical performance metrics and the respective validation aspect are given in Table 4.3. In order to compare the quality of the feature space, the t-distributed Stochastic Neighbor Embedding (t-SNE) technique was used to visualize the A-L1 regularized analytical features (input for GLM, RF, XGBoost, MLP) and the generative features from latent space (input for the probabilistic regressor of the proposed model) on a 2D plane color-coded with field estimated AGB. The t-SNE visualization were used to analyze the orientation of the latent space with respect to the field estimated AGB. These analytical and generative features were scored using feature importance derived from an independent Gradient Boosting Machine (GBM) algorithm for quantitative assessment of their feature contribution. The GBM based feature importance scores were computed using a standard Permutation Feature Importance algorithm. Finally, we also generated an AGB map of the Trentino region using the best proposed model and performed correlation analysis using 55 additional reference AGB plots for independent site validation.

4.5 Results and Discussion

4.5.1 Analytical Features Selected Post A-L1 Regularization

The list of the selected analytical features post A-L1 regularization is given in Table 4.4. The regularized datasets consisted of 26 features for MS, 11 features for SAR and 33 features for the DS dataset. The 26 features of the regularized MS dataset consisted of 10 features from the spring season, 6 features from the summer season, 7 features from the autumn season and 3 features from the winter season. Thus, the regularization showed different degree of relevance for different seasons for the MS dataset. A major contribution of the red-edge spectrum (9 features) and SWIR spectrum (4 features) was observed in the regularized MS dataset. The 11 features

of the regularized SAR dataset mainly consisted of polarimetric matrix elements (C_{12}, C_{21}) from all seasons, C_{11} from the winter season and PRVI from summer and autumn season. The 33 features of the regularized DS dataset consist of 17 MS features and 16 SAR features.

4.5.2 Predictive Analysis Using Developed Models

The regression scatterplots for the Proposed, MLP, GLM, XGBoost and RF model representing field estimated AGB versus the predicted AGB are shown in Figures 4.6 - 4.10 respectively. The 5-fold cross-validated model performance metrics for model agreement (R^2_{cv}), prediction precision (RMSD%), overfitting (R2R), and prediction bias (MAE) for all developed models and different datasets are given in Table 4.5.

The proposed model delivered best results with respect to all considered performance metrics for all the three datasets. Particularly, the proposed model obtained the best performance on DS data with an agreement score $R^2_{cv} = 0.63$ and least overfitting score R2R = 1.3. The proposed model also delivered the best prediction precision RMSD% = 34.6 and least prediction bias MAE = 73.6 Mg/ha. The proposed model achieved better prediction precision (RMSD% = 34.6 – 38.9) and less overfitting (R2R = 1.3 – 1.4) as compared to MLP (RMSD% = 72.9 – 75.4 and R2R = 1.8 – 1.9) in spite of an identical neural structure of the regressor unit of the proposed model and the MLP. Also, the analytical features had a higher dimensionality (n) compared to the generative features (n-6) but the latter improved model performance metrics for all the three datasets.

With respect to the datasets considered, all examined models performed least accurately on SAR data (RMSD = 38.9 – 75.4 and MAE = 87.1 – 172.7). This result was anticipated as C-band data are characterized by low canopy penetration. However, the proposed model assisted in improving the model predictions with SAR data (RMSD% = 38.9 and MAE = 87.1).

Table 4.4: List of Selected Analytical Features from MS, SAR and DS data post A-L1 Regularization

Selected MS features	Selected SAR features	Selected DS features
1. BWDRVI_march	1. C11_december	1. BWDRVI_march
2. GLI_march	2. C12_december	2. GLI_march
3. normG_march	3. C21_december	3. normG_march
4. VIgreen_march	4. C12_june	4. CIRE3_march
5. CIRE3_march	5. C21_june	5. NDVIre1_march
6. NDVIre1_march	6. PRVI_june	6. NBR2_march
7. NDVIre2_march	7. C12_march	7. GLI_june
8. NDVIre3_march	8. C21_march	8. normG_june
9. NBR1_march	9. PRVI_march	9. VIgreen_june
10. NBR2_march	10. C12_september	10. CCCI1_june
11. GLI_june	11. C21_september	11. NBR2_june
12. normG_june		12. VIgreen_september
13. VIgreen_june		13. RI_september
14. CCCI1_june		14. CCCI1_septemeber
15. CCCI2_june		15. NDVIre1_september
16. NBR2_june		16. NBR2_september
17. GLI_september		17. CIRE3_december
18. normG_september		18. C11_december
19. VIgreen_september		19. C12_december
20. RI_september		20. C21_december
21. CCCI1_september		21. PRVI_december
22. NDVIre1_september		22. C12_june
23. NBR2_september		23. C21_june
24. normR_december		24. PRVI_june
25. VIgreen_december		25. C11_march
26. CIRE3_december		26. C12_march
		27. C21_march
		28. C22_march
		29. PRVI_march
		30. C12_september
		31. C21_september
		32. DOP_september
		33. PRVI_september

Table 4.5: Results of Performance Metrics computed for Proposed Model and Conventional Models (GLM, MLP, RF and XGBoost) developed using MS, SAR and DS data

	MS				**SAR**				**DS**			
Models	R^2_{cv}	RMSD%	R2R	MAE	R^2_{cv}	RMSD%	R2R	MAE	R^2_{cv}	RMSD%	R2R	MAE
Proposed	0.53	37.8	1.3	83.1	0.37	38.9	1.4	87.1	0.63	34.6	1.3	73.6
MLP	0.15	72.9	1.9	162.5	0.08	75.4	1.8	172.7	0.21	74.6	1.9	165.7
GLM	0.44	46.7	1.4	80.9	0.19	51.7	1.8	117.6	0.45	36.9	1.5	82.7
XGBoost	0.34	39.6	2.5	88.2	0.18	53.1	4.3	117.3	0.36	39.6	2.6	91.9
RF	0.44	39.7	1.8	85.3	0.1	47.66	12.4	107.9	0.45	37.16	1.8	83.5

The prediction bias was reduced by the proposed model (MAE = 87.1) as compared to the conventional models despite the intrinsic wavelength dependent limitations of the SAR data. All models performed better with MS data as compared to SAR data but the combined DS data delivered highest model performances in terms of precision (R^2_{cv} = 0.63, RMSD% = 34.6, R2R = 1.3 and MAE = 73.6). Thus, the proposed model demonstrated a successful and more effective approach for seamless data fusion for modelling AGB.

Among the considered tree based models (RF and XGBoost), the RF model achieved better overall agreement (R^2_{cv} = 0.45) than XGBoost (R^2_{cv} = 0.36). Also, XGBoost produced higher overall prediction bias (MAE = 91.9) than RF (MAE = 83.5) for all datasets. Overall, RF model performed better than XGBoost in the tree based model category. Lastly, in the neural network category, the MLP model delivered inaccurate predictions and showed higher overfitting compared to the GLM and the proposed model. The MLP model produced the least agreement for SAR data (R^2_{cv} = 0.08) with overall high prediction errors (MAE > 160). Although, the proposed model provided an improved model performance as compared to the conventional neural network model (MLP) with respect to all assessment metrics. Overall, the proposed generative neural network performed

better than considered conventional models such as GLM and tree-based algorithms.

Figure 4.12 shows the AGB map obtained by the proposed model on the analyzed area. A strong correlation can be observed between the predictions of the proposed model and 55 independent reference data plots (locations shown in Figure 4.12). Figure 4.13 shows the correlation scatter plot and the computed Pearson correlation coefficient (R). The Pearson correlation coefficient for the 55 reference plots and the mapped AGB values was R = 0.66. Thus, a high correlation achieved for these independent reference plots (not used for training/testing models) indicated a robustness of the proposed model on new data or potentially on different sites.

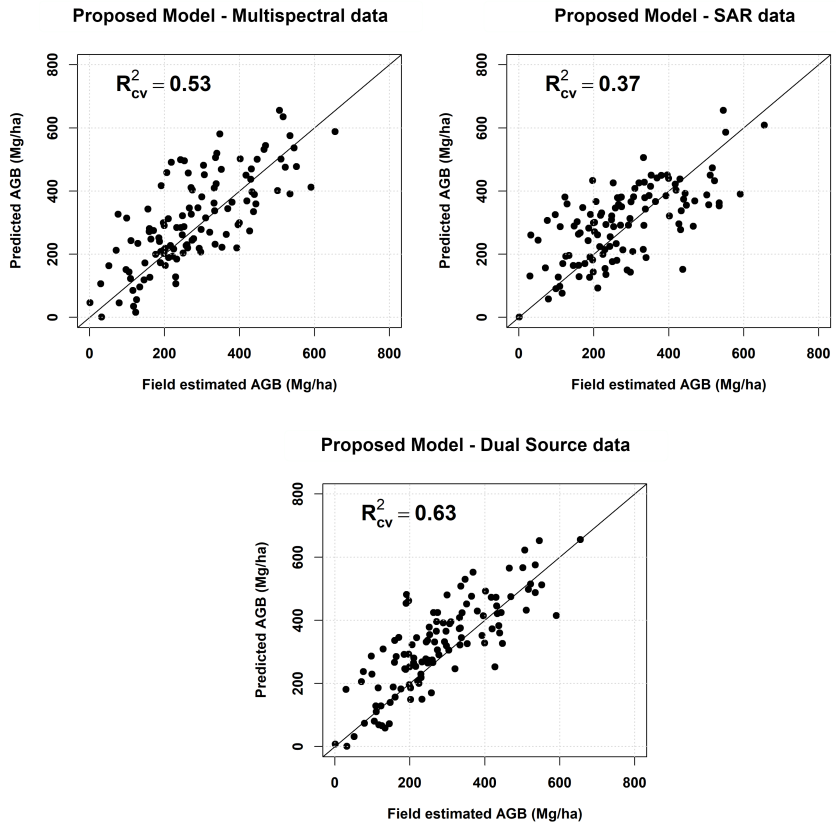

Figure 4.6: Regression scatterplots of the Field Estimated versus Model Predicted AGB for the proposed model using single source and DS data

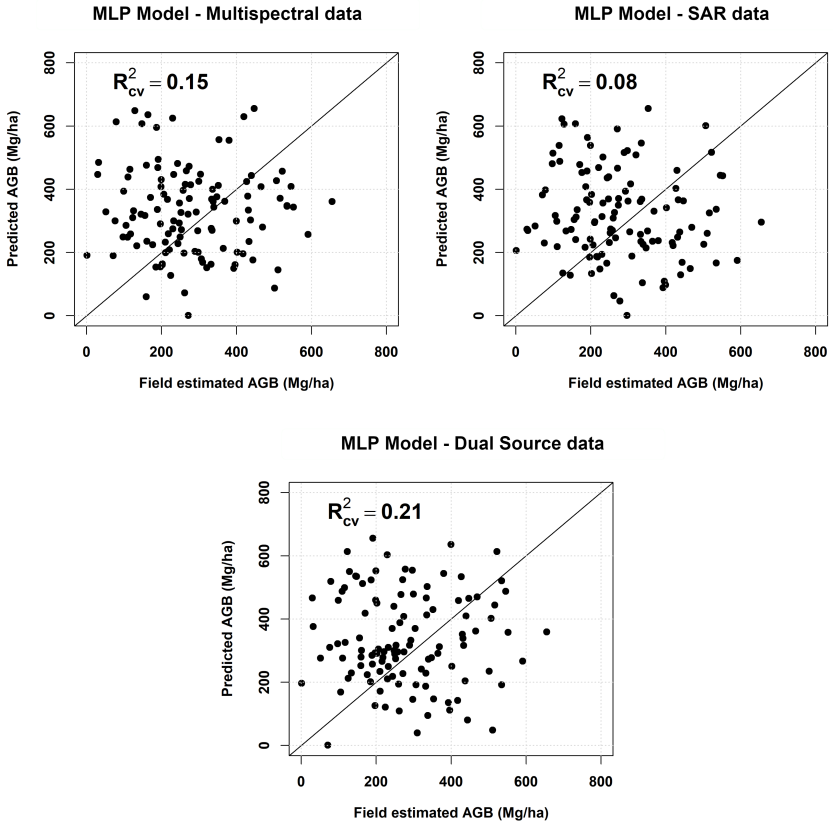

Figure 4.7: Regression scatterplots of the Field Estimated versus Model Predicted AGB for the MLP model using single source and DS data

4.5.3 Two dimensional t-SNE Visualization of Latent Space

The t-SNE scatterplots of generative features and analytical features for MS, SAR and DS datasets are shown in Figures 4.11. The t-SNE scatterplots for all three datasets indicated that the latent space of the proposed model depicting generative features is highly ordered as compared to the input feature space depicting analytical features. The two-dimensional t-SNE visualization of a higher dimensional latent space indicated that the arrangement of the generative features was directional and oriented towards the target AGB. However, the two-dimensional t-SNE visualization of input feature space was non-directional and showed no specific orienta-

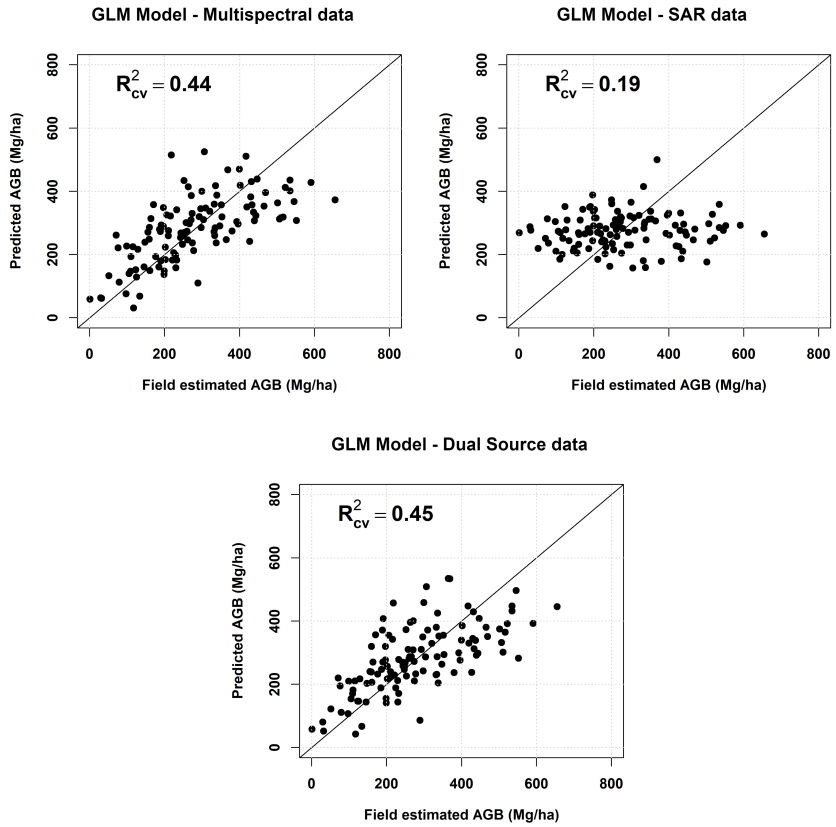

Figure 4.8: Regression scatterplots of the Field Estimated versus Model Predicted AGB for the GLM model using single source and DS data

tion towards AGB. The directional orientation of generative features with respect to target AGB values improved the generalization of the proposed model by decreasing model overfitting and prediction bias simultaneously (R2R = 1.3 – 1.4 and MAE = 73.6 – 83.1).

The t-SNE scatterplots of the input feature space for all datasets (Figures 4.11 – Right) show an un-ordered distribution and produced no AGB associated variations. Although, the t-SNE scatterplots of latent space for all datasets (Figure 4.11 – Left) have one dimension clearly associated with variability of target AGB. Also, a greater degree of target association was observed for MS and DS latent space as compared to SAR latent space. A few observations from the SAR latent space from the range of 0 to 100

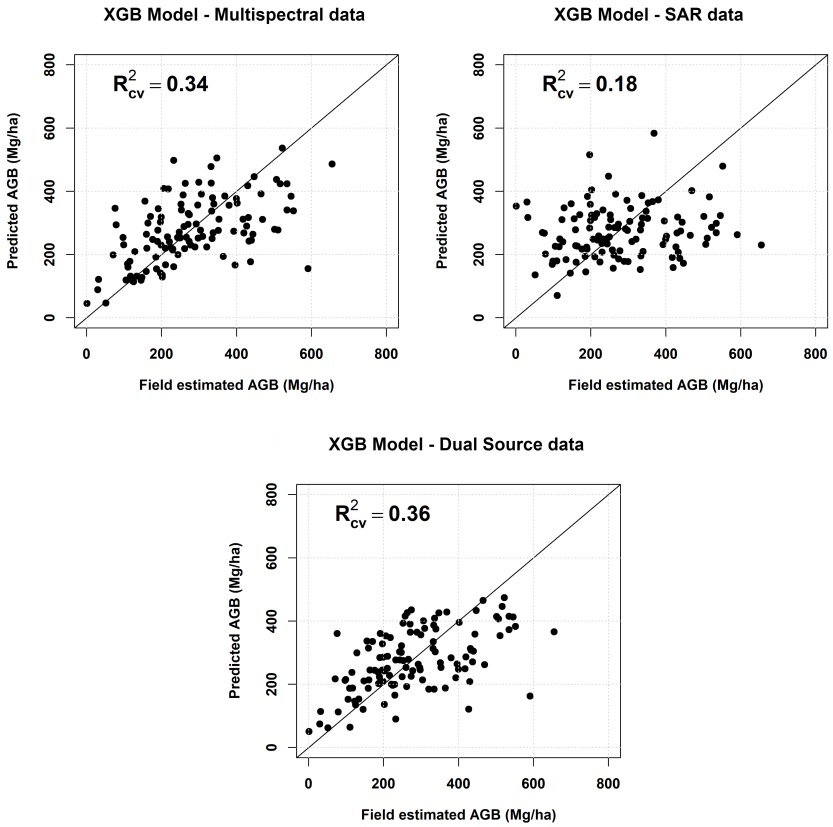

Figure 4.9: Regression scatterplots of the Field Estimated versus Model Predicted AGB for the XGB model using single source and DS data

Mg/ha were at the point (x, y) = (0,0) that had higher frequency of observations from mid-range AGB (300 to 400 Mg/ha). These disassociated observations explained the inferior results for SAR data as compared to the other data using the proposed model.

The plots of scaled feature importance versus the ten most important analytical and generative features are shown in Figure 4.14. The slope of the feature importance trend line for analytical features is greater than that for generative features. The relative feature importance of analytical features scaled down at higher rate as compared to generative features. This is because only a few among all the analytical features significantly contributed to the accurate prediction of AGB. However, in case of generative

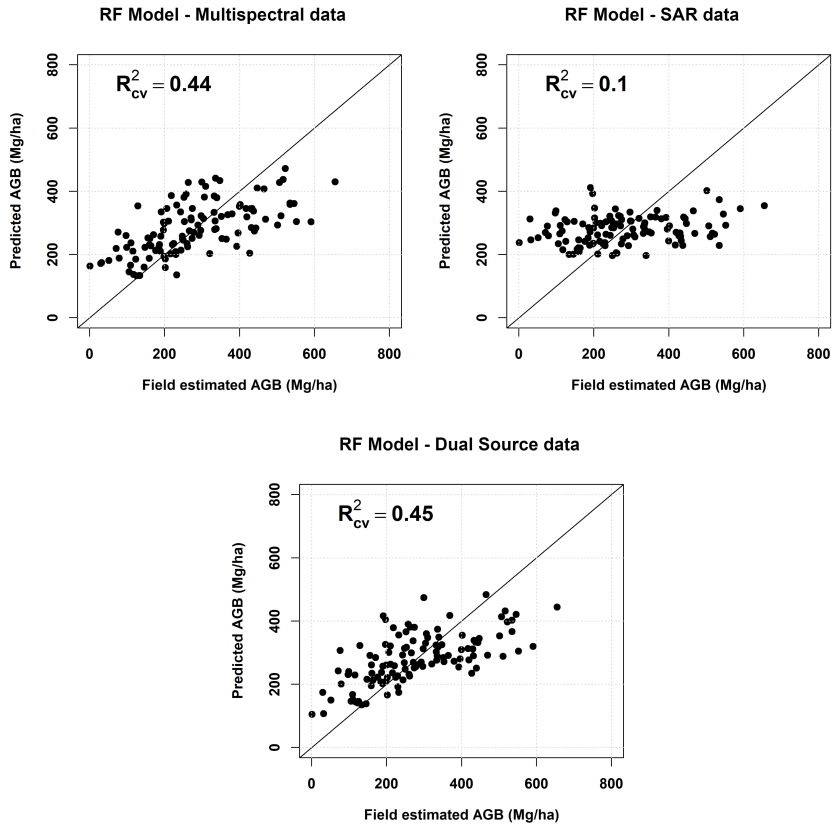

Figure 4.10: Regression scatterplots of the Field Estimated versus Model Predicted AGB for the RF model using single source and DS data

features, a greater number of features provided significant contribution (see Figure 4.14) for an accurate AGB prediction as compared to the analytical features. This indicated that the overall feature importance of generative features was higher as compared to that of analytical features. This also explains the role of generative features in delivering better prediction results from the performed experiments.

4.5.4 Discussion

In this thesis contribution,a generative neural network was proposed with a dynamic architecture that has been used on SMS and SAR remote sensing

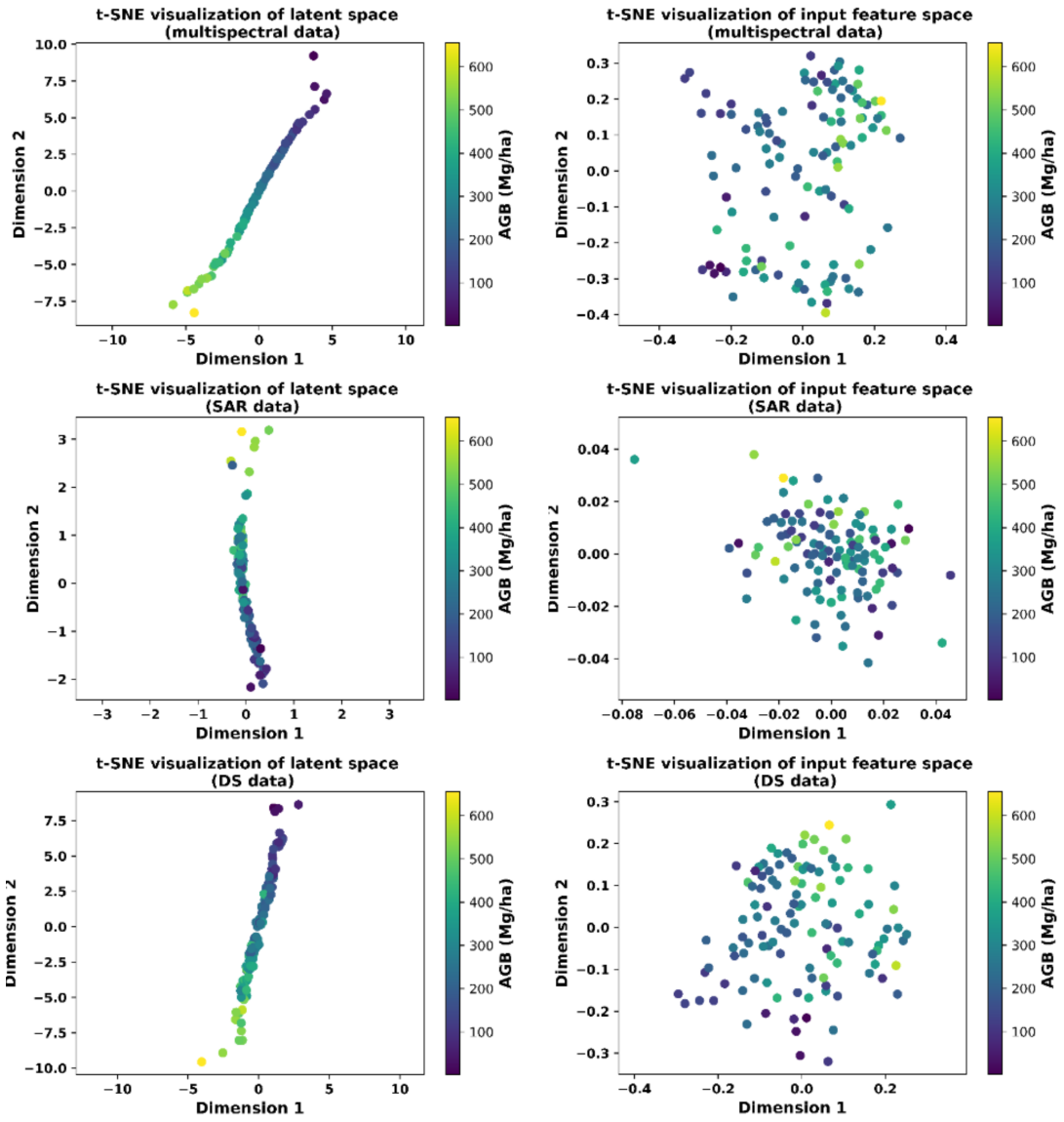

Figure 4.11: Color-coded t-SNE plots of reduced dimensionality latent space and input feature space of MS, SAR and DS data

data for modelling plot level forest AGB. The key elements of modelling consisted of operationalizing the triple loss mechanism, generating target oriented features and using a probabilistic regressor to deliver predictions. This framework was proposed for dealing with the various issues identified in the literature related to the use of satellite remote sensing data for forest AGB mapping. The results achieved from the performed experiments have

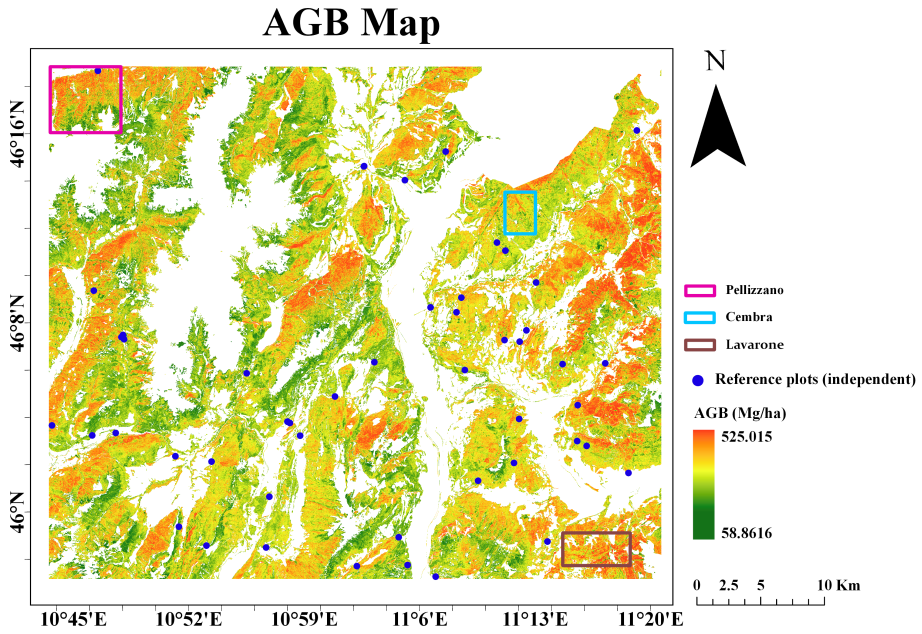

Figure 4.12: AGB Map of the study area produced using the Best Proposed Model and location of the independent reference plots

been analyzed with respect to the contemporary literature in this section.

Studies that used satellite remote sensing data for AGB prediction highlighted a common drawback in terms of data saturation and low prediction precision [113, 25, 155]. This contribution analyzed multiple models for different datasets and found two effective ways to reduce the bias and data saturation. The first proposed way is fusion of multi-sensor data as this increases the number of features (and hence the information) that can reduce saturation for predicting higher AGB values. The second proposed way is by extracting targeted features. In case the multi-sensor data are not available, engineering features to induce targeted properties in the feature space can also reduce the saturation and increase prediction precision with satellite remote sensing data. In this contribution, it was quantitatively proven that the two stated ways are effective for increasing prediction precision and reducing saturation. The metric RMSD% was significantly better for the proposed approach indicating the increased prediction preci-

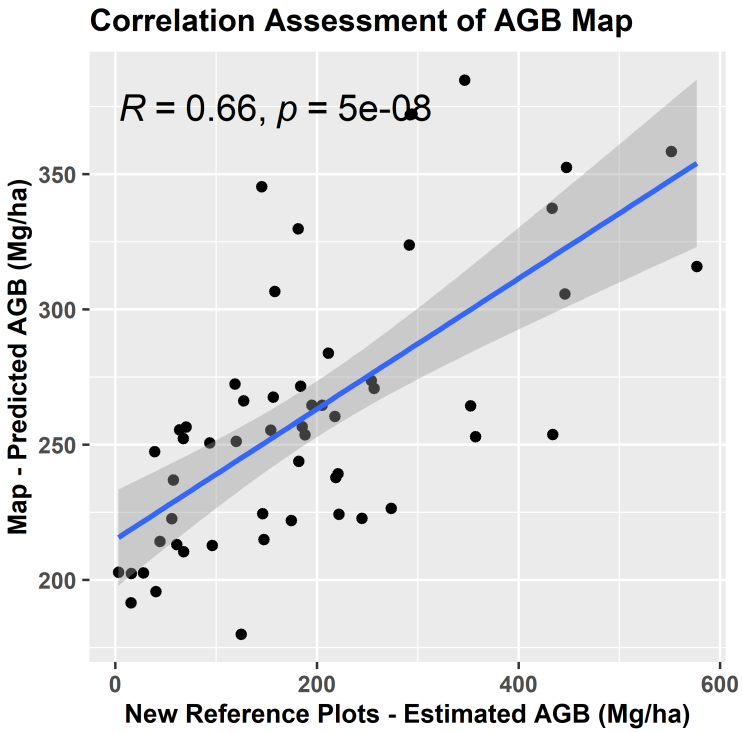

Figure 4.13: Correlation assessment of the AGB predicted using the Best Proposed Model and Field Estimated AGB for independent reference plots

sion. Moreover, a better concurrence achieved by the proposed model for a higher range of AGB values indicated reduced model saturation. This was quantitatively reflected by a higher r-squared score and can be graphically observed with regression scatterplots that show greater agreement for higher AGB values.

The analysis performed in [113] suggested calibration of the model within the range of AGB values to minimize the risk of induced prediction bias. Forest are diverse in terms of species, density, type and distribution. Thus, developing a method that can effectively learn from multi-modal data and engineer features to orient with respect to the AGB values is one of the viable solution. This aspect was presented in the form of a probabilistic regressor network that tracks the label loss and coordinates with the latent generator to produce AGB-oriented features reducing the need

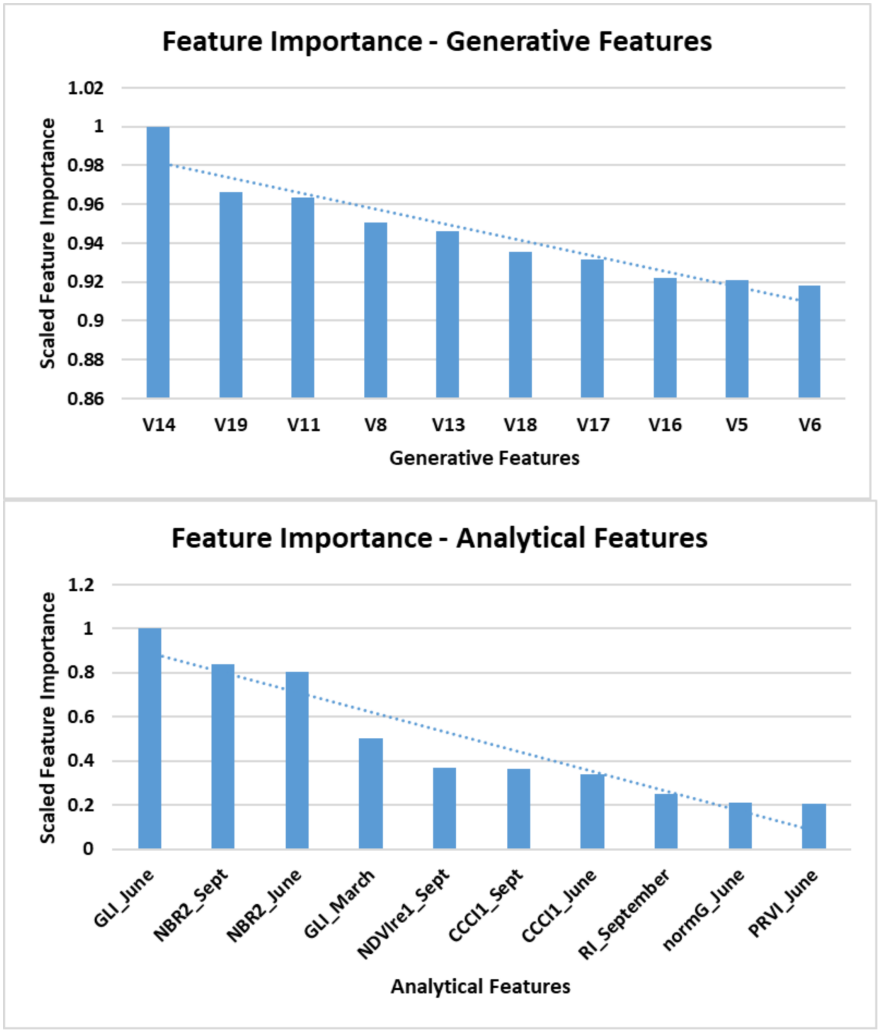

Figure 4.14: Feature Importance graph computed using an independent Gradient Boosting Machine algorithm for Generative and Analytical features

for AGB-oriented sensors. In addition to regularization and dimensionality reduction of the data, this contribution showcased a disentanglement procedure that orients one dimension of the features to display target specific variance and thereby improving the performance of the model. Moreover, the fusion of optical and SAR data for modelling forest AGB is also a challenging task to achieve the desirable modelling results. A simple additive combination of dual-source data (used for the four conventional models

- GLM, MLP, RF and XGBoost used for comparison) delivered inferior results as the models utilize only primary level information to learn two different properties. The proposed model extracts high-level information and transforms the feature space such that it exhibits target oriented properties.

The regression algorithm used for modelling predominantly affects the prediction accuracy of the model. Multiple comparative studies that used different machine learning regression algorithms for the prediction of forest AGB from satellite RS data produced diverse results [25, 151, 156]. The comparison of these studies in literature for determining the best suitable regression algorithm is difficult due to different initial conditions (data, samples, features). The number of samples used and the type of features selected can also affect the performance of a model. Thus, the same model can perform differently by changing the initial conditions. Accordingly, the identification of the baseline model remains an issue for AGB modelling. In simple terms, it is difficult to have one baseline model that can perform best for all initial conditions. Also, the study in [151] compared a few standard machine learning algorithms and a stacked sparse autoencoder (feed forward neural network) to prove the superiority of the latter. However, the study had used random splitting of dataset (3:1) and our study that used stratified cross-validation found that feed forward neural networks are highly prone to overfitting. This problem was resolved by the idea of a robust and integrated probabilistic regressor that encourages the posterior to resemble to the AGB specific prior that can reduce the prediction bias induced by the initial conditions. This can be specifically observed from Figure 3.13, where the model was applied to additional field plot samples that were not used during training of the model. The correlation plot showed that prediction bias slightly increases for extreme values of AGB but the model overall retained a strong correlation between the predicted

and field estimated AGB. This shows that the proposed approach can be effectively replicated on new data or potentially new sites.

Another important aspect of this contribution has been to automate the process of balancing the bias-variance trade-off of the prediction algorithm. In the study [157], eight machine learning models were evaluated for the prediction of forest AGB using satellite remote sensing data. It outlined the problem of stabilizing the prediction bias with change in the forest types, sampling methods and dependence of feature importance on the deployed model. It used thirteen features for modelling and observed a high variance in the importance of the same feature for different models. It also highlighted that the process of hyperparameter optimization for each model and selection of the best model was time consuming. In this contribution of the thesis, the entire process was automated for achieving optimal results with a less complex and time efficient computations.

Overall, this contribution effectively deals with various problems identified in the literature by providing a robust solution in form of a generative modelling architecture. However, a prime limitation of the proposed model is that it used a neural network based regressor unit that cannot be replaced by any other contemporary regression algorithm (e.g. RF or XGBoost). The generative architecture requires a neural network unit to update weights and produce targeted generative features. Thus, it reduces the scope of testing other regression algorithms with the proposed architecture. Moreover, the architecture of the proposed model is complex with separate inference and generative parameters. Therefore, it is difficult to use a neural architecture search algorithm to optimize the number of filters and layers of the model. This limits the strategy of determining the most optimal architecture of the proposed model.

4.6 Conclusion

This contribution of the thesis has proposed a generative approach for modelling forest AGB using satellite RS data. The results demonstrated the superiority of generative features over conventional analytical features extracted from satellite RS data for AGB prediction. The proposed dynamic architecture and the triple loss mechanism generate target specific features that showed improvement for all prediction metrics (agreement, precision, overfitting and accuracy) in the conducted experiments. Moreover, the proposed model also demonstrated its effectiveness for efficient feature fusion and compression. It was conclusive from experiments that it is difficult to completely eliminate the factor of data saturation but the proposed model substantially reduced it thereby increasing the reliability of satellite RS data for AGB prediction. The extended work requires testing the proposed framework to predict other forest biophysical variables and parameters from a different RS application. Moreover, the aspect of model saturation can be studied in more detail for all such biophysical variables including AGB using the proposed features. The proposed framework could also be modified and used for classification tasks such as tree species or land-use land-cover classification.

Chapter 5

Automated Machine Learning for Stacked Ensemble Model Development for Forest AGB Prediction

5.1 Introduction

The selection of suitable ML algorithms to solve the forest AGB problem requires domain expertise for improving and regulating model performances [16, 15, 160, 28, 34]. The task of creating an optimized ML pipeline for AGB prediction is complex and time-consuming due to diversity in forest type, distribution and species. The traditional approaches proposed for the development of ML pipelines use a trial and error mechanism for stacking and selection of models for AGB prediction [161, 162, 130]. In practice, the human element cannot be completely eliminated from the pipeline development process for performing model checks and deriving model explanation

Part of the chapter appears in:

P. Naik, M. Dalponte, and L. Bruzzone, "Automated Machine Learning Driven Stacked Ensemble Modelling for Forest Aboveground Biomass Prediction Using Multitemporal Sentinel-2 Data," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. Institute of Electrical and Electronics Engineers (IEEE), pp. 1–14, 2022. DOI: 10.1109/JSTARS.2022.3232583

with respect to the target parameter i.e. AGB. However, the human intervention can be reduced and replaced with efficient hyperparameter search and importance learning algorithms. This minimizes the number of user regulated parameters and achieve faster and efficient modelling with reduced systematic error [163, 164, 165, 166]. In this regard, the concept of Automated Machine Learning (AutoML) can be very instrumental in gaining better machine learning performance with the available computational budget and reduced human assistance to model forest AGB. AutoML automates knowledge intensive tasks such as configuring learning tools for feature engineering, architecture search and algorithm selection using an optimization-evaluation mechanism [167]. In a general architecture, the AutoML controller consists of an evaluator and an optimizer. The evaluator measures the performance of learning tools and provides feedback to the optimizer in order to update configurations for better performance. The optimizer generates configurations based on a search space that is determined by a process of targeted learning [168]. For example, if the learning process is "feature engineering", the learning tools to be configured are the "classifiers" and the search space would consist of "feature sets, feature enhancing methods (dimension reduction, feature generation, feature encoding) and related hyperparameters".

There are a wide ranging applications of AutoML such as medical image recognition [169, 170], object detection [171, 172, 173], super resolution [174, 175, 176], language modelling [177], text classification [178], semantic segmentation [179], etc. AutoML has delivered quality performance for various tasks but scarcely experimented for developing a ML pipeline for AGB prediction. A ML pipeline is a directed graph of learning elements that can be automated by using search of estimators/predictors, search of learning algorithms and search of ensemble-models [180, 181, 182]. In literature, there are various algorithms developed for automation of these learning ele-

ments of a ML pipeline [183]. The evolutionary algorithms such as Particle Swarm Model Selection (PSMS) use Particle Swarm Optimization (PSO) to automate the full model selection problem [184]. Such evolutionary algorithms inspired Ensemble PSMS [185] which is the pre-cursor to the development of the latest AutoML systems [186]. Recent AutoML systems are capable of building regression/classification pipelines, full model selection, multi-objective optimization and best architecture/hyperparameter search for deep learning models [187]. Auto-WEKA [188] system uses the Sequential Model-Based Algorithm Configuration (SMAC) method [189] which is a robust stochastic optimization framework under noisy function evaluations for lowest cross-validation misclassification. AutoSklearn [190] which is a system similar to Sequential Model Based Optimization (SMBO) has a distinctive search process initialization that exploits a meta-learner delivered best performance in many AutoML challenges. The Tree-Based Pipeline Optimization Tool (TPOT) [191] is an open source genetic AutoML system that optimizes a series of ML models for high-accuracy and compact pipelines for supervised classification. The recent surge in the application of deep neural networks led to the development of AutoML systems like Auto-Net [192] and Neural Architecture Search (NAS) [193, 194] which are built on combination of Bayesian optimization and Hyperband to automatically tune deep neural networks without human intervention.

The concept of meta-learning is integral to AutoML systems and enables them to learn from the meta-data of the learning elements [195]. AutoML systems perform faster and efficiently on a new task with meta-learning techniques that replace human-engineered ML pipelines with data-driven pipelines [196]. Meta-learning techniques have been successfully implemented in literature to automate various learning elements (feature engineering, architecture search, hyperparameter optimization) of the AutoML systems [180, 197, 198, 199, 200, 201]. In this context, there are many stud-

ies conducted using meta-learning techniques to execute various ML tasks in remote sensing applications. A trivial problem in deploying advanced ML algorithms in the RS domain is few-shot (low training samples) learning for which meta-learning techniques have provided some significant solutions [202, 203, 204, 205]. Meta-learning techniques can efficiently handle multi-scale RS data and provide efficient solutions to inversion modelling problems [206, 207]. These techniques have been used to solve more specific RS problems such as the study in [208] specifically focused on deep neural network model for semantic segmentation of circular objects in satellite images. Apart from such specific problems, meta-learning techniques are capable of dealing with combination of such problems to solve larger problems in RS. For example, a recent study [209] aimed at generating captions for RS images using meta-learning – a task that requires dealing with a combination of problems based on visual and textual features to generate RS image captions. Thus, meta-learning approaches have a significant contribution in solving complex RS problems and holds a great potential for improving AGB modelling in diverse scenarios such as mixed tree species, forest types and varying density of forests.

The possibility is unlikely for a single modelling algorithm to perform effectively on most types of AGB modelling scenarios. Ensemble models are a combination of base models built on the hypothesis that the combination of multiple (weak) models can produce a more reliable and accurate model as compared to the individual base models [210]. Ensemble models are developed based on techniques such as Bagging, Boosting and Stacking. The bagging and boosting models both involve homogeneous weak learners that are combined by a deterministic strategy. The difference is bagging algorithms follow a parallel learning strategy (e.g. RF) and boosting algorithms follow a sequential learning strategy (e.g. AdaBoost, XGBoost). Differently, stacking involves heterogeneous weak learners that

are combined using a meta-model with a parallel learning strategy. Ensemble models have been used for AGB modelling and prediction by various studies in literature [130, 69, 25]. However, developing a stacked ensemble with manual trial and error approaches to model AGB is an inefficient and time-consuming task. Meta-learning driven AutoML systems can be adequately used for developing optimal stacked ensemble models for AGB prediction [211]. AutoML systems use a collection of base models with hyperparameter tuning algorithms for producing stacked ensemble models [212]. The AutoGluon system [213] uses a multi-layer stacked ensembling with K-fold bagging that stacks models in multiple layers and trains in layer-wise manner. The AutoSklearn system [214] employs a Bayesian optimization algorithm for searching through the hyperparameter space and meta-learning for warm-starting of the search procedure. The AutoSklearn 2.0 [190] which is an improvement upon Autosklearn, uses portfolio learning after performing Bayesian optimization. Among the most recent AutoML systems, "H2O" AutoML [215] performs stacked ensembling with random forest, gradient boosting machines, linear models and deep learning models using a super learner algorithm.

The primary objective of this contribution is the prediction of AGB using Multitemporal-Multispectral (MT-MS) satellite remote sensing data. Prediction of AGB using SRS data requires a robust ML pipeline to deal with outliers in the training data and increase the generalization ability of the model. Moreover, SRS is generally used for large-scale AGB mapping and may require testing multiple models due to the spatially varying characteristics of forest. Thus, it is required to have a comprehensive evaluation framework for the multiple models that are considered. Therefore, an AutoML system is proposed for the prediction of forest AGB that enables training and evaluation of models within a single framework. In particular, a meta-learning driven AutoML system is used that automates selection

and stacking of candidate base learners for modelling AGB. Additionally, a SMBO procedure is incorporated to automatically extract features from the MT-MS satellite data to minimizes systematic error in the proposed ML pipeline.

5.2 Study Area and Datasets

5.2.1 Study Area Description and Field Data

The study area is the province of Trento (6216 km^2) situated in the north-eastern Italy in the southern part of the Alps. The area is mountainous and has a 60% of forest cover mostly owned by public institutions that are subject to broad goals of forest management (i.e. forest protection, species biodiversity, carbon storage etc.). The area has mountains with high elevations and landlocked valleys. The details regarding the tree species in the area are described in Chapters 3 and 4 (Study area section). The geographical location of the study area and the distribution of the field plots in the study area are shown in Figure 5.1. The field data consists of 315 circular plots with a fixed radius of 15 m (Figure 5.1). These data include 98 broadleaf plots, 152 coniferous plots and 65 mixed plots. The total 315 plots were divided into three categories (broadleaf, coniferous and mixed plot) such that more than 80% of the plot AGB should be derived from one particular tree-type (broadleaf or coniferous), otherwise the plot was considered as a mixed plot. The details of survey devices, measurement parameters and species specific allometric equations is given in Chapter 3 (Field data sub-section). The field estimated plot level AGB values ranged from 1.07 $Mg\ ha^{-}1$ to 711.41 $Mg\ ha^{-}1$.

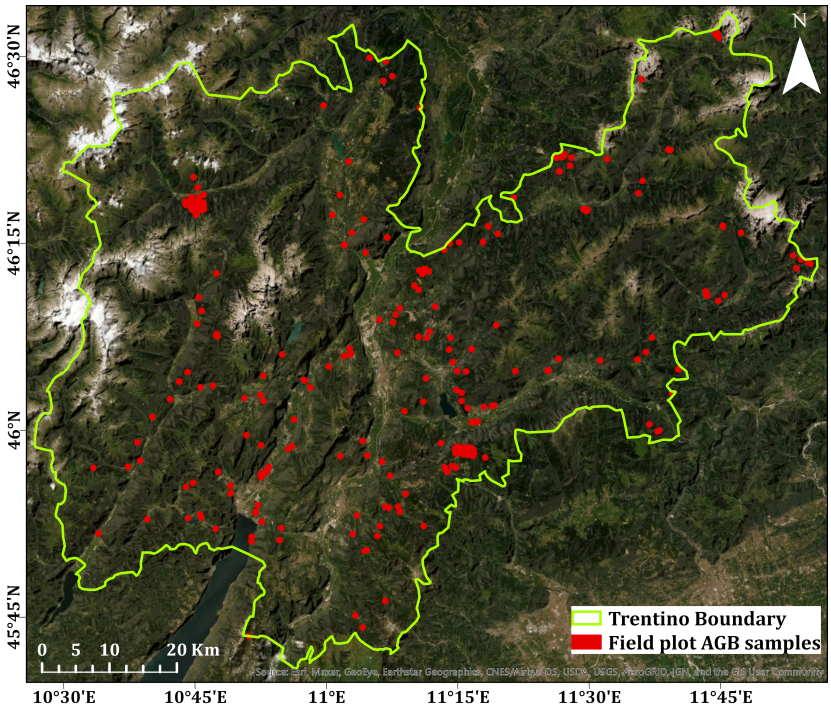

Figure 5.1: Map of the Study Area with locations of the reference AGB plots indicated as red dots

5.2.2 Remote Sensing Data

The study was performed using MT-MS Sentinel-2 images. The acquisition dates of the Sentinel-2 images are stated in Table 5.1. The specifications of Sentinel-2 images are given in from Chapter 3 (Remote sensing data sub-section).

5.3 Description of Algorithms

The developed mechanism for achieving the proposed objectives was based on coupling two algorithms: 1) Tree structured Parzen Estimator (TPE) [216] which is a hyperparameter optimization algorithm based on an iterative search process on a dynamic search space and 2) Super Learner (SL) [217] which is a V-fold cross validation based meta-learning algorithm that

Table 5.1: Acquisition dates (YYYY-MM-DD) of the Sentinel-2 images acquired over the four different seasons

Seasons	**Sentinel-2**
Spring	2016-03-20
Summer	2016-07-18
Autumn	2016-10-16
Winter	2016-12-15

selects weights for combining candidate models. The following sub-sections (5.3.1 and 5.3.2) provide a description of the two algorithms.

5.3.1 Tree Structured Parzen Estimator Algorithm for Automated Feature Extraction

The TPE algorithm is based on a SMBO approach that overcomes the limitations of computationally expensive and inefficient random search and grid search algorithms. The inputs required by the TPE algorithm are parameters (x) and loss (y) based on the prior search history to deduce hyperparameters for the next trial. The input pair (parameters and loss) is split into two densities ($\ell(x)$ and $g(x)$) based on the loss of the historical data by a γ quantile (Equation 5.1 and 5.2) of the search result. The TPE defines this split $p(x \mid y)$ using the two densities according to the following equations:

$$p(x \mid y) = \begin{cases} \ell(x) & \text{if } y < y^* \\ g(x) & \text{if } y \geq y^* \end{cases} \tag{5.1}$$

and

$$\gamma = p\left(y < y^*\right) = \int_{-\infty}^{y^*} p(y)dy \tag{5.2}$$

where, $\ell(x)$ should be maximized as it represents the density formed from the observations $x^{(i)}$ that correspond to the loss (y) smaller than the target performance (y^*), whereas $g(x)$ should be minimized as it represents the density formed by using the remaining observations. Therefore, the output of the TPE algorithm can be simply characterized as the ratio between $g(x)$ and $\ell(x)$.

The idea of the TPE algorithm is based on maximizing the Expected Improvement $EI_{y^*}(x)$ computed by (Equation 5.3) based on convergence of loss (y) and target performance (y^*) in subsequent trials:

$$EI_{y^*}(x) = \int_{-\infty}^{y^*} \max\left(y^* - y, 0\right) p(y \mid x) dy = \int_{-\infty}^{y^*} \left(y^* - y\right) \frac{p(x \mid y)p(y)}{p(x)} dy \tag{5.3}$$

Where, p(x) can be written as:

$$p(x) = \int_R p(x \mid y)p(y)dy = \gamma\ell(x) + (1-\gamma)g(x) \tag{5.4}$$

Therefore, we can deduce:

$$\begin{aligned} EI_{y^*}(x) &= \int_{-\infty}^{y^*} \left(y^* - y\right) p(x \mid y)p(y)dy \\ &= \ell(x) \int_{-\infty}^{y^*} \left(y^* - y\right) p(y)dy \\ &= \gamma y^* \ell(x) - \ell(x) \int_{-\infty}^{y^*} p(y)dy \end{aligned} \tag{5.5}$$

Finally combining (5.4) and (5.5) we obtain:

$$EI_{y^*}(x) = \left(\gamma + \frac{g(x)}{\ell(\boldsymbol{x})}(1-\gamma)\right)^{-1} \cdot \left(\gamma y^* - \int_{-\infty}^{y^*} p(y)ydy\right) \propto \left(\gamma + \frac{g(x)}{\ell(\boldsymbol{x})}(1-\gamma)\right)^{-1} \tag{5.6}$$

The expression in equation (5.6) is a product of two terms of which the second term is independent of 'x' and hence the expression is directly proportional to the first term and implies that to maximize improvement

points 'x' with high probability under $\ell(x)$ and low probability under $g(x)$ are desirable. Also, this expression shows that the ratio $\frac{(g(x)}{\ell(x)}$ should me minimized in order to increase the expected improvement.

5.3.2 Super Learner Algorithm for Automated Ensemble Modelling

The SL algorithm can be used for ensemble modelling by considering a set of diverse modelling algorithms as base learners and selecting an optimal ensemble through V-fold cross-validation. A library $\mathcal{L}$ of 'K' modelling algorithms is defined to solve a regression problem of estimating the expectation of target 'B' given the observed variable 'A'. The optimal value of the parameter of interest '$\psi_0(A)$' that is required to be estimated and the related loss function '$L(O, \psi)$' that indicates the difference between the observed and the predicted value (squared error loss) can be written as,

$$\psi_0(A) = E(B \mid A) \tag{5.7}$$

$$L(O, \psi) = (B - \psi(A))^2 \tag{5.8}$$

Thus, the optimal values of the parameter of interest for the 'K' individual algorithms included in library $\mathcal{L}$ can be written as '$\widehat{\psi}_k(A)$'where k = 1, 2,. . . .,K. The number of algorithms 'K' to be considered in the library is dependent on the sample size of the data. Note that all the considered algorithms estimate the same parameter '$\widehat{\psi}_k(A)$' but may use different subsets of 'A', different basis function, estimation procedures and range of tuning parameters. The identification of the best predicting algorithm from the library over the data distribution (P_o) is determined by minimizing the expected risk difference 'd_n', i.e.,

$$d_n\left(\widehat{\psi}_k, \psi_0\right) = \int \left\{ L\left(B, \widehat{\psi}_k\right) - L\left(B, \psi_0\right) \right\} dP_o \tag{5.9}$$

$$\tilde{k}_n = \arg\min_k \left\{ d_n\left(\widehat{\psi}_k, \psi_0\right) \right\} \tag{5.10}$$

The use of the same data to estimate '$\widehat{\psi}_k(A)$' and '$d_n\left(\widehat{\psi}_k, \psi_0\right)$' generates bias in the estimation of the true risk to determine the best algorithm. Thus a V-fold cross-validation selector is used for an unbiased estimation of the risk. Therefore, given the empirical distributions for the training and validation sets for each V-fold, the cross-validation selector can be defined as $\hat{k}$,

$$\hat{k} = \arg\min_k \left\{ \mathrm{E}_{C_n} \sum_{i:C_n(i)=1} L\left(B_i, \hat{\psi}_k\left(A_i \mid P^0_{n,C_n}\right)\right) \right\} \tag{5.11}$$

where, P^0_{n,C_n} and P^1_{n,C_n} are the distributions of training set ($C_n(i) = 0$) and validation set ($C_n(i) = 1$) respectively.

5.4 Proposed Approach

The flowchart of the proposed approach given in Figure 5.2 shows the process of integrated implementation of TPE and SL algorithms to create a ML pipeline for AGB prediction. In the following sub-sections, the pre-processing of Sentinel-2 data, implementation of TPE and SL algorithms and model explanation parameters are described in detail.

5.4.1 Data Pre-Processing

The Sentinel-2 images were acquired in Level-1C (Top of the atmosphere reflectance) format and converted to Level-2A (Bottom of the atmosphere reflectance) format with atmospheric and terrain correction using the Sen2cor

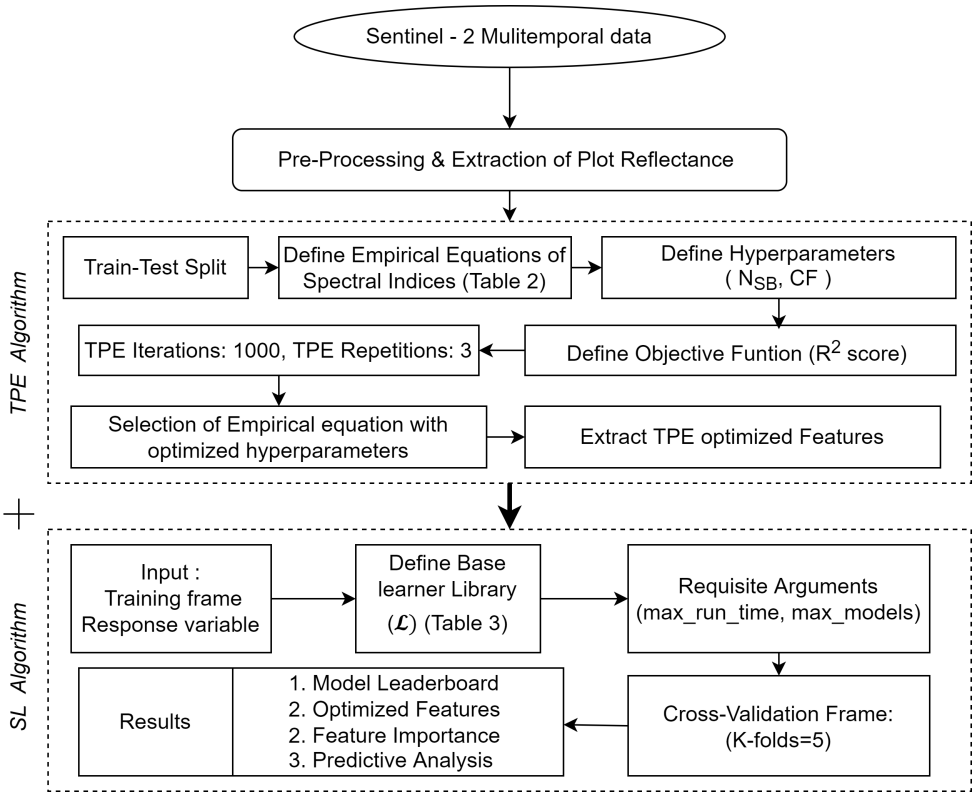

Figure 5.2: A flowchart of the Proposed Approach demonstrating sequential implementation of the TPE and SL algorithms for AGB prediction

processor [218]. The spectral bands at 20 m spatial resolution were resampled at 10 m for spatial consistency and performing computations. The resampled data were used for extracting plot reflectances to produce season-wise analysis ready data frames using 'rgeos' and 'rgdal' packages of R software. Particularly, the 'readOGR' and 'rgeos' functions were used for reading the plot points and creating a buffer of 30 m diameter around plot points, respectively. The analysis ready data frames were prepared using 'extract' function (Raster package of R) with 'mean' argument to compute average of pixels covered by the buffer around plot points. The pixels are included by the function if and only if the plot boundary covers the centroid of the pixels. This procedure was carried out uniformly on all seasonal images. The analysis ready data frames of each season consisted

of plot reflectance values for each spectral band for all sample plots.

5.4.2 Implementation of Algorithms

The two core frameworks of the proposed approach are based on TPE and SL algorithms implemented sequentially to produce a robust ML pipeline for AGB prediction. Firstly, the TPE algorithm generates automatically optimized features from the pre-processed analysis ready data frames. These optimized features and the response variable are used for training the SL algorithm that automates the process of training a large number of base models and performs stacked ensembling to produce a leaderboard of models. The TPE algorithm was implemented using "Optuna" framework that was developed in [219] and the Super Learner algorithm was implemented using "H2O-3" framework that was developed in [215]. The sequential implementation of the two frameworks is explained in detail in the following paragraphs.

The implementation of the TPE algorithm to extract features from the input spectral bands requires a model, i.e. an empirical equation for which it can generate a set of parameters and select a combination of optimal spectral bands to accurately predict the target variable. The considered empirical equations are given in Table 5.2. The library of 33 empirical equations (indexed as I_n) framed from an exhaustive database of 500+ spectral indices [220] were availed from [221]. The TPE randomly selects an empirical equation and generates parameters depending on the number of spectral bands (N_{SB}) and the coefficients ($CF = \alpha, \beta, \gamma, \rho$ and σ) in the equation. A generalized linear regression model is fitted to select the optimal spectral bands and coefficient values to maximize the coefficient of determination (objective function). The coefficient of determination enables to quantify the proportion of the variance in extracted optimized feature with respect to the target AGB values. The coefficient of determi-

Table 5.2: List of 33 Empirical Equations of spectral indices categorized depending on the number of spectral bands used for the TPE algorithm

Number of Spectral Bands (N_{SB} **= 2-6**)	**Index No.** (I_n **= 1-33**)	**Empirical Equations**
Two	1	$(B1 - B2)/(B1 + B2)$
	2	$(B1/B2)$
	3	$B1 - B2$
	4	$\alpha + \beta(B1 - B2)/(B1 + B2)$
	5	$(B1/B2) - \alpha$
	6	$log(1/B1) - log(1/B2)$
	7	$log(1/B1) + log(1/B2)$
	8	$\alpha[(B1 - \alpha * B2 - \beta)/$ $(\alpha * B1 + B2 - \alpha + \rho)]$
	9	$B1 - \alpha[B2/(B1 + \beta * B2)]$
	10	$[(B1/B2) - 1]/sqrt[(B1/B2) + 1]$
Three	11	$(B1 - B2)/(B3 - B2)$
	12	$\alpha(\beta(B1 + B2)–B3)$
	13	$B1(B2/B3^2)$
	14	$2 * B1 - B2 - B3)/(2 * B1 + B2 + B3)$
	15	$[(B1 - B2)/(B1 + B2)]/[(B1 - B3)/(B1 + B3)]$
	16	$(B1 - B2 - \alpha(B2 - B3))/(B1 + B2 - \alpha(B2 - B3))$
	17	$(B1 - B2)/B3$
	18	$(B1 - B2 + B3)/(B1 + B2 + B3)$
	19	$B1/(B1 + B2 + B3)$
	20	$arctan([(\alpha * B1 - B2 - B3)/\beta][B2 - B3])$
	21	$\alpha[\beta(B1 - B2) - \rho(B3 - B2)]/$ $sqrt[(2 * B1 + 1)^2 - (\gamma * B1 - \sigma * sqrt(B3)) - \alpha/3]$
Four	22	$(B1/B2)/[(B3 - B4)/(B3 + B4)]$
	23	$B1/(B2 + B3 + B4)$
	24	$[(B1 + B2) - B3]/(B4 - B3)$
	25	$[(B1 - B2) - \alpha(B1 - B3)(B1/B2)]/$ $[(1 + \alpha)(B4 - B2)/(B4 + B2 + \alpha)]$
Five	26	$[(B1 - B2) - \alpha(B1 - B3)(B1/B2)]/$ $[\alpha[\beta(B4 - B3) - \rho(B2 - B3)/$ $sqrt((B4 + 1)^2 - (B5 - \rho * 2 * sqrt(B2)) - 0.5)]]$
	27	$B1/(B2 + B3 + B4 + B5)$
	28	$(B1 - B2)/(B2 + B3 + B4 + B5)$
	29	$(B1 + B2) - B3)/(B4 - B5)$
Six	30	$(B1 * B2 * B3)/(B4 * B5 * B6)$
	31	$B1/(B2 + B3 + B4 + B5 + B6)$
	32	$(B1 - B2)/(B2 + B3 + B4 + B5 + B6)$
	33	$[(B1 - B2) - \alpha(B1 - B3)(B1/B2)]/$ $[(1 + \alpha)(B4 - B2)/(B4 + B2 + B5 + B6 + \alpha)]$

Table 5.3: List of Base models considered for the SL algorithm provided by H2O AI Cloud Platform

Models	Count	Type
eXtreme Gradient Boosting Machine (XGBoost)	3	Pre-specified
Generalized Linear Model (GLM)	1	Fixed Grid
Distributed Random Forest (DRF)	1	Default
Gradient Boosting Machine (GBM)	5	Pre-specified
Deep Neural Net (DNN)	1	Near-default
eXtreme Gradient Boosting Machine (XGBoost)	1	Random Grid
Gradient Boosting Machine (GBM)	1	Random Grid
Deep Neural Net (DNN)	1	Random Grid

nation should be maximized to determine the best fit in terms of spectral bands and coefficient values for the empirical equations. The TPE initialization provides a pair of parameters and loss as stated in sub-section 5.3.1 that are split into two densities ($l(x)$ and $g(x)$) as per equations (5.1) and (5.2). The ratio of the two densities should be minimized (thus maximizing the objective function (R^2) in our case) through an iterative process and the best empirical equation is determined with the related optimum set of parameters (spectral bands, coefficient value). However, the random initialization creates a selection bias in the TPE algorithm that favors the empirical equations with less number of parameters i.e. empirical equations involving less spectral bands. This selection bias problem was resolved using a meta-learning strategy by performing optimization process in groups. In total, the empirical equations were divided in five groups based on the number of spectral bands (N_{SB} = 2, 3, 4, 5, 6) and 1000 iterations were performed for each group. The best empirical equation in each group was identified and after performing three repetitions of this process, the overall best empirical equation was identified. The selected empirical equation, optimal spectral bands and coefficient values were used to compute the features and directed to the subsequent framework.

The SL framework receives the computed TPE optimized features and the target response variable as input training frame. As previously mentioned, we used the H2O-3 platform to deploy the SL algorithm which consists of a library of base models stated in Table 5.3. Additional details regarding the specifications of the base models and model parameters can be accessed from H2O.ai documentation (`https://docs.h2o.ai/`). The SL algorithm exploits the base models and derives an optimal Stacked Ensemble Model (SEM) that minimizes the expected risk difference as per equations (5.10) and (5.11). In order to achieve the defined objective of this contribution, all the available base models were used in the library (Table 5.3). All base models were trained until their convergence with 5-fold cross validation on a 134 GB NVIDIA GeForce GTX 3090 GPU and Linux based operating system. However, the maximum number of models and maximum runtime for models can be specified before the training process depending on the available computational budget and time restrictions.

5.4.3 Model Explanations

The process of sequential deployment of TPE and SL frameworks followed in the proposed approach results in TPE optimized features, model rankings, model predictions and prediction assessment metrics. Model explanations such as model rankings, evaluation statistics, regression scatterplots and feature importance chart are provided as results. The performance of the base models and the SEM were evaluated on the basis of different evaluation metrics and ranked according to the model agreement score (coefficient of determination). All models were cross-validated using 5-fold cross-validation method. The metrics used to evaluate these models are coefficient of determination (R^2_{cv}), Root Mean Squared Error ($RMSE$), Root Mean Squared Log Error ($RMSLE$), Mean Absolute Error (MAE), Mean Absolute Percentage Error ($MAPE$) and Relative Absolute Error

(RAE). The R^2_{cv} shows the goodness-of-fit of the regression model, the $RMSE$ shows the standard deviation of the residuals, $RMSLE$ shows the log-transformed standard deviation of residuals, MAE shows the mean of absolute values of residuals, $MAPE$ shows the accuracy of prediction as a percentage and RAE shows accuracy of measurements relative to the range of the observed variable.

5.5 Results and Discussion

5.5.1 TPE Optimized Features and Feature Importance

The features for the SL framework were extracted based on the hyperparameter optimization results of the TPE algorithm. The empirical equations with respective spectral bands and coefficient values for each season that were selected post TPE optimization procedure are given in Table 5.4. The selected empirical equations have been referenced using Index numbers (I_n) mentioned in Table 5.2. The spectral bands selected for the respective equations have been referenced with respect to their central wavelengths (in nm). The corresponding Sentinel-2 spectral band for each central wavelength can be identified from Copernicus website (`https://sentinels.copernicus.eu/`). The empirical equations that had multiple selected combinations of spectral bands were referenced as "a", "b" and "c".

In total, 24 extracted features were selected for training the base models for AGB prediction following the TPE based optimization procedure. Particularly, there were six features from autumn, seven from spring, six from summer and five from winter seasons. Interestingly, all extracted features consisted of spectral bands from the Vegetation Red-Edge (VRE) and SWIR spectrum of the Sentinel-2 data. The TPE algorithm that was conditional on the target variable for the optimization process suggested an

Table 5.4: TPE Optimized Features for each season computed as per Empirical Equations Indexed as I_n stated in Table 5.2

Index	**Spectral Band**	**Season**
(I_n)	**(Central Wavelength (λ_c) in nm)**	**(aut, spr, sum, win)**
17	842, 1610, 740	autumn (aut)
23	842, 1610 740, 2190	autumn
28	a = 1610, 842, 665, 740, 740 b = 1610, 842, 705, 740, 740 c = 1610, 842, 740, 2190, 665	autumn
32	842, 1610, 665, 740, 783, 740	autumn
16	a = 705, 842, 2190 ($\alpha = 1$) b= 2190, 842, 705 ($\alpha = 1$)	spring (spr)
23	a = 842, 705, 665, 1610 b = 842, 665, 1610, 740	spring
27	842, 665, 1610, 705, 740	spring
26	842, 1610, 665, 705, 705	spring
32	842, 705, 2190, 1610, 665, 1610	spring
7	1610, 705	summer (sum)
8	a = 705, 1610 ($\alpha = 0.4, \beta = 0.7, \rho = 0.9$) b = 665, 1610 ($\alpha = 0.6, \beta = 0.6, \rho = 0.9$)	summer
12	a = 705, 490, 665 ($\alpha = 0.8, \beta = 1$) b = 865, 665, 1610 ($\alpha = 0.03, \beta = 0.1$)	summer
33	1610, 665, 490, 865, 560, 560 ($\alpha = 0.01$)	summer
23	842, 705, 1610, 1610	winter (win)
27	842, 1610, 1610, 1610, 705	winter
28	a = 842, 1610, 2190, 665, 1610 b = 842, 1610, 705, 665, 1610	winter
32	842, 1610, 665, 1610, 705, 1610	winter

important role of the VRE and SWIR spectral bands for modelling forest AGB. Precisely, the SWIR band at $\lambda_c = 1610$ were selected for 21 out of the 24 extracted features and the contribution of VRE was distributed at $\lambda_c = 705$, 740 and 865. The season-wise analysis indicated that VRE at $\lambda_c = 740$ and SWIR at $\lambda_c = 1610$ was selected for all autumn features. The spring features were extracted using VRE at $\lambda_c = 705$ selected for six and out of the total seven features. Also, SWIR at $\lambda_c = 1610$ was selected for five features and $\lambda_c = 2190$ for the remaining two features. For the summer season, VRE at $\lambda_c = 705$ and 865 were selected for five out of total six features and SWIR at $\lambda_c = 1610$ for all features. The winter features had VRE at $\lambda_c = 705$ and SWIR at $\lambda_c = 1610$ selected for all extracted features. Thus, the clear observable pattern in the specific spectral contributions of the extracted features account for their target-oriented properties.

The stacked bar chart in Figure 5.3 shows the computed feature importance for all the considered base models. The features have been referenced as 'index_season' and the scaled feature importance values of the base models have been color coded for analyzing the percentage contributions. The highest average feature importance was observed for '12a_sum' (79%) that was extracted with $\lambda_c = 705$, 490 and 665. The feature achieved specific feature importance of 74% for DNN, 24% for DRF and 100% for GBM, XGBoost and GLM. This indicated that feature '12_sum' contributed significantly for most of the base models and has a dominant role in modelling AGB. A total of five features for DNN (26_spr, 12a_sum, 8b_sum, 8a_sum, 12b_sum), two features for DRF (28c_aut, 33_sum), five features for GBM (12a_sum, 32_aut, 23_aut, 28b_aut, 7_sum), three features for XGBoost (12a_sum, 7_sum, 12b_sum) and five features for GLM (12a_sum, 7_sum, 33_sum, 23_aut, 17_aut) were characterized by feature importance score greater than 50%. The feature '12a_sum' achieved highest feature importance for three (GBM, XGBoost and GLM) out of the five base models and

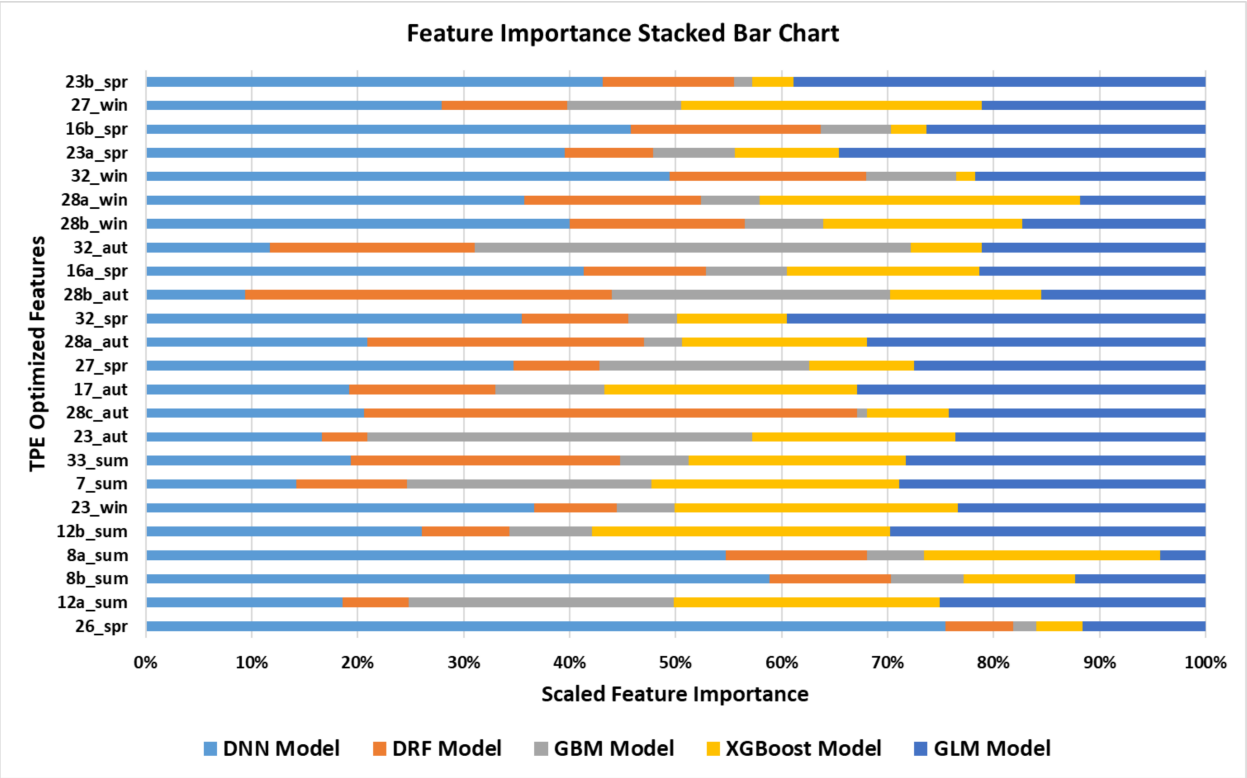

Figure 5.3: Stacked Bar Chart of Scaled Feature Importance computed for TPE Optimized Features for all Base Models

the features '26_spr' and '28b_aut' achieved highest feature importance for DNN and DRF respectively. All these features were associated with either summer, autumn or spring season indicating comparatively low importance of winter features in predicting AGB. Overall, the summer features were the most dominant with a total of 12 out of 20 features with a feature importance greater than 50%.

5.5.2 AutoML Leaderboard and Predictive Analysis

The H2O-3 based AutoML framework used to implement the SL algorithm trains all the base models (Table 5.3) and also produces an optimal SEM for predicting AGB.

The model ranking results and computed assessment metrics indicated

Table 5.5: Model Leaderboard rankings for SEM and Base Models with six Assessment Metrics

Models	**Rank**	R^2cv	**RMSE** ($Mgha^-1$)	**RMSLE**	**MAE** ($Mgha^-1$)	**MAPE**	**RAE**
SEM	1	0.71	74.44	0.56	48.58	0.30	0.43
DNN	2	0.66	81.27	0.56	48.48	0.30	0.43
DRF	3	0.65	82.01	0.56	51.17	0.36	0.45
GBM	4	0.50	98.49	0.65	74.28	0.38	0.66
XGBoost	5	0.48	99.95	0.64	73.70	0.39	0.65
GLM	6	0.38	109.44	0.72	84.94	0.42	0.75

that SEM achieved the best overall performance for predicting forest AGB. The SEM model achieved highest agreement of $R^2_{cv} = 0.71$ and best prediction precision $RMSE = 74.44$ Mg ha-1. The $RMSLE = 0.56$ was equal for the top three ranked models despite a significant difference in the $RMSE$ indicated the existence of outliers in the data. This explained the slightly higher MAE of the SEM despite achieving better model agreement as compared to DNN and DRF. The same explanation is applicable to the DNN and DRF that have a significant difference in the MAE and $MAPE$ despite of an identical $RMSLE$ score. This suggests that SEM is more robust to outliers and the meta-learning process enabled the model to achieve better model fitting. The DNN model achieved the second best performance on the leaderboard and the overall results suggested DNN to be more prone to outliers as compared to SEM. The bagging based DRF model achieved third position on the leaderboard with $R^2_{cv} = 0.65$ and $RMSE = 82.01$Mg ha-1. The next two positions were achieved by boosting based GBM and XGBoost models. Both GBM and XGBoost had an identical performance and GBM outperformed XGBoost by a small margin. The GBM and XGBoost model achieved $R^2_{cv} = 0.5, 0.48$ and $RMSE = 98.49, 99.95$ Mg ha-1 respectively. The GLM model achieved the last position on leaderboard with$R^2_{cv} = 0.38$ and $RMSE = 109.44$ Mg ha-1.

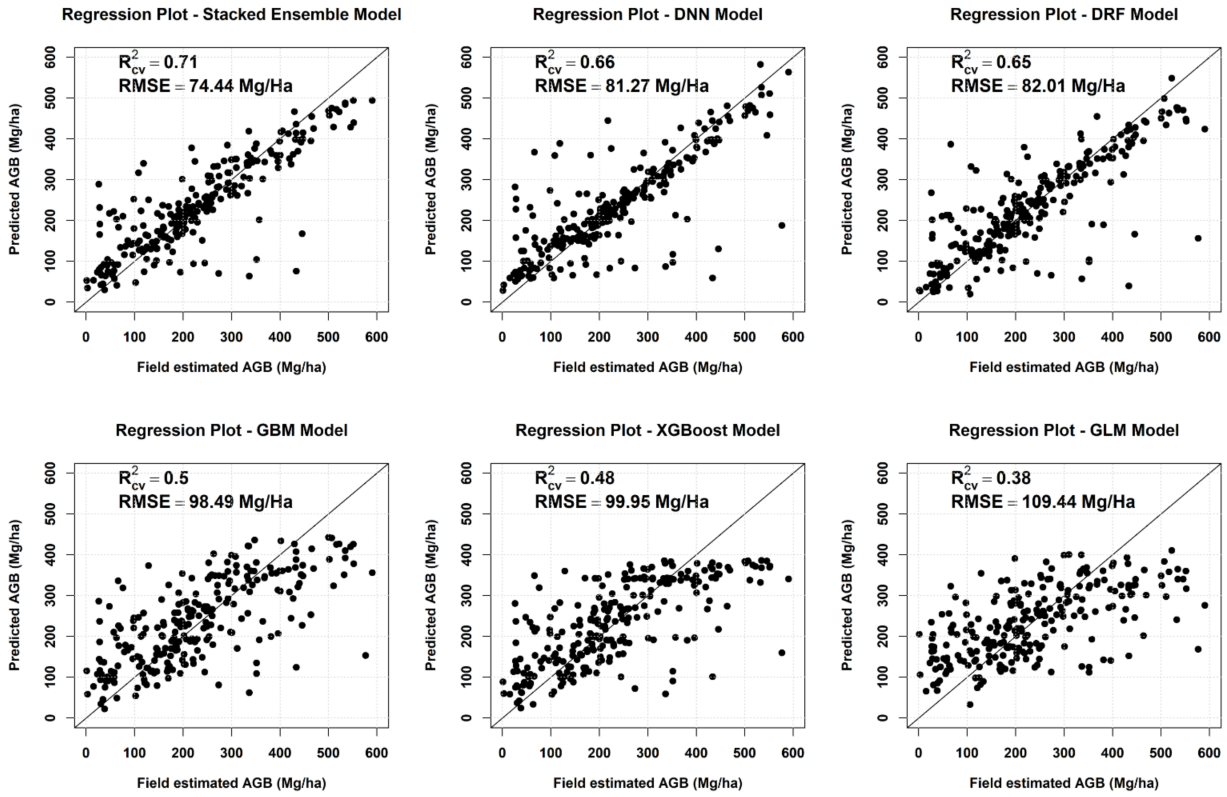

Figure 5.4: Regression scatterplots of Field Estimated AGB versus Predicted AGB for SEM and Base Models

The *RAE* score for SEM model was 0.43 and successively increased to 0.75 for other models on the leaderboard. This indicated lowest saturation from SEM and successive increment in saturation for other models to predict AGB. The low *RAE* score for SEM and DNN models show that these models can predict greater AGB values with better accuracy as compared to the other models. This is also evident from the scatterplots of the regression models shown in Figure 5.4. The scatterplots clearly indicate that SEM, DNN and DRF models efficiently predicted large values of AGB as compared to GBM, XGBoost and GLM models. Thus, the proposed automated ML pipeline yielded SEM that outperformed single base models and efficiently modelled AGB. The results also pointed out a crucial role

of meta-learning process in modelling AGB for eliminating uncertainties associated with the data and producing robust models.

The forest plot-type based analysis of model performance was carried out by independently computing $RMSE$ for broadleaves, coniferous and mixed plots. The results (shown in Table 5.6) indicate that the SEM model (which showed best overall performance) was sensitive to the type of forest plot. Overall, the least $RMSE$ errors were observed for the mixed-type plots and highest $RMSE$ errors were recorded for broadleaf plots. This type of a behavior was consistent with the results from previous studies that obtained more accurate AGB estimations results for conifers as compared to the broadleaves using spectral data [228, 229]. Broadleaves have multiple canopy layers and a more complex structure that possess a challenge at a fundamental level with passive spectral data. Therefore, the SEM model showed least $RMSE$ error for coniferous type plot. It also delivered a comparable performance for the mixed type plots ($RMSE$ = 48.38 Mg ha-1), where the DRF model achieved $RMSE$ score of 47.75 Mg ha-1. All base models secured better $RMSE$ scores as compared to SEM for broadleaf type. But, it is important to note that all these models were not trained independently on each distinct type of plot separately. This showed that the SEM model is capable of learning critical features and generalizing them for training data with multiple types of plots. This enabled SEM to achieve best overall performance in spite of underperforming for a particular type of plot (i.e. broadleaf).

5.5.3 Discussion

In this thesis contribution, an automated ML pipeline for developing a SEM was proposed for prediction of forest AGB using MT Sentinel-2 data. The key elements of the ML pipeline were the TPE and SL algorithms that automated the process of feature extraction from the data and training a

Table 5.6: Forest plot-type (Broadleaves, Coniferous and Mixed) based Prediction RMSE for developed models

Models	Broadleaves RMSE ($Mgha^{-}1$)	Coniferous RMSE ($Mgha^{-}1$)	Mixed RMSE ($Mgha^{-}1$)
SEM	112.55	58.59	48.38
DNN	121.21	63.26	57.22
DRF	118.34	69.71	47.75
GBM	111.41	101.04	66.11
XGBoost	113.00	103.77	61.19
GLM	105.72	119.85	76.25

library of base models leading to development of a stacked ensemble for modelling AGB. The automated ML pipeline was proposed for dealing with various issues identified in the literature related to the systematic error in AGB modelling and systematic evaluation of models. In this section, an extensive analysis of the results with respect to the contemporary literature is provided by precisely identifying the advancements delivered by this contribution for AGB modelling using SRS data.

The choice of features is very crucial in any modelling process and the performance of the models is highly dependent on their quality. The use of MS data for AGB modelling has led to the development and testing of various features. Practically, there are numerous combination of spectral bands that could be used for extracting a feature from MS data for modelling AGB. Studies in literature use a few standard vegetation indices as features for modelling and mapping forest AGB [113, 222, 63, 114]. A comparative analysis of these studies indicates an unstable response of vegetation indices depending upon the spatial resolution, sensor specifications and available spectral bands of the data. To state simply, a vegetation index identified as effective for a particular area or a certain radiometric specification may not be as effective for other study areas or radiometric specifications. Thus, the human-intelligence based selection of these vege-

tation indices can be highly inefficient and ambiguous for AGB modelling. In order to overcome this issue, this contribution proposed an automated mechanism for extracting such features. The TPE algorithm based optimization procedure extracts features that are highly target-oriented and involves less human intervention. It is capable of overcoming the shortcomings identified in the literature with regard to extraction and selection of effective features from SMS data for AGB modelling. The reduction of systematic error from the process enables the development of more robust and reliable features. Moreover, it performs necessary changes in composing the features based on the specifications of the data such as spatial, spectral and radiometric resolution. Thus, this contribution demonstrated the success of the proposed automated approach that led to consistent and accurate modelling results.

The second component of the modelling process after developing robust features is the choice of a modelling algorithm. The type of modelling algorithm substantially affects the precision and accuracy of the predictions. As per literature, there are several studies that use different ML models for predicting forest AGB [162, 69, 25] and provides a comparative assessment of the models in order to identify the best modelling algorithm for AGB prediction. Apart from the range of ML models that can be used, each model has associated hyperparameters that require tuning. A faulty hyperparameter tuning can adversely affect the performance of an ML model and restrict its generalization capability. Moreover, studies that use same ML models but with a different combination of hyperparameters or architecture sometimes lead to contrasting performance on an identical task. For example, the study in [25] identified XGBoost to deliver best performance as compared to the other competing models. However, the XGBoost model has multiple associated hyperparameters that were defined manually for finding the best combination of hyperparameters using a grid search.

This introduces a systematic error in the process and reduces the chances for reproducing the results for other data or scenarios. To deal with this problem, this contribution proposed the use of an AutoML approach that effectively automates the iterative tasks associated with the development of a model. Precisely, it automates the process of hyperparameter selection and a range of ML models are trained in the same pipeline for effective comparison and reproducible results.

Studies have identified that a combination of models (stacked ensemble) can produce more efficient results as compared to a single model [223, 224]. There are only a few studies that focused on using SEM's for remote sensing based forest applications [161, 225]. These studies used a manual or semi-automatic approach for identifying an optimal combination of models for AGB prediction. An optimal SEM requires a library of diverse models and a systematic algorithm to evaluate the combination of models with optimized hyperparameters and generate meaningful explanation of models. The solution for this complex problem was provided by the proposed use of SL algorithm implemented using H2O-3 framework that produced a stacked ensemble of base models. The developed SEM model delivered the best performance as compared to the individual base models. Moreover, it limited the required number of user-defined parameters reducing systematic error in the ensemble selection. Thus, a reliable and automated pipeline for robust stacked ensembling and model training was established in this study.

5.6 Conclusion

This contribution proposed an end-to-end ML pipeline for modelling forest AGB using MT-MS satellite remote sensing data. The results demonstrated that by reducing the systematic error from the modelling process

and deploying a comprehensive model evaluation strategy under a single framework can provide better model explanations. The derived model explanations can be instrumental to frame effective schemes for accurately mapping AGB on large areas with diverse forest characteristics. Moreover, instead of using pre-defined features for modelling, an automated optimization procedure can produce more effective features by weighting more on the spectrum of the data that holds greater importance in explaining the target AGB. The future developments can possibly aim at improving the robustness of the proposed pipeline by addition of optimization elements and by replacing or improving the deployed meta-learning strategies. The advances in latest AutoML systems can be possibly incorporated to derive additional model explanations for framing better modelling schemes for large scale AGB mapping.

Chapter 6

Conclusion and Future Developments

The thesis presents systematic series of experiments and advanced methods developed for prediction of forest AGB using SRS data. The developed methods and performed experiments for AGB prediction are categorized into the three distinct contributions. The first contribution of the thesis consists of the development of a procedure based on spectral features from MS images to perform evaluation of their response to AGB considering their different temporal, spectral and spatial properties. Oracle procedures were used for simultaneous estimation, regularization and feature selection for assessment of temporal, spectral and spatial properties of the data. The second contribution focused on generation of robust features from SRS data for modelling forest AGB. A variational autoencoder based model architecture was proposed for generating AGB oriented features from SMS and SAR data and dual-source feature fusion was demonstrated. Lastly, the third contribution dealt with selection and optimization of meta-algorithm and automatic model ensembling using the concept of AutoML for modelling AGB. A combination of hyperparameter optimization algorithms and meta-learning procedures were implemented to create an end-to-end pipeline for AGB modelling. The proposed methods and experiments suc-

cessfully achieved the three main objectives of the thesis presented as three main contributions for resolving the stated research problems.

The first contribution showed that the spectral features used for modelling AGB are influenced by each distinct property of the MS data. The Red-Edge and SWIR bands have a disproportionate contribution as compared to other bands in increasing the confidence of spectral features extracted from the MS data. Thus, Sentinel-2 features (with Red-edge and SWIR capacity) induced least model saturation, whereas the Dove features (w/o Red-edge and SWIR capacity) resulted in highest model saturation. With respect to the temporal property, the single-time FS-2 spectral features had significant disparity in agreement and prediction precision depending on the season of data acquisition. On the contrary, the MT FS-2 spectral features produced better model agreement and delivered greater prediction precisions for all MS data. This indicate that the temporal property of the MS data has a significant role for reliable AGB modelling. The Pearson correlation of key spectral features determined using adaptive regularization were analyzed across all available spatial resolutions (i.e. at 10, 5 and 3 m). The marginally higher correlation of key spectral features at 10 m as compared to 3 m may be a result of overlapping pixels at border with the plot boundary. This difference in Pearson correlation of key spectral variables is not significant for 10 m and 5 m. Overall, the temporal properties of SRS data and the extended spectral properties of Sentinel-2 above NIR wavelength contributed to significant improvements in AGB prediction by restricting saturation. It would be worth noting that the achieved model performance can be further enhanced by considering an additional source of data for fusion with MS source. Also, developing a procedure that can improve the quality of the extracted features and induce AGB oriented properties can further enhance model performance.

The second contribution of the thesis generated target oriented features

for modelling AGB using the developed generative neural network. The t-SNE plot of generative features was highly directional and ordered, i.e. there is an association between a particular range of generative features and a particular range of AGB values. This process resulted in a proposed model with superior results compared to the other tested models (MLP, GLM, XGBoost and RF) that used conventional analytical features. The generative network fused MS and SAR features at the bottle-neck of the network (latent space). The fused generative features provided the best modelling results in terms of agreement, precision, accuracy and overfitting. The fused generative features were more effective compared to the combined analytical features as a result of their highly ordered nature. However, the developed generative model requires an integral regressor network and it cannot be replaced with other type of regression algorithms (e.g., tree-based or kernel based) as it requires backpropagation of weights. Further, there is a requirement of automation that deploys a search algorithm to optimize the architecture and network parameters.

The third contribution of the thesis presents an automated machine learning pipeline that performs auto-feature generation and auto-ensemble selection of base models. The stacked ensemble model produced from the developed pipeline delivered best AGB modelling results on the scale of all assessment metrics. This supported the notion that a single algorithm may be inefficient for modelling on complex and diverse features. The produced SEM pointed out that an ensemble of models is more effective in scenarios that involved trees of different species and data acquired in different seasons. Moreover, the automated procedure (TPE and SL algorithms) assured derivation of features and regulation of model parameters with minimum systematic error. The effect of seasonal data and the dominant spectral bands could be more effectively determined from the developed pipeline.

All three contributions were sequentially organized to achieve the overall aim of effective and robust modelling of AGB. The contributions were linked to overcome certain limitations of the previous contribution and thus improved the qualitative and quantitative results. The results from first contribution indicated that all characteristics of the MS data have a distinct impact on the model performance. The statistical evaluation concluded that spectral features based on VRE and SWIR bands were more effective in improving the overall modelling accuracy. A coherent conclusion was deduced from the third contribution that derived greater number of spectral features with VRE and SWIR bands using an automated TPE framework. The first contribution used pre-defined features from literature based on the Red-edge and SWIR bands such NDVIre, CCCI, CIRE and NBR that delivered improved results for all assessment metrics. The third contribution used an automated framework for defining features observed greater contribution of bands with central wavelength of 705, 740, 783, 1610 and 2190 nm. Both the contributions mutually assert that VRE and SWIR based spectral features had an instrumental role in AGB prediction. The seasonal effects were prominent and produced high variance in the model performance. MT information served as an acceptable solution to deal with the variance and saturation problem of the single-time models and the approach was extended for the other two contributions. The third contribution identified dominant role of summer, autumn and spring seasons in providing adequate MT information to explain the variance in AGB. The first contribution shows that spectral features are both positively and negatively correlated with AGB. The magnitude of correlation had certain variance with respect to the spatial resolution. This produced a un-ordered feature space that limits the performance of model in predicting AGB. This problem was resolved in the second contribution that proposes a network to produce a highly ordered feature space delivering

better modelling results as compared to models developed on un-ordered feature space. In the second contribution, a MS-SAR feature fusion mechanism was proposed that reduced the data saturation level (observed in first contribution) to further improve overall accuracy of AGB prediction. The model developed in the second contribution was tested on independent reference samples distributed across different geographical areas that delivered accurate AGB predictions. These results provide a positive development with respect to the findings in the first contribution that showed inconsistent results for the two different study areas and diverse characteristics of forest. Thus, the generative process resulted in increasing the robustness of models delivering consistent performance for diverse terrain and target properties. The training of model parameters and hyperparameter tuning of ML models are critical for obtaining optimal prediction results. The second contribution considered four different ML models for comparison with the proposed model. However, a uniform framework for training model parameters and tuning hyperparameters can alter the performance and ranking of the models. The AutoML driven framework and ensemble learning approach proposed in the third contribution provided more reliable comparison and evaluation mechanism for different models. Thus, all contributions effectively achieved the stated objectives of the thesis with sequential and systematic design of experiments to overcome the existing research problems for AGB prediction using SRS data.

The increasing trend of small satellites and forestry focused missions will increase the availability of spaceborne remote sensing data. ESA's 7th Earth Explorer mission "BIMOASS" [226] aims to determine the global distribution of forest biomass by reducing the uncertainty in the calculation of carbon stocks and fluxes associated with the terrestrial biosphere. This future mission will consist of a 'P-band' SAR instrument with high capacbilty to penetrate through the forest canopy. Indeed, the lower frequency

signals are more sensitive to the forest vertical structure compared to -X, -C and -L bands. However, P-band SAR alone does not guarantee more accurate AGB models and certainly requires effective feature extraction mechanisms. The mechanism for extraction of effective secondary information from -C band SAR for AGB prediction demonstrated in Chapter 4 can be easily extended to the data from the "BIOMASS" mission to reduce uncertainty in calculation of AGB using target-oriented features from the data. In the passive EO category, the "HORUS" satellite mission [227] aims to develop a low-cost CubeSat constellation to acquire a combination of multispectral and multiangle data. This provides an upgrade over the traditional MS data acquisition systems with continuous multi-angle coverage of forest ecosystems. The MS data based models developed in the thesis can be reformulated to incorporate multi-angle information and deliver more accurate AGB predictions by further reducing the uncertainty with MS data. For example, the acquisitions at nadir can offer less distorted images with minimal influence of atmospheric scattering, which are the main sources of noise in the data. Such images can be used as reference to navigate the calibration of other off-nadir images that can improve the minor irregularities observed in the data samples from the MS images used for modelling. This will result in producing more reliable MS data based AGB models and promote the use of SRS for mapping AGB. The other future developments of this work should be focused to handle such "Big data" for maximum exploitation of its potential and achieving state-of-the-art performance. The adaptive lasso algorithm is a popular technique used in data mining for model selection and feature regularization. However, it is challenging to apply massive data samples with high dimensional features that are effective for achieving desirable AGB prediction results. A possible future work must be focused to extend such technique using parallel solvers to run on multiple cores on shared memory to

deal with high dimensionality and perform faster computation on big data. There are encouraging possibilities in representation learning for deriving AGB-oriented features from satellite remote sensing data. The second contribution used supervised generative neural network based representation learning that can be further modified to learn in semi-supervised mode. The process of collecting reference field data from forest is expensive and time consuming. Generative neural networks can be successfully extended for semi-supervised learning and boost the modelling results with limited reference data. Moreover, advanced methods based on ensemble learning used for automatic ensemble selection of base models require high computational capacity. The requirement of computational capacity will increase for better optimization of the ensemble and to operate with more number of base models. The future developments in this area should be focused on increasing the computational efficiency of ensemble learning algorithms to achieve superior results with limited computational requirements. A global level deployment of such advanced models can improve the current global AGB estimates. The models can be developed to be portable and orchestrated across various machines using cloud environments in the future.

List of Publications

International Journal Publications

1. P. Naik, M. Dalponte, and L. Bruzzone, "Automated Machine Learning Driven Stacked Ensemble Modelling for Forest Aboveground Biomass Prediction Using Multitemporal Sentinel-2 Data," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. Institute of Electrical and Electronics Engineers (IEEE), pp. 1–14, 2022. DOI: 10.1109/JSTARS.2022.3232583

2. P. Naik, M. Dalponte, and L. Bruzzone, "Generative Feature Extraction From Sentinel 1 and 2 Data for Prediction of Forest Aboveground Biomass in the Italian Alps," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15. Institute of Electrical and Electronics Engineers (IEEE), pp. 4755–4771, 2022. doi: 10.1109/jstars.2022.3179027.

3. P. Naik, M. Dalponte, and L. Bruzzone, "Prediction of Forest Aboveground Biomass Using Multitemporal Multispectral Remote Sensing Data," Remote Sensing, vol. 13, no. 7. MDPI AG, p. 1282, Mar. 27, 2021. doi: 10.3390/rs13071282.

4. P. Sivaraj, A. Kumar, S. R. Koti, and P. Naik, "Effects of Training Parameter Concept and Sample Size in Possibilistic c-Means Classifier for Pigeon Pea Specific Crop Mapping," Geomatics, vol. 2, no. 1. MDPI AG, pp. 107–124, Feb. 22, 2022. doi: 10.3390/geomatics2010007.

Conferences

1. P. Naik, M. Dalponte, and L. Bruzzone, "A Disentangled Variational Autoencoder for Prediction of Above Ground Biomass from Hyperspectral Data," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, Jul. 11, 2021. doi: 10.1109/igarss47720.2021.9554415.

2. P. Naik, M. Dalponte, and L. Bruzzone, “A comparison on the use of different satellite multispectral data for the prediction of aboveground biomass,” Image and Signal Processing for Remote Sensing XXVI. SPIE, Sep. 20, 2020. doi: 10.1117/12.2572807.

3. P. Naik and A. Kumar, “A Stochastic Approach for Automatic Collection of Precise Training Data for a Soft Machine Learning Algorithm Using Remote Sensing Images,” Advances in Intelligent Systems and Computing. Springer Singapore, pp. 285–297, 2021. doi: 10.1007/978-981-16-2712-524.

Bibliography

[1] GCOS Second report on the Adequacy of the Global Observing System for Climate in Support of the UNFCCC. World Meteorological Organization, 2003.

[2] H. D. R. Sessa, Ed., Terrestrial essential climate variables for climate change assessment, mitigation and adaptation. Rome: FAO, 2008.

[3] Y. Malhi and O. L. Phillips, "Tropical forests and global atmospheric change: a synthesis," Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, vol. 359, no. 1443. The Royal Society, pp. 549–555, Mar. 29, 2004. doi: 10.1098/rstb.2003.1449.

[4] D. E. Bunker, F. DeClerck, J. C. Bradford, R. K. Colwell, I. Perfecto, O. L. Phillips, M. Sankaran, and S. Naeem, "Species Loss and Aboveground Carbon Storage in a Tropical Forest," Science, vol. 310, no. 5750. American Association for the Advancement of Science (AAAS), pp. 1029–1031, 11-Nov-2005.

[5] T. H. Nguyen, S. Jones, M. Soto-Berelov, A. Haywood, and S. Hislop, "Landsat Time-Series for Estimating Forest Aboveground Biomass and Its Dynamics across Space and Time: A Review," Remote Sensing, vol. 12, no. 1. MDPI AG, p. 98, Dec. 27, 2019. doi: 10.3390/rs12010098.

[6] D. Lu, "Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon," International Journal of Remote Sensing, vol. 26, no. 12. Informa UK Limited, pp. 2509–2525, Jun. 20, 2005. doi: 10.1080/01431160500142145.

[7] P. Rodríguez-Veiga, J. Wheeler, V. Louis, K. Tansey, and H. Balzter, "Quantifying Forest Biomass Carbon Stocks From Space," Current Forestry Reports, vol. 3, no. 1. Springer Science and Business Media LLC, pp. 1–18, Feb. 09, 2017. doi: 10.1007/s40725-017-0052-5.

[8] D. Peng, H. Zhang, L. Liu, W. Huang, A. R. Huete, X. Zhang, F. Wang, L. Yu, Q. Xie, C. Wang, S. Luo, C. Li, and B. Zhang, "Estimating the Aboveground Biomass for Planted Forests Based on Stand Age and Environmental Variables," Remote Sensing, vol. 11, no. 19. MDPI AG, p. 2270, 28-Sep-2019.

[9] X. Zhu and D. Liu, "Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 102. Elsevier BV, pp. 222–231, Apr. 2015. doi: 10.1016/j.isprsjprs.2014.08.014.

[10] A. Günlü, Ercanli, E. Z. Başkent, and G. Çakır, "Estimating aboveground biomass using landsat TM imagery: A case study of Anatolian Crimean pine forests in Turkey," Annals of Forest Research, vol. 57, no. 2, pp. 289–298, 2014.doi: 10.15287/afr.2014.278

[11] I. Chrysafis, G. Mallinis, S. Siachalou, and P. Patias, "Assessing the relationships between growing stock volume and Sentinel-2 imagery in a Mediterranean forest ecosystem," Remote Sensing Letters, vol. 8, no. 6. Informa UK Limited, pp. 508–517, Feb. 22, 2017. doi: 10.1080/2150704x.2017.1295479.

[12] T. Majasalmi and M. Rautiainen, "The potential of Sentinel-2 data for estimating biophysical variables in a boreal forest: a simulation study," Remote Sensing Letters, vol. 7, no. 5. Informa UK Limited, pp. 427–436, Feb. 18, 2016. doi: 10.1080/2150704x.2016.1149251.

[13] S. Wittke, X. Yu, M. Karjalainen, J. Hyyppä, and E. Puttonen, "Comparison of two-dimensional multitemporal Sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest," International Journal of Applied Earth Observation and Geoinformation, vol. 76. Elsevier BV, pp. 167–178, Apr. 2019. doi: 10.1016/j.jag.2018.11.009.

[14] K. Wang, S. E. Franklin, X. Guo, and M. Cattet, "Remote Sensing of Ecology, Biodiversity and Conservation: A Review from the Perspective of Remote Sensing Specialists," Sensors, vol. 10, no. 11. MDPI AG, pp. 9647–9667, Nov. 01, 2010. doi: 10.3390/s101109647.

[15] G. M. Foody, D. S. Boyd, and M. E. J. Cutler, "Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions," Remote Sensing of Environment, vol. 85, no. 4. Elsevier BV, pp. 463–474, Jun. 2003. doi: 10.1016/s0034-4257(03)00039-7.

[16] T. Dube and O. Mutanga, "Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 101. Elsevier BV, pp. 36–46, Mar. 2015. doi: 10.1016/j.isprsjprs.2014.11.001.

[17] S. Abbas, M. S. Wong, J. Wu, N. Shahzad, and S. Muhammad Irteza, "Approaches of Satellite Remote Sensing for the Assessment of Above-Ground Biomass across Tropical Forests: Pan-tropical to National Scales," Remote Sensing, vol. 12, no. 20. MDPI AG, p. 3351, Oct. 14, 2020. doi: 10.3390/rs12203351.

[18] N. G. Silleos, T. K. Alexandridis, I. Z. Gitas, and K. Perakis, "Vegetation Indices: Advances Made in Biomass Estimation and Vegetation Monitoring in the Last 30 Years," Geocarto International, vol. 21, no. 4. Informa UK Limited, pp. 21–28, Dec. 2006. doi: 10.1080/10106040608542399.

[19] X. Li, M. Zhang, J. Long, and H. Lin, "A Novel Method for Estimating Spatial Distribution of Forest Above-Ground Biomass Based on Multispectral Fusion Data and Ensemble Learning Algorithm," Remote Sensing, vol. 13, no. 19. MDPI AG, p. 3910, Sep. 30, 2021. doi: 10.3390/rs13193910.

[20] Askar, N. Nuthammachot, W. Phairuang, P. Wicaksono, and T. Sayektiningsih, "Estimating Aboveground Biomass on Private Forest Using Sentinel-2 Imagery," Journal of Sensors, vol. 2018. Hindawi Limited, pp. 1–11, Dec. 16, 2018. doi: 10.1155/2018/6745629.

[21] J. Chang and M. Shoshany, "Mediterranean shrublands biomass estimation using Sentinel-1 and Sentinel-2," 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, Jul. 2016. doi: 10.1109/igarss.2016.7730380.

[22] R. J. L. Argamosa et al., "Modelling Above Ground Biomass Of Mangrove Forest Using Sentinel-1 Imagery," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. IV–3. Copernicus GmbH, pp. 13–20, Apr. 23, 2018. doi: 10.5194/isprs-annals-iv-3-13-2018.

[23] G. De Luca, J. M. N. Silva, S. Di Fazio, and G. Modica, "Integrated use of Sentinel-1 and Sentinel-2 data and open-source machine learning algorithms for land cover mapping in a Mediterranean region," European Journal of Remote Sensing, vol. 55, no. 1. Informa UK Limited, pp. 52–70, Jan. 12, 2022. doi: 10.1080/22797254.2021.2018667.

[24] Md. L. R. Sarker, J. Nichol, H. B. Iz, B. B. Ahmad, and A. A. Rahman, "Forest Biomass Estimation Using Texture Measurements of High-Resolution Dual-Polarization C-Band SAR Data," IEEE Transactions on Geoscience and Remote Sensing, vol. 51, no. 6. Institute of Electrical and Electronics Engineers (IEEE), pp. 3371–3384, Jun. 2013. doi: 10.1109/tgrs.2012.2219872.

[25] Y. Li, M. Li, C. Li, and Z. Liu, "Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms," Scientific Reports, vol. 10, no. 1. Springer Science and Business Media LLC, Jun. 19, 2020. doi: 10.1038/s41598-020-67024-3.

[26] J. Xue and B. Su, "Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications," Journal of Sensors, vol. 2017. Hindawi Limited, pp. 1–17, 2017. doi: 10.1155/2017/1353691.

[27] C. Kalaitzidis, V. Heinzel, and D. Zianis, "A Review of Multispectral Vegetation Indices for Biomass Estimation," 29th Symposium of the European Association of Remote Sensing Laboratorie, Chania, Greece pp. 201–208, 2010.

[28] Md. L. R. Sarker, J. Nichol, B. Ahmad, I. Busu, and A. A. Rahman, "Potential of texture measurements of two-date dual polarization PALSAR data for the improvement of forest biomass estimation," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 69. Elsevier BV, pp. 146–166, Apr. 2012. doi: 10.1016/j.isprsjprs.2012.03.002.

[29] M. M. Rahman and J. Tetuko Sri Sumantyo, "Retrieval of tropical forest biomass information from ALOS PALSAR data," Geocarto International, vol. 28, no. 5. Informa UK Limited, pp. 382–403, Aug. 2013. doi: 10.1080/10106049.2012.710652.

[30] G. Sandberg, L. M. H. Ulander, J. E. S. Fransson, J. Holmgren, and T. Le Toan, "L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest," Remote Sensing of Environment, vol. 115, no. 11. Elsevier BV, pp. 2874–2886, Nov. 2011. doi: 10.1016/j.rse.2010.03.018.

[31] N. Ghasemi, M. R. Sahebi, and A. Mohammadzadeh, "A review on biomass estimation methods using synthetic aperture radar data," International Journal of Geomatics and Geosciences 2010 Vol.1 No.4 pp.776-788 ref.56

[32] D. Lu, Q. Chen, G. Wang, E. Moran, M. Batistella, M. Zhang, G. Vaglio Laurin, and D. Saah, "Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates," International Journal of Forestry Research, vol. 2012. Hindawi Limited, pp. 1–16, 2012.

[33] G. Wang, M. Zhang, G. Z. Gertner, T. Oyana, R. E. McRoberts, and H. Ge, "Uncertainties of mapping aboveground forest carbon due to plot locations using national forest inventory plot and remotely sensed data," Scandinavian Journal of Forest Research, vol. 26, no. 4. Informa UK Limited, pp. 360–373, Mar. 21, 2011. doi: 10.1080/02827581.2011.564204.

[34] M. Zhang, H. Lin, S. Zeng, J. Li, J. Shi, and G. Wang, "Impacts of Plot Location Errors on Accuracy of Mapping and Scaling Up Aboveground Forest Carbon Using Sample Plot and Landsat TM Data," IEEE Geoscience and Remote Sensing Letters, vol. 10, no. 6. Institute of Electrical and Electronics Engineers (IEEE), pp. 1483–1487, Nov. 2013. doi: 10.1109/lgrs.2013.2260719.

[35] S. Saatchi, Y. Malhi, B. Zutta, W. Buermann, and L. Anderson, "Mapping the distribution of forest structure, biomass, and productivity in Amazonia," Biogeosciences, vol. in press, no. March, 2009.

[36] Y. Gao, D. Lu, G. Li, G. Wang, Q. Chen, L. Liu, and D. Li, "Comparative Analysis of Modeling Algorithms for Forest Aboveground Biomass Estimation in a Subtropical Region," Remote Sensing, vol. 10, no. 4. MDPI AG, p. 627, 18-Apr-2018.

[37] L. Zhao, W. Zhou, Y. Peng, Y. Hu, T. Ma, Y. Xie, L. Wang, J. Liu, and Z. Liu, "A new AG-AGB estimation model based on MODIS and SRTM data in Qinghai Province, China," Ecological Indicators, vol. 133. Elsevier BV, p. 108378, Dec-2021.

[38] O. Mutanga Correspond and A. K. Skidmore, "Narrow band vegetation indices overcome the saturation problem in biomass estimation," International Journal of Remote Sensing, vol. 25, no. 19. Informa UK Limited, pp. 3999–4014, Oct. 2004. doi: 10.1080/01431160310001654923.

[39] S. Pandit, S. Tsuyuki, and T. Dube, "Estimating Above-Ground Biomass in Sub-Tropical Buffer Zone Community Forests, Nepal, Using Sentinel 2 Data," Remote Sensing, vol. 10, no. 4. MDPI AG, p. 601, Apr. 12, 2018. doi: 10.3390/rs10040601.

[40] P. Rodríguez-Veiga, S. Quegan, J. Carreiras, H. J. Persson, J. E. S. Fransson, A. Hoscilo, D. Ziółkowski, K. Stereńczak, S. Lohberger, M. Stängel, A. Berninger, F. Siegert, V. Avitabile, M. Herold, S. Mermoz, A. Bouvet, T. Le Toan, N. Carvalhais, M. Santoro, O. Cartus, Y. Rauste, R. Mathieu, G. P. Asner, C. Thiel, C. Pathe, C. Schmullius, F. M. Seifert, K. Tansey, and H. Balzter, "Forest biomass retrieval approaches from earth observation in different biomes," International Journal of Applied Earth Observation and Geoinformation, vol. 77. Elsevier BV, pp. 53–68, May-2019. doi: 10.1016/j.jag.2018.12.008.

[41] P. Zhao, D. Lu, G. Wang, C. Wu, Y. Huang, and S. Yu, "Examining Spectral Reflectance Saturation in Landsat Imagery and Corresponding Solutions to Improve Forest Aboveground Biomass Estimation," Remote Sensing, vol. 8, no. 6. MDPI AG, p. 469, Jun. 02, 2016. doi: 10.3390/rs8060469.

[42] M. C. Dobson, F. T. Ulaby, T. LeToan, A. Beaudoin, E. S. Kasischke, and N. Christensen, "Dependence of radar backscatter on coniferous forest biomass," IEEE Transactions on Geoscience and Remote Sensing, vol. 30, no. 2. Institute of Electrical and Electronics Engineers (IEEE), pp. 412–415, Mar. 1992. doi: 10.1109/36.134090.

[43] F. Ulaby and M. El-rayes, "Microwave Dielectric Spectrum of Vegetation - Part II: Dual-Dispersion Model," IEEE Transactions on Geoscience and Remote Sensing, vol. GE-25, no. 5. Institute of Electrical and Electronics Engineers (IEEE), pp. 550–557, Sep. 1987. doi: 10.1109/tgrs.1987.289833.

[44] O. Cartus, M. Santoro, and J. Kellndorfer, "Mapping forest aboveground biomass in the Northeastern United States with ALOS PALSAR dual-polarization L-band," Re-

mote Sensing of Environment, vol. 124. Elsevier BV, pp. 466–478, Sep. 2012. doi: 10.1016/j.rse.2012.05.029.

[45] S. Periasamy, "Significance of dual polarimetric synthetic aperture radar in biomass retrieval: An attempt on Sentinel-1," Remote Sensing of Environment, vol. 217. Elsevier BV, pp. 537–549, Nov. 2018. doi: 10.1016/j.rse.2018.09.003.

[46] R. J. L. Argamosa, A. C. Blanco, A. B. Baloloy, C. G. Candido, J. B. L. C. Dumalag, L. L. C. Dimapilis, and E. C. Paringit, "Modelling above ground biomass of mangrove forest using Sentinel-1 imagery," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. IV–3. Copernicus GmbH, pp. 13–20, 23-Apr-2018. doi: 10.5194/isprs-annals-iv-3-13-2018.

[47] S. M. Ghosh and M. D. Behera, "Aboveground biomass estimates of tropical mangrove forest using Sentinel-1 SAR coherence data - The superiority of deep learning over a semi-empirical model," Computers amp; Geosciences, vol. 150. Elsevier BV, p. 104737, May 2021. doi: 10.1016/j.cageo.2021.104737.

[48] S. Roy et al., "Estimating Above Ground Biomass (AGB) and Tree Density using Sentinel-1 Data," Spatial Modeling in Forest Resources Management. Springer International Publishing, pp. 259–280, Oct. 09, 2020. doi: 10.1007/978-3-030-56542-811.

[49] Z. Liu, O. O. Michel, G. Wu, Y. Mao, Y. Hu, and W. Fan, "The Potential of Fully Polarized ALOS-2 Data for Estimating Forest Above-Ground Biomass," Remote Sensing, vol. 14, no. 3. MDPI AG, p. 669, Jan. 30, 2022. doi: 10.3390/rs14030669.

[50] M. K. Nesha, Y. A. Hussin, L. M. van Leeuwen, and Y. B. Sulistioadi, "Modeling and mapping aboveground biomass of the restored mangroves using ALOS-2 PALSAR-2 in East Kalimantan, Indonesia," International Journal of Applied Earth Observation and Geoinformation, vol. 91. Elsevier BV, p. 102158, Sep. 2020. doi: 10.1016/j.jag.2020.102158.

[51] M. Hayashi, T. Motohka, and Y. Sawada, "Aboveground Biomass Mapping Using ALOS-2/PALSAR-2 Time-Series Images for Borneo's Forest," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 12. Institute of Electrical and Electronics Engineers (IEEE), pp. 5167–5177, Dec. 2019. doi: 10.1109/jstars.2019.2957549.

[52] H. L. G. Cassol, J. M. de B. Carreiras, E. C. Moraes, L. E. O. e C. de Aragão, C. V. de J. Silva, S. Quegan, and Y. E. Shimabukuro, "Retrieving Secondary Forest Aboveground Biomass from Polarimetric ALOS-2 PALSAR-2 Data in the Brazilian Amazon," Remote Sensing, vol. 11, no. 1. MDPI AG, p. 59, 29-Dec-2018. doi: 10.3390/rs11010059.

[53] E. Blomberg, L. M. H. Ulander, S. Tebaldini, and L. Ferro-Famil, "Evaluating P-Band TomoSAR for Biomass Retrieval in Boreal Forest," IEEE Transactions on Geoscience and

Remote Sensing, vol. 59, no. 5. Institute of Electrical and Electronics Engineers (IEEE), pp. 3793–3804, May 2021. doi: 10.1109/tgrs.2020.3020775.

[54] G. Sandberg, L. M. H. Ulander, J. Wallerman, and J. E. S. Fransson, "Measurements of Forest Biomass Change Using P-Band Synthetic Aperture Radar Backscatter," IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 10. Institute of Electrical and Electronics Engineers (IEEE), pp. 6047–6061, Oct. 2014. doi: 10.1109/tgrs.2013.2294684.

[55] H. T. M. Dinh, F. Rocca, S. Tebaldini, M. M. d'Alessandro, T. Le Toan, and L. Villard, "Relating tropical forest biomass to P-band SAR tomography," 2012 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Jul. 2012. doi: 10.1109/igarss.2012.6351871.

[56] H. Agata, L. Aneta, Z. Dariusz, S. Krzysztof, L. Marek, S. Christiane, and P. Carsten, "Forest Aboveground Biomass Estimation Using a Combination of Sentinel-1 and Sentinel-2 Data," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Jul-2018. doi: 10.1109/igarss.2018.8517965.

[57] R. K. M. Malhi, A. Anand, P. K. Srivastava, S. K. Chaudhary, M. K. Pandey, M. D. Behera, A. Kumar, P. Singh, and G. Sandhya Kiran, "Synergistic evaluation of Sentinel 1 and 2 for biomass estimation in a tropical forest of India," Advances in Space Research, vol. 69, no. 4. Elsevier BV, pp. 1752–1767, Feb-2022. doi: 10.1016/j.asr.2021.03.035.

[58] D. Stratoulias, N. Nuthammachot, T. Suepa, and K. Phoungthong, "Assessing the Spectral Information of Sentinel-1 and Sentinel-2 Satellites for Above-Ground Biomass Retrieval of a Tropical Forest," ISPRS International Journal of Geo-Information, vol. 11, no. 3. MDPI AG, p. 199, Mar. 16, 2022. doi: 10.3390/ijgi11030199.

[59] T. D. Pham, N. N. Le, N. T. Ha, L. V. Nguyen, J. Xia, N. Yokoya, T. T. To, H. X. Trinh, L. Q. Kieu, and W. Takeuchi, "Estimating Mangrove Above-Ground Biomass Using Extreme Gradient Boosting Decision Trees Algorithm with Fused Sentinel-2 and ALOS-2 PALSAR-2 Data in Can Gio Biosphere Reserve, Vietnam," Remote Sensing, vol. 12, no. 5. MDPI AG, p. 777, 29-Feb-2020. doi: 10.3390/rs12050777.

[60] S. Deng, M. Katoh, Q. Guan, N. Yin, and M. Li, "Estimating Forest Aboveground Biomass by Combining ALOS PALSAR and WorldView-2 Data: A Case Study at Purple Mountain National Park, Nanjing, China," Remote Sensing, vol. 6, no. 9. MDPI AG, pp. 7878–7910, Aug. 25, 2014. doi: 10.3390/rs6097878.

[61] L. Kumar, P. Sinha, S. Taylor, and A. F. Alqurashi, "Review of the use of remote sensing for biomass estimation to support renewable energy generation," Journal of Applied Remote Sensing, vol. 9, no. 1. SPIE-Intl Soc Optical Eng, p. 097696, Jun. 16, 2015. doi: 10.1117/1.jrs.9.097696.

[62] I. Indirabai, M. V. Harindranathan Nair, Jaishankar. R. Nair, and R. R. Nidamanuri, "Optical Remote Sensing for Biophysical Characterisation in Forests: A Review," International Journal of Applied Engineering Research, vol. 14, no. 2. Research India Publications, p. 344, Feb. 28, 2019. doi: 10.37622/ijaer/14.2.2019.344-354.

[63] A. B. Baloloy, A. C. Blanco, C. G. Candido, R. J. L. Argamosa, J. B. L. C. Dumalag, L. L. C. Dimapilis, and E. C. Paringit, "Estimation Of Mangrove Forest Aboveground Biomass Using Multispectral Bands, Vegetation Indices And Biophysical Variables Derived From Optical Satellite Imageries: Rapideye, Planetscope And Sentinel-2," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. IV–3. Copernicus GmbH, pp. 29–36, Apr. 23, 2018. doi: 10.5194/isprs-annals-iv-3-29-2018.

[64] E. Tomppo, M. Nilsson, M. Rosengren, P. Aalto, and P. Kennedy, "Simultaneous use of Landsat-TM and IRS-1C WiFS data in estimating large area tree stem volume and aboveground biomass," Remote Sensing of Environment, vol. 82, no. 1. Elsevier BV, pp. 156–171, Sep. 2002. doi: 10.1016/s0034-4257(02)00031-7.

[65] D. Zheng, J. Rademacher, J. Chen, T. Crow, M. Bresee, J. Le Moine, and S.-R. Ryu, "Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA," Remote Sensing of Environment, vol. 93, no. 3. Elsevier BV, pp. 402–411, Nov-2004. doi: 10.1016/j.rse.2004.08.008.

[66] A. Baccini, M. A. Friedl, C. E. Woodcock, and R. Warbington, "Forest biomass estimation over regional scales using multisource data," Geophysical Research Letters, vol. 31, no. 10. American Geophysical Union (AGU), p. n/a-n/a, May 2004. doi: 10.1029/2004gl019782.

[67] B. Zhang, L. Zhang, D. Xie, X. Yin, C. Liu, and G. Liu, "Application of Synthetic NDVI Time Series Blended from Landsat and MODIS Data for Grassland Biomass Estimation," Remote Sensing, vol. 8, no. 1. MDPI AG, p. 10, Dec. 24, 2015. doi: 10.3390/rs8010010.

[68] P. S. Thenkabail et al., "Biomass estimations and carbon stock calculations in the oil palm plantations of African derived savannas using IKONOS data," International Journal of Remote Sensing, vol. 25, no. 23. Informa UK Limited, pp. 5447–5472, Dec. 2004. doi: 10.1080/01431160412331291279.

[69] H. Su, W. Shen, J. Wang, A. Ali, and M. Li, "Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests," Forest Ecosystems, vol. 7, no. 1. Elsevier BV, Nov. 26, 2020. doi: 10.1186/s40663-020-00276-7.

[70] P. S. Thenkabail, N. Stucky, B. W. Griscom, M. S. Ashton, J. Diels, B. van der Meer, and E. Enclona, "Biomass estimations and carbon stock calculations in the oil palm plantations of African derived savannas using IKONOS data," International Journal of Remote Sensing, vol. 25, no. 23. Informa UK Limited, pp. 5447–5472, Dec-2004. doi: 10.3390/f13040616.

[71] M. K. Steininger, "Satellite estimation of tropical secondary forest above-ground biomass: Data from Brazil and Bolivia," International Journal of Remote Sensing, vol. 21, no. 6–7. Informa UK Limited, pp. 1139–1157, Jan. 2000. doi: 10.1080/014311600210119.

[72] R. F. Nelson, D. S. Kimes, W. A. Salas, and M. Routheir, "Secondary Forest Age and Tropical Forest Biomass Estimation Using Thematic Mapper Imagery," BioScience, vol. 50, no. 5. Oxford University Press (OUP), p. 419, 2000. doi: 10.1641/0006-3568(2000)050[0419:sfaatf]2.0.co;2.

[73] M.-H. Phua and H. Saito,"Estimation of biomass of a mountainous tropical forest using Landsat TM data," Canadian Journal of Remote Sensing, 29:4, 429-440, DOI: 10.5589/m03-005

[74] C. Song, M. B. Dickinson, L. Su, S. Zhang, and D. Yaussey, "Estimating average tree crown size using spatial information from Ikonos and QuickBird images: Across-sensor and across-site comparisons," Remote Sensing of Environment, vol. 114, no. 5. Elsevier BV, pp. 1099–1107, May 2010. doi: 10.1016/j.rse.2009.12.022.

[75] J. A. Greenberg, S. Z. Dobrowski, and S. L. Ustin, "Shadow allometry: Estimating tree structural parameters using hyperspatial image analysis," Remote Sensing of Environment, vol. 97, no. 1. Elsevier BV, pp. 15–25, Jul. 2005. doi: 10.1016/j.rse.2005.02.015.

[76] S. A. Soenen, D. R. Peddle, R. J. Hall, C. A. Coburn, and F. G. Hall, "Estimating above-ground forest biomass from canopy reflectance model inversion in mountainous terrain," Remote Sensing of Environment, vol. 114, no. 7. Elsevier BV, pp. 1325–1337, Jul. 2010. doi: 10.1016/j.rse.2009.12.012.

[77] C. Li, L. Zhou, and W. Xu, "Estimating Aboveground Biomass Using Sentinel-2 MSI Data and Ensemble Algorithms for Grassland in the Shengjin Lake Wetland, China," Remote Sensing, vol. 13, no. 8. MDPI AG, p. 1595, Apr. 20, 2021. doi: 10.3390/rs13081595.

[78] S. Puliti, J. Breidenbach, J. Schumacher, M. Hauglin, T. F. Klingenberg, and R. Astrup, "Above-ground biomass change estimation using national forest inventory data with Sentinel-2 and Landsat," Remote Sensing of Environment, vol. 265. Elsevier BV, p. 112644, Nov. 2021. doi: 10.1016/j.rse.2021.112644.

[79] S. Muhe and M. Argaw, "Estimation of above-ground biomass in tropical afro-montane forest using Sentinel-2 derived indices," Environmental Systems Research, vol. 11, no. 1. Springer Science and Business Media LLC, Mar. 22, 2022. doi: 10.1186/s40068-022-00250-y.

[80] S. Labrecque, R. A. Fournier, J. E. Luther, and D. Piercey, "A comparison of four methods to map biomass from Landsat-TM and inventory data in western Newfoundland," Forest

Ecology and Management, vol. 226, no. 1–3. Elsevier BV, pp. 129–144, May 2006. doi: 10.1016/j.foreco.2006.01.030.

[81] D. Zheng, L. S. Heath, and M. J. Ducey, "Forest biomass estimated from MODIS and FIA data in the Lake States: MN, WI and MI, USA," Forestry, vol. 80, no. 3. Oxford University Press (OUP), pp. 265–278, Jun. 18, 2007. doi: 10.1093/forestry/cpm015.

[82] G. Yin, Y. Zhang, Y. Sun, T. Wang, Z. Zeng, and S. Piao, "MODIS Based Estimation of Forest Aboveground Biomass in China," PLOS ONE, vol. 10, no. 6. Public Library of Science (PLoS), p. e0130143, Jun. 26, 2015. doi: 10.1371/journal.pone.0130143.

[83] A. Baccini, N. Laporte, S. J. Goetz, M. Sun, and H. Dong, "A first map of tropical Africa's above-ground biomass derived from satellite imagery," Environmental Research Letters, vol. 3, no. 4. IOP Publishing, p. 045011, Oct. 2008. doi: 10.1088/1748-9326/3/4/045011.

[84] M. Dalponte and D. A. Coomes, "Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data," Methods in Ecology and Evolution, vol. 7, no. 10. Wiley, pp. 1236–1245, May 14, 2016. doi: 10.1111/2041-210x.12575.

[85] M. Dalponte, T. Jucker, S. Liu, L. Frizzera, and D. Gianelle, "Characterizing forest carbon dynamics using multi-temporal lidar data," Remote Sensing of Environment, vol. 224. Elsevier BV, pp. 412–420, Apr. 2019. doi: 10.1016/j.rse.2019.02.018.

[86] P. Morcillo-Pallarés, J. P. Rivera-Caicedo, S. Belda, C. De Grave, H. Burriel, J. Moreno, and J. Verrelst, "Quantifying the Robustness of Vegetation Indices through Global Sensitivity Analysis of Homogeneous and Forest Leaf-Canopy Radiative Transfer Models," Remote Sensing, vol. 11, no. 20. MDPI AG, p. 2418, 18-Oct-2019. doi: 10.3390/rs11202418.

[87] S. J. Becker, C. S. T. Daughtry, and A. L. Russ, "Robust Forest Cover Indices for Multispectral Images," Photogrammetric Engineering amp; Remote Sensing, vol. 84, no. 8. American Society for Photogrammetry and Remote Sensing, pp. 505–512, Aug. 01, 2018. doi: 10.14358/pers.84.8.505.

[88] H. Zou, "The Adaptive Lasso and Its Oracle Properties," Journal of the American Statistical Association, vol. 101, no. 476. Informa UK Limited, pp. 1418–1429, Dec. 01, 2006. doi: 10.1198/016214506000000735.

[89] J. Friedman ,T. Hastie ,R. Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent", J Stat Softw. 2010;33(1):1-22. PMID: 20808728; PMCID: PMC2929880.

[90] Z. Zhang, M. Liu, X. Liu, and G. Zhou, "A New Vegetation Index Based on Multitemporal Sentinel-2 Images for Discriminating Heavy Metal Stress Levels in Rice," Sensors, vol. 18, no. 7. MDPI AG, p. 2172, Jul. 06, 2018. doi: 10.3390/s18072172.

[91] F. Wang, J. Huang, Y. Tang, and X. Wang, "New Vegetation Index and Its Application in Estimating Leaf Area Index of Rice," Rice Science, vol. 14, no. 3. Elsevier BV, pp. 195–203, Sep. 2007. doi: 10.1016/s1672-6308(07)60027-4.

[92] Z. Jiang, A. Huete, K. Didan, and T. Miura, "Development of a two-band enhanced vegetation index without a blue band," Remote Sensing of Environment, vol. 112, no. 10. Elsevier BV, pp. 3833–3845, Oct. 15, 2008. doi: 10.1016/j.rse.2008.06.006.

[93] M. J. P. Sullivan, S. L. Lewis, W. Hubau, L. Qie, T. R. Baker, L. F. Banin, J. Chave, A. Cuni-Sanchez, T. R. Feldpausch, G. Lopez-Gonzalez, E. Arets, P. Ashton, J. Bastin, N. J. Berry, J. Bogaert, R. Boot, F. Q. Brearley, R. Brienen, D. F. R. P. Burslem, C. Canniere, M. Chudomelová, M. Dančák, C. Ewango, R. Hédl, J. Lloyd, J. Makana, Y. Malhi, B. S. Marimon, B. H. M. Junior, F. Metali, S. Moore, L. Nagy, P. N. Vargas, C. A. Pendry, H. Ramírez-Angulo, J. Reitsma, E. Rutishauser, K. A. Salim, B. Sonké, R. S. Sukri, T. Sunderland, M. Svátek, P. M. Umunay, R. V. Martinez, R. R. E. Vernimmen, E. V. Torre, J. Vleminckx, V. Vos, and O. L. Phillips, "Field methods for sampling tree height for tropical forest biomass estimation," Methods in Ecology and Evolution, vol. 9, no. 5. Wiley, pp. 1179–1189, 13-Feb-2018. doi: 10.1111/2041-210x.12962.

[94] A. Banskota, N. Kayastha, M. J. Falkowski, M. A. Wulder, R. E. Froese, and J. C. White, "Forest Monitoring Using Landsat Time Series Data: A Review," Canadian Journal of Remote Sensing, vol. 40, no. 5. Informa UK Limited, pp. 362–384, Sep. 03, 2014. doi: 10.1080/07038992.2014.987376.

[95] A. Günlü, I. Ercanli, E. Başkent, and G. Çakır, "Estimating aboveground biomass using Landsat TM imagery: A case study of Anatolian Crimean pine forests in Turkey," Annals Of Forest Research, vol. 57, no.2, pp. 289-298. doi:10.15287/afr.2014.278, 2014.

[96] S. Tanaka, T. Takahashi, T. Nishizono, F. Kitahara, H. Saito, T. Iehara, E. Kodani, and Y. Awaya, "Stand Volume Estimation Using the k-NN Technique Combined with Forest Inventory Data, Satellite Image Data and Additional Feature Variables," Remote Sensing, vol. 7, no. 1. MDPI AG, pp. 378–394, 31-Dec-2014. doi: 10.3390/rs70100378.

[97] J. A. A. Castillo, A. A. Apan, T. N. Maraseni, and S. G. Salmo III, "Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines using Sentinel imagery," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 134. Elsevier BV, pp. 70–85, Dec. 2017. doi: 10.1016/j.isprsjprs.2017.10.016.

[98] L. Korhonen, Hadi, P. Packalen, and M. Rautiainen, "Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index," Remote Sensing of Environment, vol. 195. Elsevier BV, pp. 259–274, Jun. 2017. doi: 10.1016/j.rse.2017.03.021.

[99] U. B. Gewali, S. T. Monteiro, and E. Saber, "Gaussian Processes for Vegetation Parameter Estimation from Hyperspectral Data with Limited Ground Truth," Remote Sensing, vol. 11, no. 13. MDPI AG, p. 1614, Jul. 08, 2019. doi: 10.3390/rs1113161.

[100] B. Petrovska, E. Zdravevski, P. Lameski, R. Corizzo, I. Štajduhar, and J. Lerga, "Deep Learning for Feature Extraction in Remote Sensing: A Case-Study of Aerial Scene Classification," Sensors, vol. 20, no. 14. MDPI AG, p. 3906, Jul. 14, 2020. doi: 10.3390/s20143906.

[101] J. Heaton, "An empirical analysis of feature engineering for predictive modeling," SoutheastCon 2016. IEEE, Mar. 2016. doi: 10.1109/secon.2016.7506650.

[102] P. Naik and A. Kumar, "A Stochastic Approach for Automatic Collection of Precise Training Data for a Soft Machine Learning Algorithm Using Remote Sensing Images," Advances in Intelligent Systems and Computing. Springer Singapore, pp. 285–297, 2021. doi: 10.1007/978-981-16-2712-524.

[103] P. Sivaraj, A. Kumar, S. R. Koti, and P. Naik, "Effects of Training Parameter Concept and Sample Size in Possibilistic c-Means Classifier for Pigeon Pea Specific Crop Mapping," Geomatics, vol. 2, no. 1. MDPI AG, pp. 107–124, Feb. 22, 2022. doi: 10.3390/geomatics2010007.

[104] C. Poultney, S. Chopra, Y. L. Cun, and others, "Efficient learning of sparse representations with an energy-based model," Advances in Neural Information Processing Systems 19, pp. 1137–1144, 2006.

[105] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," Nature, vol. 323, no. 6088. Springer Science and Business Media LLC, pp. 533–536, Oct. 1986. doi: 10.1038/323533a0.

[106] W. La Cava, T. R. Singh, J. Taggart, S. Suri, and J. H. Moore, "Learning concise representations for regression by evolving networks of trees." arXiv, 2018. doi: 10.48550/ARXIV.1807.00981.

[107] S. Mohanty and G. Singh, "Fully polarimetric synthetic aperture radar indices for scintillation observation," 2019 URSI Asia-Pacific Radio Science Conference (AP-RASC). IEEE, Mar. 2019. doi: 10.23919/ursiap-rasc.2019.8738514.

[108] M. García, S. Saatchi, A. Casas, A. Koltunov, S. Ustin, C. Ramirez, and H. Balzter, "Extrapolating Forest Canopy Fuel Properties in the California Rim Fire by Combining Airborne LiDAR and Landsat OLI Data," Remote Sensing, vol. 9, no. 4. MDPI AG, p. 394, 22-Apr-2017. doi: 10.3390/rs9040394.

[109] E. Halme, P. Pellikka, and M. Mõttus, "Utility of hyperspectral compared to multispectral remote sensing data in estimating forest biomass and structure variables in Finnish boreal

forest," International Journal of Applied Earth Observation and Geoinformation, vol. 83. Elsevier BV, p. 101942, Nov. 2019. doi: 10.1016/j.jag.2019.101942.

[110] R. Sheridan, S. Popescu, D. Gatziolis, C. Morgan, and N.-W. Ku, "Modeling Forest Aboveground Biomass and Volume Using Airborne LiDAR Metrics and Forest Inventory and Analysis Data in the Pacific Northwest," Remote Sensing, vol. 7, no. 1. MDPI AG, pp. 229–255, Dec. 24, 2014. doi: 10.3390/rs70100229.

[111] L. Ruiz, T. Hermosilla, F. Mauro, and M. Godino, "Analysis of the Influence of Plot Size and LiDAR Density on Forest Structure Attribute Estimates," Forests, vol. 5, no. 5. MDPI AG, pp. 936–951, May 16, 2014. doi: 10.3390/f5050936.

[112] Z. Wu, D. Dye, J. Stoker, J. Vogel, M. Velasco, and B. Middleton, "Evaluating Lidar Point Densities for Effective Estimation of Aboveground Biomass," International Journal of Advanced Remote Sensing and GIS, vol. 5, no. 1. Cloud Publications, pp. 1483–1499, 09-Feb-2016. doi: 10.23953/cloud.ijarsg.40.

[113] N. Jha, N. K. Tripathi, N. Barbier, S. G. P. Virdis, W. Chanthorn, G. Viennois, W. Y. Brockelman, A. Nathalang, S. Tongsima, N. Sasaki, R. Pélissier, and M. Réjou-Méchain, "The real potential of current passive satellite data to map aboveground biomass in tropical forests," Remote Sensing in Ecology and Conservation, vol. 7, no. 3. Wiley, pp. 504–520, 08-Apr-2021. doi: 10.1002/rse2.203.

[114] P. Naik, M. Dalponte, and L. Bruzzone, "Prediction of Forest Aboveground Biomass Using Multitemporal Multispectral Remote Sensing Data," Remote Sensing, vol. 13, no. 7. MDPI AG, p. 1282, Mar. 27, 2021. doi: 10.3390/rs13071282.

[115] T. Tuong, H. Tani, X. Wang, N. Thang, and H. Bui, "Combination of SAR Polarimetric Parameters for Estimating Tropical Forest Aboveground Biomass," Polish Journal of Environmental Studies, vol. 29, no. 5. HARD Publishing Company, pp. 3353–3365, May 12, 2020. doi: 10.15244/pjoes/112900.

[116] P. Naik, M. Dalponte, and L. Bruzzone, "A comparison on the use of different satellite multispectral data for the prediction of aboveground biomass," Image and Signal Processing for Remote Sensing XXVI. SPIE, Sep. 20, 2020. doi: 10.1117/12.2572807.

[117] T. Pham, N. Yokoya, D. Bui, K. Yoshino, and D. Friess, "Remote Sensing Approaches for Monitoring Mangrove Species, Structure, and Biomass: Opportunities and Challenges," Remote Sensing, vol. 11, no. 3. MDPI AG, p. 230, Jan. 22, 2019. doi: 10.3390/rs11030230.

[118] Y. Zhang, S. Liang, and L. Yang, "A Review of Regional and Global Gridded Forest Biomass Datasets," Remote Sensing, vol. 11, no. 23. MDPI AG, p. 2744, Nov. 22, 2019. doi: 10.3390/rs11232744.

[119] S. Sinha, C. Jeganathan, L. K. Sharma, and M. S. Nathawat, "A review of radar remote sensing for biomass estimation," International Journal of Environmental Science and Technology, vol. 12, no. 5. Springer Science and Business Media LLC, pp. 1779–1792, Jan. 20, 2015. doi: 10.1007/s13762-015-0750-0.

[120] P. Naik, M. Dalponte, and L. Bruzzone, "A Disentangled Variational Autoencoder for Prediction of Above Ground Biomass from Hyperspectral Data," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, Jul. 11, 2021. doi: 10.1109/igarss47720.2021.9554415.

[121] M. A. Tanase, R. Panciera, K. Lowell, J. Hacker, and J. P. Walker, "Estimation of forest biomass from L-band polarimetric decomposition components," 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS. IEEE, Jul. 2013. doi: 10.1109/igarss.2013.6721318.

[122] I. Huuva, J. E. S. Fransson, H. J. Persson, J. Wallerman, L. M. H. Ulander, E. Blomberg, and M. J. Soja, "Measurements of forest biomass change using L- and P-band sar backscatter," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, Jul-2017. doi: 10.1109/igarss.2017.8128331.

[123] M. Schlund and M. Davidson, "Aboveground Forest Biomass Estimation Combining L- and P-Band SAR Acquisitions," Remote Sensing, vol. 10, no. 7. MDPI AG, p. 1151, Jul. 20, 2018. doi: 10.3390/rs10071151.

[124] M. E. J. Cutler, D. S. Boyd, G. M. Foody, and A. Vetrivel, "Estimating tropical forest biomass with a combination of SAR image texture and Landsat TM data: An assessment of predictions between regions," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 70. Elsevier BV, pp. 66–77, Jun. 2012. doi: 10.1016/j.isprsjprs.2012.03.011.

[125] J. Lei, X. Song, L. Sun, M. Song, N. Li, and C. Chen, "Learning deep classifiers with deep features," 2016 IEEE International Conference on Multimedia and Expo (ICME). IEEE, Jul. 2016. doi: 10.1109/icme.2016.7552910.

[126] X.-Y. Tong, G.-S. Xia, F. Hu, Y. Zhong, M. Datcu, and L. Zhang, "Exploiting Deep Features for Remote Sensing Image Retrieval: A Systematic Investigation," IEEE Transactions on Big Data, vol. 6, no. 3. Institute of Electrical and Electronics Engineers (IEEE), pp. 507–521, Sep. 01, 2020. doi: 10.1109/tbdata.2019.2948924.

[127] L. Xu, Y. Chen, S. Srinivasan, N. de Freitas, A. Doucet, and A. Gretton, "Learning Deep Features in Instrumental Variable Regression." arXiv, 2020. doi: 10.48550/ARXIV.2010.07154.

[128] N. T. Ha, M. Manley-Harris, T. D. Pham, and I. Hawes, "The use of radar and optical satellite imagery combined with advanced machine learning and metaheuristic optimization techniques to detect and quantify above ground biomass of intertidal seagrass in a New Zealand estuary," International Journal of Remote Sensing, vol. 42, no. 12. Informa UK Limited, pp. 4712–4738, Mar. 23, 2021. doi: 10.1080/01431161.2021.1899335.

[129] S. Vafaei, J. Soosani, K. Adeli, H. Fadaei, H. Naghavi, T. Pham, and D. Tien Bui, "Improving Accuracy Estimation of Forest Aboveground Biomass Based on Incorporation of ALOS-2 PALSAR-2 and Sentinel-2A Imagery and Machine Learning: A Case Study of the Hyrcanian Forest Area (Iran)," Remote Sensing, vol. 10, no. 2. MDPI AG, p. 172, 25-Jan-2018. doi: 10.3390/rs10020172.

[130] R. C. Sharma, "Ensemble Learning of Multi-Source Satellite Sensors Dataset for Estimating Forest Biomass in New England Region." MDPI AG, Feb. 16, 2021. doi: 10.20944/preprints202102.0338.v1.

[131] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources," IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4. Institute of Electrical and Electronics Engineers (IEEE), pp. 8–36, Dec-2017. doi: 10.1109/mgrs.2017.2762307.

[132] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 152. Elsevier BV, pp. 166–177, Jun. 2019. doi: 10.1016/j.isprsjprs.2019.04.015.

[133] G. Tsagkatakis, A. Aidini, K. Fotiadou, M. Giannopoulos, A. Pentari, and P. Tsakalides, "Survey of Deep-Learning Approaches for Remote Sensing Observation Enhancement," Sensors, vol. 19, no. 18. MDPI AG, p. 3929, Sep. 12, 2019. doi: 10.3390/s19183929.

[134] A. Romero, C. Gatta, and G. Camps-Valls, "Unsupervised Deep Feature Extraction for Remote Sensing Image Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 3. Institute of Electrical and Electronics Engineers (IEEE), pp. 1349–1362, Mar. 2016. doi: 10.1109/tgrs.2015.2478379.

[135] S. Ghosh, L. Bruzzone, S. Patra, F. Bovolo, and A. Ghosh, "A Context-Sensitive Technique for Unsupervised Change Detection Based on Hopfield-Type Neural Networks," IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 3. Institute of Electrical and Electronics Engineers (IEEE), pp. 778–789, Mar. 2007. doi: 10.1109/tgrs.2006.888861.

[136] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community," Journal of Applied

Remote Sensing, vol. 11, no. 04. SPIE-Intl Soc Optical Eng, p. 1, Sep. 23, 2017. doi: 10.1117/1.jrs.11.042609.

[137] M. Zhu, Y. He, and Q. He, "A Review of Researches on Deep Learning in Remote Sensing Application," International Journal of Geosciences, vol. 10, no. 01. Scientific Research Publishing, Inc., pp. 1–11, 2019. doi: 10.4236/ijg.2019.101001.

[138] T. Hoeser, F. Bachofer, and C. Kuenzer, "Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review—Part II: Applications," Remote Sensing, vol. 12, no. 18. MDPI AG, p. 3053, Sep. 18, 2020. doi: 10.3390/rs12183053.

[139] L. Zhu, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Generative Adversarial Networks for Hyperspectral Image Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 9. Institute of Electrical and Electronics Engineers (IEEE), pp. 5046–5063, Sep. 2018. doi: 10.1109/tgrs.2018.2805286.

[140] X. Liu, Y. Wang, and Q. Liu, "Psgan: A Generative Adversarial Network for Remote Sensing Image Pan-Sharpening," 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, Oct. 2018. doi: 10.1109/icip.2018.8451049.

[141] S. Hinterstoisser, V. Lepetit, N. Rajkumar, and K. Konolige, "Going Further with Point Pair Features," Computer Vision – ECCV 2016. Springer International Publishing, pp. 834–848, 2016. doi: 10.1007/978-3-319-46487-951.

[142] Z. Wang, Q. She, and T. E. Ward, "Generative Adversarial Networks in Computer Vision," ACM Computing Surveys, vol. 54, no. 2. Association for Computing Machinery (ACM), pp. 1–38, Mar. 31, 2022. doi: 10.1145/3439723.

[143] C. Li, K. Xu, J. Zhu, J. Liu, and B. Zhang, "Triple Generative Adversarial Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence. Institute of Electrical and Electronics Engineers (IEEE), pp. 1–1, 2021. doi: 10.1109/tpami.2021.3127558.

[144] B. Rodríguez-Suárez, P. Quesada-Barriuso, and F. Argüello, "Design of CGAN Models for Multispectral Reconstruction in Remote Sensing," Remote Sensing, vol. 14, no. 4. MDPI AG, p. 816, Feb. 09, 2022. doi: 10.3390/rs14040816.

[145] Y. Yoo, S. Yun, H. J. Chang, Y. Demiris, and J. Y. Choi, "Variational Autoencoded Regression: High Dimensional Regression of Visual Data on Complex Manifold," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Jul. 2017. doi: 10.1109/cvpr.2017.314.

[146] X. Li, X. Jin, J. Lin, T. Yu, S. Liu, Y. Wu, W. Zhou, and Z. Chen, "Learning Disentangled Feature Representation for Hybrid-distorted Image Restoration." arXiv, 2020. doi: 10.48550/ARXIV.2007.11430.

[147] Y. Liu, F. Wei, J. Shao, L. Sheng, J. Yan, and X. Wang, "Exploring Disentangled Feature Representation Beyond Face Identification," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, Jun. 2018. doi: 10.1109/cvpr.2018.00222.

[148] C. T. Marx, R. L. Phillips, S. A. Friedler, C. Scheidegger, and S. Venkatasubramanian, "Disentangling Influence: Using Disentangled Representations to Audit Model Predictions." arXiv, 2019. doi: 10.48550/ARXIV.1906.08652.

[149] A. H. Liu, Y.-C. Liu, Y.-Y. Yeh, and Y.-C. F. Wang, "A Unified Feature Disentangler for Multi-Domain Image Translation and Manipulation." arXiv, 2018. doi: 10.48550/ARXIV.1809.01361.

[150] F. Del Frate and D. Solimini, "On Neural Network Algorithms for Retrieving Forest Biomass From SAR Data," IEEE Transactions on Geoscience and Remote Sensing, vol. 42, no. 1. Institute of Electrical and Electronics Engineers (IEEE), pp. 24–34, Jan. 2004. doi: 10.1109/tgrs.2003.817220.

[151] L. Zhang, Z. Shao, J. Liu, and Q. Cheng, "Deep Learning Based Retrieval of Forest Aboveground Biomass from Combined LiDAR and Landsat 8 Data," Remote Sensing, vol. 11, no. 12. MDPI AG, p. 1459, Jun. 20, 2019. doi: 10.3390/rs11121459.

[152] Z. Shao, L. Zhang, and L. Wang, "Stacked Sparse Autoencoder Modeling Using the Synergy of Airborne LiDAR and Satellite Optical and SAR Data to Map Forest Above-Ground Biomass," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 12. Institute of Electrical and Electronics Engineers (IEEE), pp. 5569–5582, Dec. 2017. doi: 10.1109/jstars.2017.2748341.

[153] M. Main-Knorn, B. Pflug, J. Louis, V. Debaecker, U. Müller-Wilm, and F. Gascon, "Sen2Cor for Sentinel-2," Image and Signal Processing for Remote Sensing XXIII. SPIE, Oct. 04, 2017. doi: 10.1117/12.2278218.

[154] N. Bhogapurapu, S. Dey, D. Mandal, A. Bhattacharya, and Y. Rao, "PolSAR tools: A QGIS plugin for generating SAR descriptors," Journal of Open Source Software, vol. 6, no. 60. The Open Journal, p. 2970, Apr. 08, 2021. doi: 10.21105/joss.02970.

[155] T. D. Nguyen and M. Kappas, "Estimating the Aboveground Biomass of an Evergreen Broadleaf Forest in Xuan Lien Nature Reserve, Thanh Hoa, Vietnam, Using SPOT-6 Data and the Random Forest Algorithm," International Journal of Forestry Research, vol. 2020. Hindawi Limited, pp. 1–13, Aug. 27, 2020. doi: 10.1155/2020/4216160.

[156] C. Wu, H. Shen, A. Shen, J. Deng, M. Gan, J. Zhu, H. Xu, and K. Wang, "Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery,"

Journal of Applied Remote Sensing, vol. 10, no. 3. SPIE-Intl Soc Optical Eng, p. 035010, 08-Aug-2016. doi: 10.1117/1.jrs.10.035010.

[157] Y. Zhang, J. Ma, S. Liang, X. Li, and M. Li, "An Evaluation of Eight Machine Learning Regression Algorithms for Forest Aboveground Biomass Estimation from Multiple Satellite Data Products," Remote Sensing, vol. 12, no. 24. MDPI AG, p. 4015, Dec. 08, 2020. doi: 10.3390/rs12244015.

[158] Scrinzi, G.; Galvagni, D.; Marzullo, L. I Nuovi Modelli Dendrometrici per la Stima Delle Masse Assestamentali in Provincia di Trento;Provincia Autonoma di Trento-Servizio Foreste e fauna: Trento, Italy, 2010; ISBN 978-88-7702-271-4.

[159] IPCC. (2003). Good Practice Guidance for Land Use , Land-Use Change and Forestry (eds J. Penman, M. Gytarsky, T. Hiraishi, T. Krug , D. Kruger , R. Pipatti, L. Buendia , K. Miwa , T. Ngara , K. Tanabe Fabian Wagner). IPCC National Greenhouse Gas Inventories Programme, Kanagawa, Japan.

[160] N. Ghasemi, M. R. Sahebi, and A. Mohammadzadeh, "A review on biomass estimation methods using synthetic aperture radar data," International journal of Geomatics and Geosciences, vol. 1, no. 4, pp. 776–788, 2011.

[161] Y. Zhang, J. Ma, S. Liang, X. Li, and J. Liu, "A stacking ensemble algorithm for improving the biases of forest aboveground biomass estimations from multiple remotely sensed datasets," GIScience amp; Remote Sensing, vol. 59, no. 1. Informa UK Limited, pp. 234–249, Jan. 02, 2022. doi: 10.1080/15481603.2021.2023842.

[162] L. Chen, Y. Wang, C. Ren, B. Zhang, and Z. Wang, "Optimal Combination of Predictors and Algorithms for Forest Above-Ground Biomass Mapping from Sentinel and SRTM Data," Remote Sensing, vol. 11, no. 4. MDPI AG, p. 414, Feb. 18, 2019. doi: 10.3390/rs11040414.

[163] C. Liu, B. Zoph, M. Neumann, J. Shlens, W. Hua, L.-J. Li, L. Fei-Fei, A. Yuille, J. Huang, and K. Murphy, "Progressive Neural Architecture Search." arXiv, 2017. doi: 10.48550/ARXIV.1712.00559.

[164] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," IEEE Transactions on Evolutionary Computation, vol. 1, no. 1. Institute of Electrical and Electronics Engineers (IEEE), pp. 67–82, Apr. 1997. doi: 10.1109/4235.585893.

[165] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of Go with deep neural networks and tree search," Nature,

vol. 529, no. 7587. Springer Science and Business Media LLC, pp. 484–489, 27-Jan-2016. doi: 10.1038/nature16961.

[166] D. Maclaurin, D. Duvenaud, and R. P. Adams, "Gradient-based Hyperparameter Optimization through Reversible Learning." arXiv, 2015. doi: 10.48550/ARXIV.1502.03492.

[167] X. He, K. Zhao, and X. Chu, "AutoML: A survey of the state-of-the-art," Knowledge-Based Systems, vol. 212. Elsevier BV, p. 106622, Jan. 2021. doi: 10.1016/j.knosys.2020.106622.

[168] Q. Yao, M. Wang, Y. Chen, W. Dai, Y.-F. Li, W.-W. Tu, Q. Yang, and Y. Yu, "Taking Human out of Learning Applications: A Survey on Automated Machine Learning." arXiv, 2018. doi: 10.48550/ARXIV.1810.13306.

[169] X. He, S. Wang, X. Chu, S. Shi, J. Tang, X. Liu, C. Yan, J. Zhang, and G. Ding, "Automated Model Design and Benchmarking of 3D Deep Learning Models for COVID-19 Detection with Chest CT Scans." arXiv, 2021. doi: 10.48550/ARXIV.2101.05442.

[170] L. Faes, S. K. Wagner, D. J. Fu, X. Liu, E. Korot, J. R. Ledsam, T. Back, R. Chopra, N. Pontikos, C. Kern, G. Moraes, M. K. Schmid, D. Sim, K. Balaskas, L. M. Bachmann, A. K. Denniston, and P. A. Keane, "Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study," The Lancet Digital Health, vol. 1, no. 5. Elsevier BV, pp. e232–e242, Sep-2019. doi: 10.1016/s2589-7500(19)30108-6.

[171] Y. Chen, T. Yang, X. Zhang, G. Meng, X. Xiao, and J. Sun, "DetNAS: Backbone Search for Object Detection." arXiv, 2019. doi: 10.48550/ARXIV.1903.10979.

[172] H. Xu, L. Yao, Z. Li, X. Liang, and W. Zhang, "Auto-FPN: Automatic Network Architecture Adaptation for Object Detection Beyond Classification," 2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Oct. 2019. doi: 10.1109/iccv.2019.00675.

[173] J. Guo, K. Han, Y. Wang, C. Zhang, Z. Yang, H. Wu, X. Chen, and C. Xu, "Hit-Detector: Hierarchical Trinity Architecture Search for Object Detection," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Jun-2020. doi: 10.1109/cvpr42600.2020.01142.

[174] X. Chu, B. Zhang, H. Ma, R. Xu, and Q. Li, "Fast, Accurate and Lightweight Super-Resolution with Neural Architecture Search," 2020 25th International Conference on Pattern Recognition (ICPR). IEEE, Jan. 10, 2021. doi: 10.1109/icpr48806.2021.9413080.

[175] D. Song, C. Xu, X. Jia, Y. Chen, C. Xu, and Y. Wang, "Efficient Residual Dense Block Search for Image Super-Resolution," Proceedings of the AAAI Conference on Artificial In-

telligence, vol. 34, no. 07. Association for the Advancement of Artificial Intelligence (AAAI), pp. 12007–12014, Apr. 03, 2020. doi: 10.1609/aaai.v34i07.6877.

[176] Y. Guo, Y. Luo, Z. He, J. Huang, and J. Chen, "Hierarchical Neural Architecture Search for Single Image Super-Resolution," IEEE Signal Processing Letters, vol. 27. Institute of Electrical and Electronics Engineers (IEEE), pp. 1255–1259, 2020. doi: 10.1109/lsp.2020.3003517.

[177] Y. Jiang, C. Hu, T. Xiao, C. Zhang, and J. Zhu, "Improved Differentiable Architecture Search for Language Modeling and Named Entity Recognition," Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Association for Computational Linguistics, 2019. doi: 10.18653/v1/d19-1367.

[178] J. Chen, K. Chen, X. Chen, X. Qiu, and X. Huang, "Exploring Shared Structures and Hierarchies for Multiple NLP Tasks." arXiv, 2018. doi: 10.48550/ARXIV.1808.07658.

[179] C. Liu, L.-C. Chen, F. Schroff, H. Adam, W. Hua, A. L. Yuille, and L. Fei-Fei, "Auto-DeepLab: Hierarchical Neural Architecture Search for Semantic Image Segmentation," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Jun-2019. doi: 10.1109/cvpr.2019.00017.

[180] C. Yang, J. Fan, Z. Wu, and M. Udell, "AutoML Pipeline Selection: Efficiently Navigating the Combinatorial Space," Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery; Data Mining. ACM, Aug. 20, 2020. doi: 10.1145/3394486.3403197.

[181] H. J. Escalante, W.-W. Tu, I. Guyon, D. L. Silver, E. Viegas, Y. Chen, W. Dai, and Q. Yang, "AutoML @ NeurIPS 2018 Challenge: Design and Results," The NeurIPS '18 Competition. Springer International Publishing, pp. 209–229, 30-Nov-2019. doi: 10.1007/978-3-030-29135-88.

[182] M. Feurer, A. Klein, K. Eggensperger, J. T. Springenberg, M. Blum, and F. Hutter, "Efficient and robust automated machine learning," Advances in Neural Information Processing Systems,2015.

[183] N. Pillay and R. Qu, "Automated Design of Machine Learning and Search Algorithms." Springer International Publishing, 2021. doi: 10.1007/978-3-030-72069-8.

[184] H. J. Escalante, M. Montes, L. E. Sucar, 'Particle Swarm Model Selection', The Journal of Machine Learning Research,vol. 10, pp. 405–440, 2009.

[185] H. J. Escalante, M. Montes, and E. Sucar, "Ensemble particle swarm model selection," The 2010 International Joint Conference on Neural Networks (IJCNN). IEEE, Jul. 2010. doi: 10.1109/ijcnn.2010.5596915.

[186] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, F. Hutter, 'Efficient and Robust Automated Machine Learning', Advances in Neural Information Processing Systems, vol. 2, pp. 2755–2763, 2015.

[187] D. A. Masood, Automated Machine Learning: Hyperparameter optimization, neural architecture search, and algorithm selection with cloud platforms. Packt Publishing Ltd, 2021.

[188] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Auto-WEKA," Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, Aug. 11, 2013. doi: 10.1145/2487575.2487629.

[189] F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Sequential Model-Based Optimization for General Algorithm Configuration," Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 507–523, 2011. doi: 10.1007/978-3-642-25566-340.

[190] M. Feurer, K. Eggensperger, S. Falkner, M. Lindauer, and F. Hutter, "Auto-Sklearn 2.0: Hands-free AutoML via Meta-Learning." arXiv, 2020. doi: 10.48550/ARXIV.2007.04074.

[191] R. S. Olson and J. H. Moore, "TPOT: A Tree-Based Pipeline Optimization Tool for Automating Machine Learning," Automated Machine Learning. Springer International Publishing, pp. 151–160, 2019. doi: 10.1007/978-3-030-05318-58.

[192] H. Mendoza, A. Klein, M. Feurer, J. T. Springenberg, M. Urban, M. Burkart, M. Dippel, M. Lindauer, and F. Hutter, "Towards Automatically-Tuned Deep Neural Networks," Automated Machine Learning. Springer International Publishing, pp. 135–149, 2019. doi: 10.1007/978-3-030-05318-57.

[193] B. Zoph and Q. V. Le, "Neural Architecture Search with Reinforcement Learning." arXiv, 2016. doi: 10.48550/ARXIV.1611.01578.

[194] H. Jin, Q. Song, and X. Hu, "Auto-Keras: An Efficient Neural Architecture Search System," Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery amp; Data Mining. ACM, Jul. 25, 2019. doi: 10.1145/3292500.3330648.

[195] Joaquin Vanschoren, AutoML Book - Meta-Learning. 2018.

[196] J. P. Monteiro, D. Ramos, D. Carneiro, F. Duarte, J. M. Fernandes, and P. Novais, "Meta-learning and the new challenges of machine learning," International Journal of Intelligent Systems, vol. 36, no. 11. Wiley, pp. 6240–6272, Jun. 30, 2021. doi: 10.1002/int.22549.

[197] K. A. Smith-Miles, "Cross-disciplinary perspectives on meta-learning for algorithm selection," ACM Computing Surveys, vol. 41, no. 1. Association for Computing Machinery (ACM), pp. 1–25, Jan. 15, 2009. doi: 10.1145/1456650.1456656.

[198] M. Maher and S. Sakr, "SmartML: A Meta Learning-Based Framework for Automated Selection and Hyperparameter Tuning for Machine Learning Algorithms." OpenProceedings.org, 2019. doi: 10.5441/002/EDBT.2019.54.

[199] C. Yang, Y. Akimoto, D. W. Kim, and M. Udell, "OBOE," Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery amp; Data Mining. ACM, Jul. 25, 2019. doi: 10.1145/3292500.3330909.

[200] K. Li and J. Malik, "Learning to Optimize." arXiv, 2016. doi: 10.48550/ARXIV.1606.01885.

[201] F. Nargesian, H. Samulowitz, U. Khurana, E. B. Khalil, and D. Turaga, "Learning Feature Engineering for Classification," Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Aug. 2017. doi: 10.24963/ijcai.2017/352.

[202] H. Li, Z. Cui, Z. Zhu, L. Chen, J. Zhu, H. Huang, and C. Tao, "RS-MetaNet: Deep Metametric Learning for Few-Shot Remote Sensing Scene Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 8. Institute of Electrical and Electronics Engineers (IEEE), pp. 6983–6994, Aug-2021., doi: 10.1109/TGRS.2020.3027387.

[203] L. Benrais and N. Baha, "High level visual scene classification using background knowledge of objects," Multimedia Tools and Applications, vol. 81, no. 3. Springer Science and Business Media LLC, pp. 3663–3692, Nov. 18, 2021. doi: 10.1007/s11042-021-11701-6.

[204] P. Zhang, Y. Bai, D. Wang, B. Bai, and Y. Li, "Few-Shot Classification of Aerial Scene Images via Meta-Learning," Remote Sensing, vol. 13, no. 1. MDPI AG, p. 108, Dec. 31, 2020. doi: 10.3390/rs13010108.

[205] M. Ruswurm, S. Wang, M. Korner, and D. Lobell, "Meta-Learning for Few-Shot Land Cover Classification," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, Jun. 2020. doi: 10.1109/cvprw50498.2020.00108.

[206] M. Werther, E. Spyrakos, S. G. H. Simis, D. Odermatt, K. Stelzer, H. Krawczyk, O. Berlage, P. Hunter, and A. Tyler, "Meta-classification of remote sensing reflectance to estimate trophic status of inland and nearshore waters," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 176. Elsevier BV, pp. 109–126, Jun-2021. doi: 10.1016/j.isprsjprs.2021.04.003.

[207] X. Tang, W. Lin, C. Liu, X. Han, W. Wang, J. Ma, and L. Jiao, "Multi-Scale Meta-Learning-Based Networks for High-Resolution Remote Sensing Scene Classification," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, 11-Jul-2021. doi: 10.1109/IGARSS47720.2021.9555134.

[208] A. Méndez, 'Meta-Learning for Instance Segmentation on Satellite Imagery', 2020.

[209] Q. Yang, Z. Ni, and P. Ren, "Meta captioning: A meta learning based remote sensing image captioning framework," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 186. Elsevier BV, pp. 190–200, Apr. 2022. doi: 10.1016/j.isprsjprs.2022.02.001.

[210] L. Rokach, Pattern Classification Using Ensemble Methods, Vol. 75. Singapore: World Scientific Publishing Co. Pte. Ltd., 2018.

[211] K. T. Ngo, "Stacking Ensemble for auto ml," Virginia Polytechnic Institute and State University, 2018.

[212] J. Yoo, T. Joseph, D. Yung, S. A. Nasseri, and F. Wood, "Ensemble Squared: A Meta AutoML System." arXiv, 2020. doi: 10.48550/ARXIV.2012.05390.

[213] N. Erickson, J. Mueller, A. Shirkov, H. Zhang, P. Larroy, M. Li, and A. Smola, "AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data." arXiv, 2020. doi: 10.48550/ARXIV.2003.06505.

[214] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, F. Hutter, 'Efficient and Robust Automated Machine Learning', Advances in Neural Information Processing Systems, vol.2, pp. 2755–2763, 2015.

[215] E. LeDell, 'H2O AutoML: Scalable Automatic Machine Learning', 2020.

[216] J. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, 'Algorithms for Hyper-Parameter Optimization', Advances in Neural Information Processing Systems, vol. 24, pp. 2546–2554, 2011.

[217] M. J. van der Laan, E. C. Polley, and A. E. Hubbard, "Super Learner," Statistical Applications in Genetics and Molecular Biology, vol. 6, no. 1. Walter de Gruyter GmbH, Jan. 16, 2007. doi: 10.2202/1544-6115.1309.

[218] J. Louis, "Sentinel-2 SEN2COR: L2A processor for users," Eur. Sp. Agency, (Special Publ. ESA SP, vol. SP-740, no. August, pp. 9–13, 2016.

[219] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A Next-generation Hyperparameter Optimization Framework." arXiv, 2019. doi: 10.48550/ARXIV.1907.10902.

[220] C. Henrich, V., Krauss, G., Götze, C., Sandow, "IDB - Index-Database; Development of a database for remote sensing indices." Crop Science Colloquium, University of Bonn, 2012.

[221] J. C. O. Koh, B. P. Banerjee, G. Spangenberg, and S. Kant, "Automated hyperspectral vegetation index derivation using a hyperparameter optimisation framework for high-throughput plant phenotyping," New Phytologist, vol. 233, no. 6. Wiley, pp. 2659–2670, Jan. 20, 2022. doi: 10.1111/nph.17947.

[222] Y. Wang, X. Zhang, and Z. Guo, "Estimation of tree height and aboveground biomass of coniferous forests in North China using stereo ZY-3, multispectral Sentinel-2, and DEM data," Ecological Indicators, vol. 126. Elsevier BV, p. 107645, Jul. 2021. doi: 10.1016/j.ecolind.2021.107645.

[223] R. D. Wolfinger, P.-Y. Tan, 'Stacked Ensemble Models for Improved Prediction Accuracy', SAS, 2017.

[224] I. Kandel, M. Castelli, and A. Popovič, "Comparing Stacking Ensemble Techniques to Improve Musculoskeletal Fracture Image Classification," Journal of Imaging, vol. 7, no. 6. MDPI AG, p. 100, Jun. 21, 2021. doi: 10.3390/jimaging7060100.

[225] S. P. Healey, W. B. Cohen, Z. Yang, C. Kenneth Brewer, E. B. Brooks, N. Gorelick, A. J. Hernandez, C. Huang, M. Joseph Hughes, R. E. Kennedy, T. R. Loveland, G. G. Moisen, T. A. Schroeder, S. V. Stehman, J. E. Vogelmann, C. E. Woodcock, L. Yang, and Z. Zhu, "Mapping forest change using stacked generalization: An ensemble approach," Remote Sensing of Environment, vol. 204. Elsevier BV, pp. 717–728, Jan-2018. doi: 10.1016/j.rse.2017.09.029.

[226] T. L. Toan, J. Chave, J. Dall, K. Papathanassiou, P. Paillou, M. Rechstein, S. Quegan, S. Saatchi, K. Seipel, H. Shugart, S. Tebaldini, L. Ulander, and M. Williams, "The Biomass Mission: Objectives and Requirements," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Jul-2018.

[227] A. Pellegrino, M. G. Pancalli, A. Gianfermo, P. Marzioli, F. Curianò, F. Angeletti, F. Piergentili, and F. Santoni, "HORUS: Multispectral and Multiangle CubeSat Mission Targeting Sub-Kilometer Remote Sensing Applications," Remote Sensing, vol. 13, no. 12. MDPI AG, p. 2399, 19-Jun-2021.

[228] C. Li, M. Li, J. Liu, Y. Li, and Q. Dai, "Comparative Analysis of Seasonal Landsat 8 Images for Forest Aboveground Biomass Estimation in a Subtropical Forest," Forests, vol. 11, no. 1. MDPI AG, p. 45, 31-Dec-2019.

[229] C. Li, M. Li, and Y. Li, "Improving estimation of forest aboveground biomass using Landsat 8 imagery by incorporating forest crown density as a dummy variable," Canadian Journal of Forest Research, vol. 50, no. 4. Canadian Science Publishing, pp. 390–398, Apr-2020.

Appendix

The allometric equations used for the estimation of the AGB of each tree in the field ($\widehat{AGB}_{TREE}$) were based on the stem volume equations published in [158] multiplied by the wood density (WD) of each species (IPCC 2003). The stem volume equations of [158] were developed for the region where our study areas are located. The equation for estimating $\widehat{AGB}_{TREE}$ is in the following form:

$$\widehat{AGB}_{TREE} = WD * \alpha * (DBH - d_0)^{\gamma} * H^{\delta} \tag{6.1}$$

where DBH is the diameter in centimeters, H the height in meters, and $\widehat{AGB}_{TREE}$ is the estimated aboveground biomass in kilograms. The coefficients used for the different species are in Table 6.1.

Table 6.1: Coefficients of the allometric equations of Scrinzi et al. (2010) and wood densities (WD) from (IPCC 2003). The wood density (WD) is expressed in kg/m3

	WD	α	γ	δ	d_0
Abies alba Mill.	400	0.000163	1.70656	0.941905	3.69465
Broadleaves	580	0.000055	1.942089	1.00642	4.0091
Larix decidua Mill.	460	0.000108	1.407756	1.341377	3.69465
Picea abies (L.) Karst.	400	0.000177	1.564254	1.051565	3.69465
Pinus cembra L.	420	0.000188	1.613713	0.985266	3.69465
Pinus nigra J.F.Arnold	420	0.000129	1.763086	0.938445	3.69465
Pinus sylvestris L.	420	0.000102	1.918184	0.830164	3.69465