MINISTRY OF AGRICULTURE AND RURAL DEVELOPMENT

MINISTRY OF EDUCATION AND TRAINING

VIETNAM NATIONAL UNIVERSITY OF FORESTRY

PHAM VAN DUAN

A TECHNICAL STUDY ON ESTIMATION OF FOREST WOOD VOLUME USING SATELLITE IMAGES IN DAK NONG PROVINCE

SUMMARY OF DOCTORIAL THESIS

Majors: Forest inventory and planning

Code: 9620208

HA NOI, 2019

The thesis is completed at:

Vietnam National University of Forestry,
Xuan Mai town – Chuong My district – Ha Noi city

Scientific supervisors:

1. Associate professor. Dr. Nguyen Trong Binh
2. Dr. Nguyen Thanh Hoan

Reviewer 1: ....

....

Reviewer 2: ....

....

Reviewer 3: ....

....

The thesis will be upheld at the University-level Assessment Council:.......

....

At the time of ……, day.............month................year......

The dissertation can be found at the library: National Library or Library of Vietnam National University of Forestry

LIST OF PUBLISHED WORKS RELATED TO THE THESIS

1. **Pham Van Duan**, Vu Thi Thin (2015), *Determinate estimate biomass and volume forests from satellite images*. Forestry Science and Technology Journal No. 3, 2015.

2. **Pham Van Duan**, Vu Thi Thin, Nguyen Quoc Huy (2016), *Estimated value of the object-oriented optimal segmentation parameters within ecognition software: experiments in satellite images SPOT-6*. Forestry Science and Technology Journal No. 6, 2016.

3. Nguyen Thanh Hoan, **Pham Van Duan**, Le Sy Doanh and Nguyen Van Dung (2017), *Determining the locations of deforestation using multi-variant change vector analysis (MCVA) on landsat-8 satellite data*. Forestry Science and Technology Journal No. 4, 2017.

4. **Pham Van Duan**, Nguyen Thanh Hoan, Nguyen Trong Binh and Pham Tien Dung (2017), *A combination of ALOS-2/PALSAR-2 and LANDSAT-8 satellite images for wood volume estimation of evergreen broadleaf forest in Dak Nong province*. Vietnam Journal of Forest Science No. 4, 2017.

5. **Pham Van Duan**, Nguyen Thanh Hoan, Nguyen Trong Binh and Vu Thi Thin (2018), *Building a model to identify the evergreen broadleaf natural wood forest in Dak Nong province using remote sensing data*. Science and Technology Journal of Agriculture and Rural Development, period 3 + 4, 2018.

TABLE OF CONTENTS

INTRODUCTION

1. The need of the thesis

Forest status map showing the frontier forest status and forest wood volume is an important tool in forest management and is one of the bases for developing policies, strategies and arrangements forest protection and development activities.

In the past, forest wood volume was determined from forest status, that is, pre-existing forest status map and forest wood volume was calculated by status. However, at present, the classification of the forest status of our country is based on forest volume, without forest volume, the forest status cannot be determined. Therefore, information of forest wood volume becomes particularly important, especially for the forest investigation and inventory programs regulated in the Forest Law. Technical innovation to ensure the identification of forest wood volume to each forest plot is a requirement being practiced by the reality.

One of the current feasible methods to identify forest wood volume on a wide scale in a short time is to use remote sensing images. There are 3 types of remote sensing images commonly used to identify forest wood volume: Optical, RADAR and LIDAR. However, the role of each type of image in determining forest wood volume is different. In particular, LIDAR image does not have satellite receiver so the application is limited. Therefore, studies identifying forest wood volume from remote sensing images mainly use Optical and RADAR images.

Identifying forest wood volume from satellite images is a complex task, including many steps including: Selection, image processing, selection of variables on images, selection of suitable algorithms that best simulate the relationship between forest wood volume and variables on images, collecting field data to build and verify models, build models, apply models to identify forest wood volume have been studied in many parts of the world. However, in Vietnam this issue has not been studied and applied satisfactorily.

According to the forest inventory results in 2014, Dak Nong province had 253.962.3 ha of forest, rate of forest cover 39.0%. Beside economic value, Dak Nong forest is particularly important with the function of protecting, protecting water sources, preventing erosion ... However, due to many different reasons, the current status of Dak Nong forest in recent years in Many places have a decline in both quantity and quality. Facing this situation, beside tightening management to maintain the existing forest area combined with afforestation on the land area planned for forestry development, the forest status maps in which forest wood volume is determined to each forest plot need to be constantly updated on a regular basis.

From the above reasons, the thesis “A technical study on estimation of forest wood volume using satellite images in Dak Nong province” is carried out with the view: technical research to identify forest wood volume from satellite images is study steps techniques and conditions for applying those technical steps so that can be identified forest wood volume from satellite images, including: image selection; processing images; select variables from image; identify forest wood volume at the scene; select algorithms to build a model for determining forest wood volume; assess the error of models and identify the main technical factors affecting the accuracy of the forest wood volume determination models; identify forest wood volume to each pixel; identify forest wood volume to each forest plot.

2. The objectives of the thesis

Research and select techniques to identify forest wood volume from satellite images to improve the quality of forest status mapping in Vietnam. Specifically: (1) Assess the technical factors affecting the efficiency of forest wood volume determination from satellite images; (2) Develop techniques to identify forest wood volume from satellite images.

3. Research subject and scope of the thesis

The object of the thesis is forest types and satellite images selected in the study area with the following scope: (1) About time: to be implemented in the period of 2013 - 2016; (2) Regarding forest type: implemented with evergreen broadleaf natural forest; (3) Satellite images materials: LANDSAT-8 and ALOS-2/PALSAR-2 Satellite images.

4. New contributions of the thesis

- Confirmed the possibility of LANDSAT-8 and ALOS-2/PALSAR-2 satellite images in forest wood volume estimation for evergreen broad-leaved natural forest in Dak Nong province.

- Selected the optimal input variables for wood volume estimation of evergreen broad-leaved natural forest in Dak Nong province from LANDSAT-8, ALOS-2/PALSAR-2 images and from combination of the two image types.

- Selected the optimal algorithm for wood volume estimation of evergreen broad-leaved natural forest in Dak Nong province from current commonly applied algorithms.

- Combining LANDSAT-8 and ALOS-2/PALSAR-2 satellite images made the error of forest volume result acceptable, which can be applied into practice in periodical forest investigation and inventory, as well as support for forest management, monitoring, updating and determining carbon sequestration capacity of the forest.

5. Scientific and practical significance of the thesis

5.1. Scientific significance

The thesis is a comprehensive technical research on estimation of forest wood volume for natural evergreen broadleaf forest in Dak Nong province including: Selection of satellite images; Processing satellite images; Selecting variables from the satellite images; Determining forest wood volume at the scene (Plots); Selecting algorithms to build models; Assessing the error of the model and identify the main technical factors affecting the accuracy of the model to determine the forest wood volume; Determining forest wood volume to each pixel; Determining forest wood volume to each forest plot.

Based on the results of the thesis, the abilities of LANDSAT-8 and ALOS-2/PALSAR-2 satellite images, also the combination of the two satellite images, have already been confirmed for wood volume estimation of evergreen broad-leaved natural forest in Dak Nong province in particular and it can be applied for the other areas in Vietnam that have similar conditions in general.

The dissertation provides a theoretical basis and methodologies for estimating forest wood volume using satellite imagery that can be used as a good reference for other studies in Dak Nong province in particular and in Vietnam in general.

5.2. Practical significance

Applying the process of forest wood volume estimation from satellite images in the thesis to determine the wood volume of evergreen broad-leaved natural forest in Dak Nong province, the results are relatively consistent with the forest inventory results. This is an important practical significance for using this process in Dak Nong.

Currently, three important tasks in forest resource management and monitoring that the forestry sector has been and will be implementing are: investigation, inventory and update of changes of forest. In which: (1) Forest investigation is conducted every 5 years; (2) Forest inventory is conducted every 10 years; (3) Update of forest changes is conducted annually. The results of the thesis will provide solutions to identify forest wood volume at low cost, which can be implemented on a large scale to support investigation and inventory of forest.

I. OVERVIEW AND ORIENTATION OF THE RESEARCH ISSUES

1.1. Overview of research issues

The results finding show that, in order to identify forest wood volume from satellite images, studies often focus on: (1) Selecting appropriate image materials; (2) Identify suitable variables from images associated with forest wood volume; (3) Identify suitable algorithms to build defined models forest wood volume; (4) Analysis of factors affecting the accuracy of identify forest wood volume.

- Selecting satellite images materials: based on wavelength, there are three main types of remote sensing image materials: Optical, RADAR and LIDAR. Each type has different strengths and weaknesses when used to identify forest wood volume. In which:

+ Optical satellite images are the most commonly used document for identifying forest wood volume. Normally, medium and low resolution images are provided free of charge and vice versa. Different types of optical images have been used by many authors to determine forest wood volume and have achieved certain results. In general, high resolution images are better for estimating forest structure attributes than low and medium resolution images. However, the limited high-resolution image is the large fluctuation value due to the shade of the tree and the shadow of the terrain, thereby causing errors for the model to identify forest wood volume. Besides, the image high resolution need data storage capacity, time for image processing and hardware configuration requirements, software for large image processing and high cost of image materials. With a wide research area, the processing ability and the cost to purchase images are important factors that influence the decision to select high-resolution satellite data in research as well as practical applications to identify forest wood volume.

+ RADAR satellite image: The wavelength determines how electromagnetic radiation interacts with the on the surface of object, so it is important information when using RADAR images to identify forest wood volume. RADAR documentation has short wavelength (channels X, C) cannot get information deep inside dense forest canopy, whereas RADAR data with long wavelengths (channels L, P) can get information deep in the leaf canopy can even get the information in the soil layer below the canopy is closely related to forest wood volume. Therefore, RADAR images are often considered to be better to identify forest wood volume than optical satellite images.

- Determine suitable variables from forest wood volume related satellite images: Many satellite image variables have been used in the estimation model forest wood volume. However, not all variables are useful in constructing this estimation indicator model. For optical satellite images, techniques such as vegetation index determination, main component analysis, spectral mixed analysis, structure analysis, etc. were used to create new variables in addition to regular universal value variables. For RADAR images, backscatter values are often used as input variables in forest wood volume estimation . On the other hand, forest wood volume is influenced by many factors such as geography, climate ... but in most cases, these factors have been ignored by assuming that forested areas are homogeneous in terms of the geography and climate. Therefore, adding some variables of geography, climate ,etc. combined with variables from satellite images can improve the determination error forest wood volume, because terrain factors, climate ... affect the vertical structure and growth of forest trees.

- Identify suitable algorithms to build the determination model forest wood volume: Many algorithms have been developed for estimating forest wood volume from satellite images, divided into two categories: parametric and non-parametric. The parametric algorithm assumes that the relationship between forest wood volume (dependent variable) and independent variable derived from satellite imagery can be modeled using univariate, multivariate, or nonlinear linear regression functions. Many authors have used the regression function to determine forest wood volume. However, in reality the relationship between forest wood volume and the independent variables identified from satellite images is often very complex, so sometimes the parameter algorithm does not show good this relationship. In contrast, non-parametric algorithms do not predetermine the model structure so it is more flexible than the empirical regression method. Non-parametric algorithms such as K-NN, ANN, SVM, RT, RF, etc. are often used to determine forest wood volume from optical satellite images but there are few studies using non-parametric algorithms to build models identify forest wood volume from RADAR image. In order to determine an optimal algorithm, many studies have conducted comparative analysis of the determination results forest wood volume from satellite images with different algorithms to determine the most suitable algorithm. However, due to many different reasons, this comparison has not made a clear effect. Therefore, the determination of the effect of the algorithm on the efficiency of determination forest wood volume is still left open.

- Factors affecting the accuracy of forest wood volume determination: identifying the origin of the error of determining forest wood volume from satellite images is of special importance and has been of interest to many scientists. The results show that: (1) the error of determining forest wood volume can vary from 5% to 30%, depending on forest ecosystems, geographic characteristics, monitoring data, spatial resolution of satellite image, methods used ... (2) The selection of different regression models to determine forest wood volume from satellite images can give errors up to 20%. (3) The sample plot size affects the accuracy of the estimate forest wood volume, the estimated precision forest wood volume increases 10% when the size of the sample plot increases from 0.25 ha to 1 ha or the estimated error forest wood volume has been decrease by 38% when sample size increased from 0.36 ha to 1 ha. In addition, the location of sample plots does not affect the accuracy of the determination of forest wood volume.

- The size of the study area affects the accuracy of the estimation of forest wood volume from satellite images through the appropriateness between sample plot size and spatial resolution of satellite image data. Theoretically, high spatial resolution satellite images do not require a large sample area, but in a forest ecosystem, a very small sample plot loses its representation and creates an error in determining forest wood volume immediately fieldwork due to its complex structure. The majority of sample plots used in the forest inventory range in size from 400-1,000 m^2. These dimensions can be: large for satellite images with high spatial resolution, resulting in large variations in spectral values on the same sample plot; Relatively suitable for satellite images with medium spatial resolution, but may not be suitable for low spatial resolution satellite images. Collecting field data is a very expensive job. Therefore, the first priority is to choose a sample plot size that represents the study area with the lowest collection cost.

1.2. Orientation of research issues

- In this study, the forest wood volume is the total volume from the root to the top of the trees in the stand, with unit m^3/ha, denoted as M.

- The wood volume of a forest are related to the spectral reflectance characteristics from that forest and its variation in space. Therefore, the development of the technique to determine forest wood volume from satellite images must first be the selection of image types and determination of reflectance indexes of spectral reflectance characteristics and its spatial variation for each type of image.

- Different types of satellite images will have different spectral and spatial resolutions.

Therefore, they are capable of identifying forest wood volume with different precision and on different scales. In general, the higher the spectral resolution, the more accurate forest wood volume is to determine, the higher the spatial resolution, the better the ability to distinguish the forest canopy surface structure and the higher the accuracy of forest wood volume.

- Determination of forest wood volume from satellite images is based on spectral characteristics and their distribution on space-based satellite images to calculate forest wood volume. Therefore, the development of the technique to determine forest wood volume from satellite images is the construction formulas, algorithm selection to calculate forest wood volume from spectral reflectance indexes and reflect their distribution in space. The best model is one that allows identifying forest wood volume to each forest plot with the lowest errors.

- Methods of forest classification determine the method of determining the status and forest wood volume. In the past, forest wood volume in an area often surveyed determined by forest status. Accordingly, first build a map showing the status of the forest, then arrange and investigate the sample plots on each status and determine the average forest wood volume for each status. All forest plots in a status are assigned reserves equal to the average volume of that state. However, under current conditions, without forest wood volume, forest status cannot be determined. Therefore, the condition of a pre-existing status map and the subsequent reserve calculation is not feasible. Therefore, the forest wood volume determination model from satellite images in this study must be the forest wood volume determination model for each location of the forest type. From there, to evaluate the error of determining forest wood volume from the satellite image, we also have to stand on the view of knowing only the forest type but not the forest status.

II. BASIC CONDITIONS OF THE RESEARCH AREA AND CHARACTERISTICS SATELLITE IMAGES MATERIALS USED

2.1. Basic conditions of Dak Nong province related to the research issue

Dak Nong has a natural area of 651.561.5 ha, with diverse terrains, alternating between majestic and rugged high mountains with vast high plateaus. The rainy season is from April to the end of October, concentrating on 85% of the annual rainfall; dry season from November to the end of March next year. The type of evergreen broadleaf natural forest is the main forest type in Dak Nong province, and also the large-area forest type in Vietnam, which is why the author chose this type of forest as the object of research.

2.2. Characteristics of satellite image data used in research

Based on the results of the general analysis and characteristics of satellite image materials, the author chooses: 1) LANDSAT-8 image - representing optical satellite images data of medium resolution; 2) ALOS-2/PALSAR-2 images represent RADAR (L) long-wavelength data to study and develop forest wood volume determination models from images for Dak Nong province. In particular, based on the climatic conditions in Dak Nong and the time to collect field data, using 5 scenes LANDSAT-8 taken from November 14th,2014 to March 3nd, 2015 and 6 scenes ALOS-2/PALSAR- 2 taken from September 21th, 2014 - January 25th, 2015 to conduct research.

III. CONTENT AND RESEARCH METHODS

3.1. Research content

(1) Investigate the relationship between variables from satellite image and non-image with forest wood volume;

(2) Study to build a model for determining forest wood volume by multivariate regression function;

(3) Study to build a model for determining forest wood volume by non-parametric algorithms;

(4) Research to combine satellite images with inventory plot boundaries to develop a model for identifying forest wood volume;

(5) Proposing procedure for determining forest wood volume from satellite images.

3.2. Research methods

The entire research process of the thesis is shown in the diagram in Figure 3.1

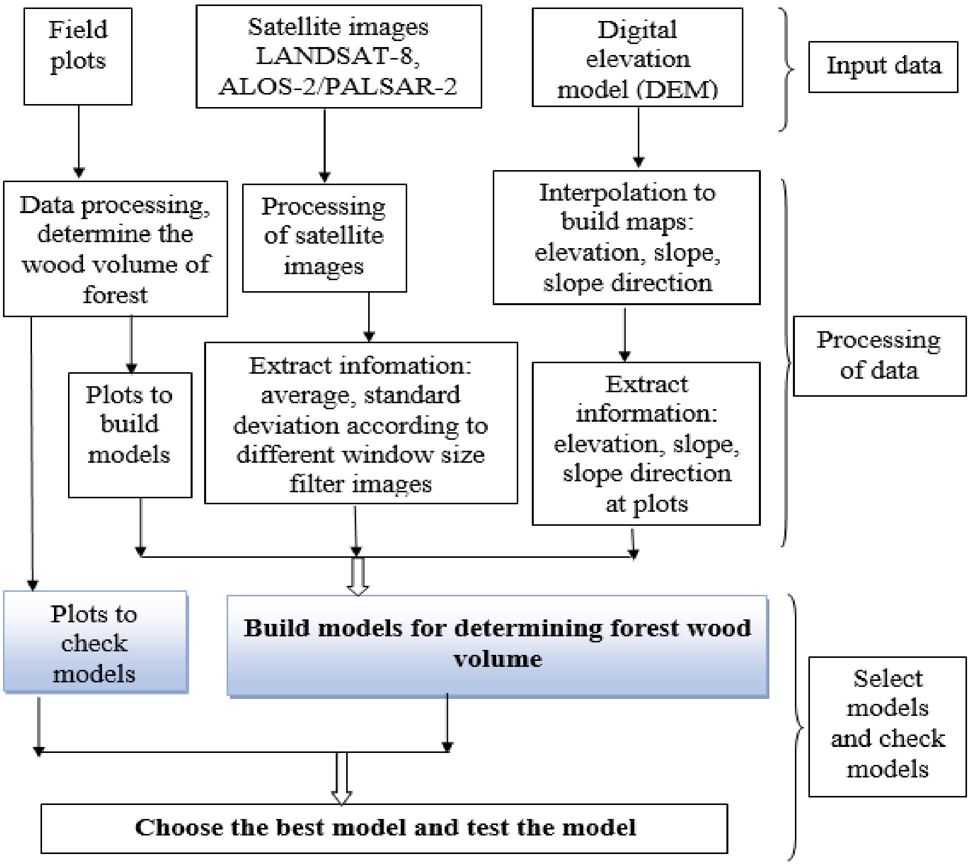

Figure 3.1. Diagram of the thesis's research process

3.2.1. Collecting and processing foreign data

- Data collection: at 214 plots evergreen broadleaf natural forest in the study area, collection time from October, 2014 to Feberuary/2015, plot size: 1,000 square meter (30mx33.3m).

- Data processing: at each plot, determine the volume of each individual tree by a volume chart of 2 factors made nationwide for evergreen broadleaf natural forest, then determine the total volume of trees in the plot and forest wood volume at the plot position. The plots list includes the following criteria: Plot position (x, y) and forest wood volume are used for the next research content.

3.2.2. Method of processing and extracting information on satellite images and non-images at the plot location

- Satellite image processing: LANDSAT-8: (1) conversion of coordinate system; (2) correcting the effects of terrain shadows; (3) mixing images to create multi-spectral channels with a resolution of 15m; (4) build the main component images and plant index images. ALOS-2/PALSAR-2: (1) conversion of coordinate system; (2) convert the digital number on the image to the backscatter value. The result creates 4 types of images: (1) Image of vegetation index (symbol NDVI); (2) Image of the main component (symbol PC);

(3) Image of backscatter HV channel (symbol HV); (4) Image of backscatter HH channel (symbol HH).

- Extract value of satellite images and non-image information at the plot location:

+ Satellite image information: spectral average value, spectral standard deviation of each type of image created after processing according to different window size (KTCS) filter images.

+ Non-image information: elevation, slope and exposure values from ASTER GDEM.

3.2.3. Research methods to build a model for determining forest wood volume

(1) - Exploring the relationship between each independent variable x_i (image and non-image variables) and dependent variable y (forest wood volume - M) and among independent variables.

Exploring the relationship between each x_i and y to establish the morphology of the relationship between them, eliminating the variables x_i not related to y, is done by 4 types of functions: Y = a + b*X; Y = a + b*Ln (X); Ln (Y) = a + b*X; Ln (Y) = a + b*Ln (X).

These relationships are evaluated by testing the significance of the correlation coefficient (r) by the standard t with the hypothesis: H_0 does not exist the relationship between xi and y.

After selecting the x_i related to y (forest wood volume - M), conduct exploration of the relationship between the x_i with each other. If the correlation coefficient between the two independent variables (r>0,7) will retain x_i with a better relationship with forest wood volume and remove x_i with a worse relationship with forest wood volume from the list of independent variables.

(2) - Selecting the algorithm used for building models: In this study, using: multivariate regression models and non-parametric algorithms: K-NN, ANN, RF to build determine model forest wood volume. In which:

With multivariate regression function: the determination model of forest wood volume is built in 4 forms:

Type (3.1) $$Y = b_0 + b_1x_1 + b_2x_2 + ... + b_px_p$$

Type (3.2) $$Y = b_0 + b_1Ln (x_1) + b_2Ln (x_2) + ... + b_pLn (x_p)$$

Type (3.3) $$Ln (Y) = b_0 + b_1x_1 + b_2x_2 + ... + b_px_p$$

Type (3.4) $$Ln (Y) = b_0 + b_1Ln (x_1) + b_2Ln (x_2) + ... + b_pLn (x_p)$$

The multivariate constructions must meet the following requirements: (1) The equation exists in accordance with the standard F; The coefficients of the equation exist according to the standard t. Based on the errors of the models to choose the best equation.

After using multivariate functions to build a defined model of forest wood volume and selecting the best model, use the input variables of these best models to test the building of the model define forest wood volume by 3 non-parametric algorithms: k-NN; ANN; RF via Weka software.

(3) Divide the data to build and verify the model: the plots is divided into 2 parts: Randomly selected according to the space 33% (71 Plots) used to verify the model; The remaining 143 plots are used to construct the determined model of forest wood volume.

(4) Building and evaluating errors of the models: Using 143 plots with the following data: forest wood volume at the plot position, spectral mean indicators, spectral standard errors, topographic conditions ... to build the models of determining forest wood volume by: multivariate regression model, algorithms K-NN, ANN, RF. Each model was calculated for error types: Mean absolute error (MAE), Mean relative error (MAE%), Root Mean

Square Error (RMSE), Root Mean Square Error percent (RMSE%).

(5) Select the optimal model: the model has 4 smallest error values (MAE, MAE%, RMSE, RMSE%). In particular, priority is given to selecting models with the smallest RMSE%, followed by the smallest MAE%, the smallest RMSE and the smallest MAE.

3.2.4. Combine satellite imagery with inventory plot boundaries to develop a determination model forest wood volume

In order to limit the disadvantages of the method of taking values on the image according to the window size filter of satellite image at the boundary position, reducing the effect of pixel noise on the RADAR image, while still obtaining the typical value on the image to each pixel, after select the best forest wood volume defined model with input variables extracted from images according to a certain window size. The author considers the part of the inventory on that window size to be a homogeneous unit. From there, the method of extracting information on images is performed as illustrated in Figure 3.2.

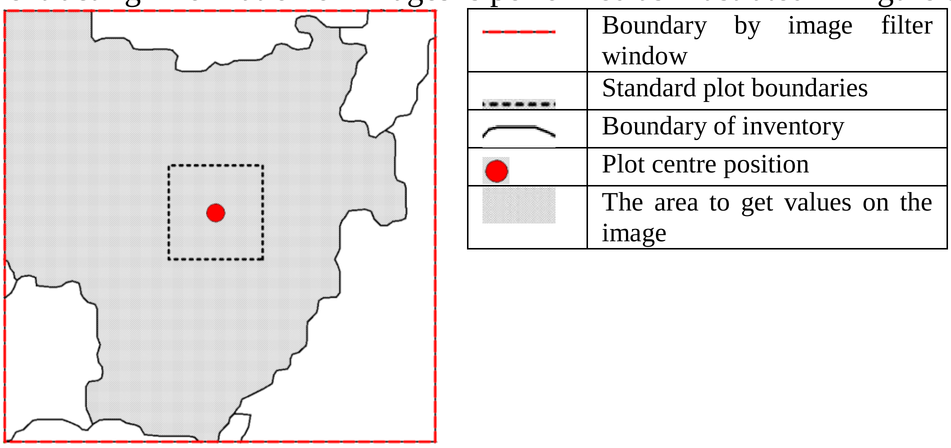

Figure 3.2. Method of extracting information combining filtered image with forest plot

3.2.5. Method of verifying the determination model forest wood volume

Using 33% of the standard plots (71 Plots) independently not participating in building the model to verify the optimal model. The verification errors of the models are calculated similarly to the model construction errors.

IV. RESEARCH RESULTS AND DISCUSSION

4.1. Determine the forest wood volume at the location of the plots and explore the relationship between the forest wood volume and the variables from satellite image and non-image

4.1.1. Determine forest wood volume at the location of the plots

Results of data collection, calculation of forest wood volume at the location of plots and plots division to build and verify the models in Table 4.1.

Table 4.1. General information about the plots used in research

Purpose	OTC Number (n)				M_{min} m³/ha	M_{max} m³/ha	M_{ave} m³/ha
	Total	<100 m³/ha	100-200 m³/ha	> 200 m³/ha			
Total	**214**	**86**	**98**	**30**	**11.8**	**301.6**	**123.3**
Explore the relationship and model building	143	57	66	20	11.8	290.8	123.8
Verify the models	71	29	32	10	20.6	301.6	122.5

4.1.2. Investigate the relationship between forest wood volume and satellite image and non-image variables

The results selected 54 variables from satellite image (33 from LANDSAT-8; 21 from ALOS-2/PALSAR-2) and 1 non-image variable (slope -DOC) related to forest wood volume. Absolute value the number r of each independent variable from the satellite image with forest wood volume varies from 0.19 to 0.64. In which: LANDSAT-8 image ranges from 0.23 to 0.62; ALOS-2/PALSAR-2 images range from 0.19 to 0.64. Show that, even though there exists a relationship between variables from image and non-image with forest wood volume, this relationship is only moderate, relatively tight, even some variables are related to forest wood volume at a weak level.

4.1.3. Exploring the relationship between independent variables and choosing input variables to develop models

Investigating the relationship between the independent variables shows: Between NDVI and PC1 on LANDSAT-8 image; between HH and HV on ALOS-2/PALSAR-2 images at each window size of satellite image filter always exists a very close relationship. Between the HV channel, HH of ALOS-2/PALSAR-2 image and the LANDSAT-8 image channel does not exist a close relationship. From there, input variables for models building are selected as follows:

- LANDSAT-8 image: PC1, PC2, DOC; NDVI, PC2, DOC.
- ALOS-2/PALSAR-2 image: HH, DOC; HV, DOC.
- Combining ALOS-2/PALSAR-2 image and LANDSAT-8 image: NDVI, PC2, HV, DOC; NDVI, PC2, HH, DOC; PC1, PC2, HV, DOC; PC1, PC2, HH, DOC.

4.2. Building a model to determine forest wood volume by multivariate regression

4.2.1. Model building with LANDSAT-8 image

As a result, 88 models to identify forest wood volume with 2 cases of input variables were established: (1) PC1, PC2, DOC; (2) NDVI, PC2, DOC according to 11 results of image filtering and 4 types of multivariate regression functions (Forms 3.1 to 3.4). Equations and coefficients of equations all exist statistically. It can be shown that LANDSAT-8 images can be used to develop forest wood volume determination models for evergreen broadleaf natural forest object in the study area.

At window size 1x1, the error of models are the largest. When the window size satellite image filter increased from 3x3 to 21x21, the error of the models decreased. In particular, the errors decreased sharply when filtered infrastructure increased from 3x3 to 11x11, when window size filtered images from 13x13 onwards, the errors of the models decreased unsignificantly.

Different types of equations affect the relative error of the model. Equation form (3.3) and (3.4) has advantages over two types of equation (3.1) and (3.2).

The error of the best model selected when the input variable is PC1, PC2, DOC is smaller than the best model selected when the input variable is NDVI, PC2, DOC.

Using LANDSAT-8 image to build forest wood volume defined models for evergreen broadleaf natural forest in Dak Nong with errors: MAE: 38 ÷ 43 m^3/ha; MAE%: 35% ÷ 41%; RMSE: 53 ÷ 56 m^3/ha; RMSE%: 51% ÷ 62%. From there, choose 2 models:

$$\mathbf{M = EXP\ (0.0001 * NDVI13TB - 0.3569)} \quad (4.1)$$

Error: MAE = 41 m^3/ha; MAE% = 39%; RMSE = 55 m^3/ha; RMSE% = 57%.

$$\mathbf{M = EXP\ [4{,}82644 * Ln\ (NDVI13TB) - 47{,}498]} \quad (4.2)$$

Error: MAE = 41 m^3/ha; MAE% = 39%; RMSE = 56 m^3/ha; RMSE% = 56%.

To identify forest wood volume from LANDSAT-8 image for evergreen broadleaf natural forest type in Dak Nong.

4.2.2. Model building with ALOS-2/PALSAR-2 image

For ALOS-2/PALSAR-2 images, 88 models of forest wood volume have been developed with 2 cases of input variables: (1) HV, DOC; (2) HH, DOC according to 11 results of filtering images and 4 types of multivariate regression functions. The results show that the equations and coefficients of the equation are statistically existent. It can be shown that ALOS-2/PALSAR-2 images can be used to develop forest wood volume identification models for evergreen broadleaf natural forest type in Dak Nong province.

Equation form (3.3) and (3.4) is better than two equation (3.1) and (3.2) when used to build forest wood volume determination models from ALOS-2/PALSAR-2 images in the study area assist.

Considering in the same window size filter images and equation form, the error of determining forest wood volume when the input variables are HV, DOC value is always smaller than the input variables are HH, DOC values.

Errors of the models for determining forest wood volume from ALOS-2/PALSAR-2 images: MAE: 35 ÷ 41 m^3/ha; MAE%: 37% ÷ 45%; RMSE: 46 ÷ 54 m^3/ha; RMSE%: 59% ÷ 77%. From there, choose 2 models:

M = EXP (0.000241 * HV_21TB + 0.019589 * DOC - 4,535) (4.3)

Error: MAE = 35 m^3/ha; MAE% = 37%; RMSE = 46 m^3/ha; RMSE% = 59%.

M = EXP [8,629208 * Ln (HV_21TB) + 0,129567 * Ln (DOC) - 86,457] (4.4)

Error: MAE = 35 m^3/ha; MAE% = 37%; RMSE = 46 m^3/ha; RMSE% = 59%.

To determine forest wood volume from ALOS-2/PALSAR-2 image for the evergreen broadleaf natural forest type in Dak Nong.

4.2.3. Building a model that combines LANDSAT-8 and ALOS-2/PALSAR-2 images

Combining ALOS-2/PALSAR-2 with LANDSAT-8, built 176 models to determine forest wood volume, the results show that all 176 equations and coefficients of each equation are statistically exist. Two types of equations (3.3) and (3.4) are better than two types of equations (3.1) and (3.2) when used to develop models to determine forest wood volume in the region.

The selected optimal forest wood volume determination models are suitable for each input variable case when combining 2 types of LANDSAT-8 and ALOS-2/PALSAR-2 images: MAE: 28-32 m^3/ha; MAE%: 27-32%; RMSE: 38-42 m^3/ha and RMSE%: 39-46%. In particular, the two best models are the model (4.5) and (4.6).

M = EXP (0,00020 * HV13TB + 0,00094 * PC1_13TB - 9,0454) (4.5)

Error: MAE = 28 m^3/ha; MAE% = 27%; RMSE = 38 m^3/ha; RMSE% = 39%.

M = EXP [7,33400 * Ln (HV11TB) + 6,00097 * Ln (PC1_11TB) - 125.44] (4.6)

Error: MAE = 29 m^3/ha; MAE% = 27%; RMSE = 39 m^3/ha; RMSE% = 39%.

4.2.4. Select and verify models defining forest wood volume

The optimal model built by combining two types of images together has smaller types of errors than the optimal model built for each type of image. Since then, the two models with the regular equations (4.5) and (4.6) are the two best models to determine forest wood volume for evergreen broadleaf natural forest type in Dak Nong province.

Using two models with regular equations (4.5) and (4.6) to determine forest wood volume at the location of independent plots not participating in model building (71 plots) and calculating the verification errors of the model , the results of determining of forest wood volume error to each pixel of the two models when achieved: MAE: 25 m^3/ha; MAE%: 29%; RMSE: 32 m^3/ha; RMSE%: 47% (model 4.5) and 48% (model 4.6). In particular, the model (4.5) is simpler model (4.6), there are the difference between MAE%, RMSE% of the model and verification are: $\Delta_{MAE\%}$ = 2%; $\Delta_{RMSE\%}$ = 9%.

4.3. Building models to determine forest wood volume by non-parametric algorithms

Using the variables of the regular equation (4.5) and (4.6) as an input variable for building 6 models to determine forest wood volume by non-parametric algorithms, construction results, and verification of comparative models with the optimal equation constructed by the multivariate regression function shows that: Although the difference in error between the models with the same input variables, the difference in applied algorithms is not large, but the models built by the regression function multivariate rules or ANN algorithms always have similar errors and are lower than models built using two algorithms K-NN and RF. Therefore, when using LANDSAT-8 and ALOS-2/PALSAR-2 images to determine wood volume for evergreen broadleaf natural forest type in Dak Nong province, it is recommended to use multivariate regression functions or ANN algorithms.

4.4. Combine satellite imagery with forest plot boundaries to build models determining of forest wood volume

Based on 4 types of errors: models built by multivariate regression for the lowest types of errors, next to the model built by ANN algorithm, model built using RF algorithm, model built by K-NN algorithm for the largest types of errors. Inside:

- Model of construction using multivariate regression with errors: MAE = 25 m^3/ha; MAE% = 25%; RMSE = 33 m^3/ha; RMSE% = 35%. Error checking the model: MAE = 21 m^3/ha; MAE% = 24%; RMSE = 29 m^3/ha; RMSE% = 41%. The difference between MAE%, RMSE% of the model and verified in turn is: $\Delta_{MAE\%}$ = 1% and $\Delta_{RMSE\%}$ = 6%.

- ANN construction model has the error: MAE = 25 m^3/ha; MAE% = 26%; RMSE = 33 m^3/ha; RMSE% = 39%. Error of checking the model is: MAE = 20 m^3/ha; MAE% = 23%; RMSE = 26 m^3/ha; RMSE% = 38%. The difference between MAE%, RMSE% of the model and verification are: $\Delta_{MAE\%}$ = 3% and $\Delta_{RMSE\%}$ = 1%, respectively.

The optimal model selected in this case is the model:

$$\mathbf{M = EXP\ (0.00022 * HV\ (K)\ TB + 0.00096 * PC1\ (K)\ TB + 0.02024 * DOC\ (K)\ TB - 10{,}191)} \quad (4.7)$$

Under the same conditions, when using the information extraction method according to the window size filter, the best forest wood volume determination model (model 4.5), giving the verification error: RMSE = 32 m^3/ha; MAE = 25 m^3/ha; MAE% = 29%; RMSE% = 48% are both larger than the errors of the best forest wood volume determination model when combining images with the inventory plot boundaries (model 4.7). It is proved that using a uniform calculation unit as a forest inventory plot on the window size 13x13 image infrastructure has increased the accuracy of the model to identify wood volume for evergreen broadleaf natural forest type in the studied area.

4.5. Select, adjust and evaluate determination model of forest wood volume

4.5.1. Select and adjust determination models of forest wood volume

The study has developed forest wood volume determination models for evergreen broadleaf natural forest type in Dak Nong province from each type of LANDSAT-8, ALOS-2/PALSAR-2 images and combined two types of images by two extraction of infomation methods. values from photos: (1) Window size of filtered; (2) Dimensions are the intersection of the satellite image filter space with the boundary of the forest plots. The two best models for forest wood volume determination that correspond to the two image filtering methods are those with regular equations (4.5) and (4.7).

Although the selected models have met the statistical criteria and have the highest correlation coefficients, the lowest errors. But due to the complex relationship between forest wood volume and satellite image value and non-image value, conventional mathematical functions may not be able to correctly simulate this relationship over the all forest type, but only on each interval of wood volume.

Empirical data show that: M is positively related to: $HV13_{TB}$, $PC1\_13_{TB}$ (model 4.5) and $HV(K)_{TB}$, $PC1(K)_{TB}$ (model 4.7). In order to make full use of the value of the input variables of the model, corresponding to the model (4.5) the author uses the variable $\sqrt{HV13_{TB} * PC1\_13_{TB}}$, the model (4.7) the author uses the variable $\sqrt{HV(K)_{TB} * PC1(K)_{TB}}$ to build two sub-models.

$$\mathbf{M = EXP\ [(-12391) / 731.94]\sqrt{HV13_{TB} * PC1_{13_{TB}}}} \quad (4.8)$$

Model available: MAE = 39 m^3/ha; MAE% = 34%; RMSE = 56 m^3/ha and RMSE% = 48% are both larger than the corresponding values of the model (4.5).

$$\mathbf{M = EXP\ [(-12452) / 728.91]\sqrt{HV(K)_{TB} * PC1(K)_{TB}}} \quad (4.9)$$

Error of model: MAE = 36 m^3/ha; MAE% = 31%; RMSE = 53 m^3/ha and RMSE% = 43% larger than the corresponding values of the model (4.7).

Although the two models (4.8) and (4.9) both have larger error types than the corresponding models (4.5) and (4.7) when calculating the forest type, the results are:

- In the actual forest wood volume range (M_{TT}) less than 200 m^3/ha between the model's forest wood volume (4.8) and (4.9) with the average of actual forest wood volume difference is significantly lower than this difference between the model (4.5) and (4.7).
- In the actual forest wood volume range greater 200 m^3/ha: calculated forest wood volume from the model (4.5) and model (4.7) are smaller than the average of actual forest wood volume. Meanwhile, calculated reserves from the model (4.8) and the model (4.9) tend to be larger than the average actual forest wood volume.

In order to limit the difference between average of actual forest wood volume and average of the model's forest wood volume according to range of the actual forest wood volume , by empirical results, the author proposes a model to ditermine wood volume of evergreen broadleaf natural forest type in Dak Nong province as follows:

- The model ditermining forest wood volume according to the window size filter images is a combination of two principal equations: (4.5) and (4.8) with the principle of association:

$$M = \begin{cases} M_{LT}\ (4.8) & \text{Within the conditions of } M_{LT}\ (4.5) \leq 200\ m^3/ha \\ M_{LT}\ (4.5) & M_{LT}\ (4.5) > 200\ m^3/ha \end{cases} \quad (4.10)$$

In which: M_{LT} (4.5), M_{LT} (4.8) are the forest wood volume calculated according to the regular equations (4.5) and (4.8).

- The model ditermining forest wood volume according to the window size filter images delivered with the forest plot is a combination of two principal equations (4.7) and (4.9) with the combination principle:

$$M = \begin{cases} M_{LT}\ (4.9) & \text{Within the conditions of } M_{LT}\ (4.7) \leq 200\ m^3/ha \\ M_{LT}\ (4.7) & M_{LT}\ (4.7): 200 \div 250\ m^3/ha \\ \dfrac{(M_{LT}(4.7) + M_{LT}(4.9))}{2} & M_{LT}\ (4.7) > 250\ m^3/ha \end{cases} \quad (4.11)$$

In which: M_{LT} (4.7), M_{LT} (4.9) are the forest wood volume calculated according to the regular equations (4.7) and (4.9)

The results of determining errors of forest wood volume determination models (4.10) and (4.11) are shown in Table 4.2.

Table 4.2. Results of determining errors of determination models forest wood volume (4.10) and (4.11)

Model name	Type of error	Type/ Forest status	Error					
			MAE	MAE%	RMSE	RMSE %	MAE_{max}	$MAE\%_{Max}$
(4.10)	Model	**Forest type**	**33**	**31**	**43**	**44**	**116**	**192**
		Poor	29	44	40	61	100	192
		medium	33	23	43	29	104	68
		Rich	40	17	54	23	116	56
	Verify	**Forest type**	**35**	**35**	**44**	**49**	**137**	**201**
		Poor	26	44	31	61	68	201
		Medium	44	32	54	43	137	136
		Rich	34	15	41	19	68	33
(4.11)	Model	**Forest type**	**32**	**29**	**44**	**41**	**172**	**181**
		Poor	27	40	37	55	102	181
		medium	28	20	39	27	111	65
		Rich	55	24	69	30	172	79
	Verify	**Forest type**	**30**	**29**	**43**	**43**	**168**	**161**
		Poor	24	40	32	55	94	161
		medium	31	23	42	33	122	122
		Rich	44	18	65	26	168	66

The relationship between: (1) Actual forest wood volume at the sample plots and the theoretical forest wood volume of the model 4.10; (2) Actual forest wood volume at the sample plots and the theoretical reserves of model 4.11 are shown in Figures 4.1 and 4.2 respectively.

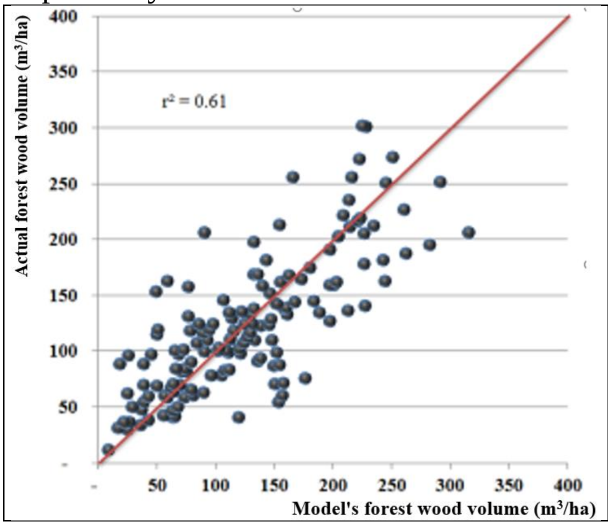

Figure 4.1. Relationship between Actual forest wood volume in the sample plots and the model's forest wood volume (4.10)

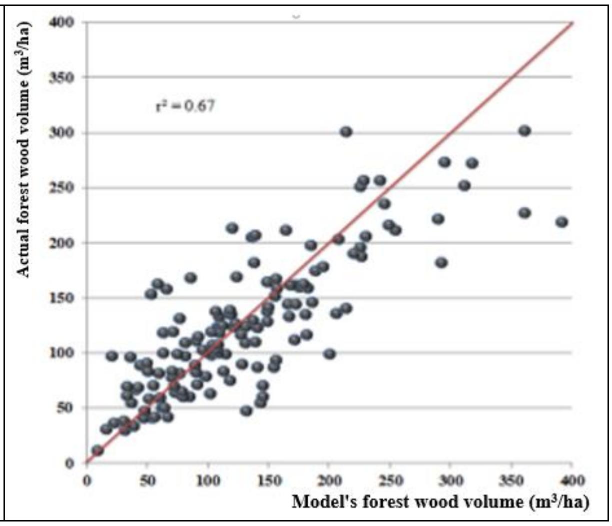

Figure 4.2. Relationship between Actual forest wood volume in the sample plots and the model's forest wood volume (4.11)

4.5.2. Evaluate the determination models of forest wood volume

To evaluate the models (4.10) and (4.11), the author allocates Actual forest wood volume and model's forest wood volume at the locations where the plots are used to build and verify the model into one of the three base forest status according to Actual forest wood volume (M_{TT}) (Poor Forest (M_{TT}≤100 m^3/ha); Medium Forest (M_{TT}: 100 ÷ 200 m^3/ha); Rich Forest (M_{TT} ≥200 m^3/ha)). In each forest status: (1) calculate the percentage of points with Mean absolute error (MAE): <20 m^3/ha; <40 m^3/ha; <60 m^3/ha; <80 m^3/ha and ≥80 m^3/ha; (2) calculate the percentage of points with Mean relative error (MAE%): <20%; <40%; <60%; <80% and ≥80%. The results are shown in tables 4.3 and 4.4.

4.5.2.1. Model (4.10) - Model 1

Table 4.3. Percentage of model points and model verification (4.10) according to the absolute and relative error threshold values

Classified by	**Type/ Status the forest**	**Percentage of control points based on error value threshold (unit:%)**									
		Threshold of Mean absolute error value (m^3/ha)					Threshold of Mean relative error value (%)				
		<20	<40	<60	<80	≥80	<20	<40	<60	<80	≥ 80
Model	**Type the forest**	**44**	**71**	**80**	**90**	**10**	**43**	**75**	**91**	**96**	**4**
	Poor	54	74	82	91	9	26	61	82	89	11
	Medium	39	71	80	91	9	48	82	95	100	0
	Rich	30	60	75	85	15	70	90	100	100	0
Verify	**Type the forest**	**34**	**65**	**85**	**93**	**7**	**42**	**72**	**83**	**92**	**8**
	Poor	45	72	97	100	0	34	59	79	86	14
	Medium	22	59	75	84	16	41	75	81	94	6
	Rich	40	60	80	100	0	70	100	100	100	0

4.5.2.2. Model (4.11) - Model 2

Table 4.4. Percentage of model points and model verification (4.11) according to the absolute and relative error threshold values

Classified by	**Type/ Status the forest**	**Percentage of control points based on error value threshold (unit:%)**									
		Threshold of Mean absolute error value (m^3/ha)					Threshold of Mean relative error value (%)				
		<20	<40	<60	<80	≥80	< 20	<40	<60	<80	≥80
Model	**Type the forest**	**47**	**72**	**84**	**91**	**9**	**50**	**76**	**90**	**97**	**3**
	Poor	53	77	86	93	7	32	65	81	91	9
	Medium	52	77	88	92	8	67	82	95	100	0
	Rich	15	40	65	80	20	50	85	95	100	0
Verify	**Type the forest**	**46**	**72**	**86**	**94**	**6**	**52**	**76**	**89**	**92**	**8**
	Poor	55	83	93	97	3	38	66	79	83	17
	Medium	41	66	84	94	6	63	81	97	97	3
	Rich	40	60	70	90	10	60	90	90	100	0

4.6. Procedure for ditermination forest wood volume from satellite images in Dak Nong province

4.6.1. The process of determining of forest wood volume according to model 4.10

The process of determining forest wood volume according to model 4.10 is illustrated in Figure 4.3. Specifically:

(1) - Input data to perform the procedure: LANDSAT-8 image; ALOS-2/PALSAR-2 image have rectified geometry of map projection grid, UTM coordinate system; Boundary

of evergreen broadleaf natural forest plots in Dak Nong province; SRTM DEM digital elevation model.

(2) - Data processing: Transfer channels (from 1 to 8) LANDSAT-8 image; HV image channel ALOS-2/PALSAR-2 from UTM system to VN2000 system; Correct the influence of terrain on the LANDSAT-8 images; Create LANDSAT-8 multi-spectral channel with spatial resolution of 15m; Building main components image of each LANDSAT-8 image from multi-spectral channels with resolution of 15m; Convert HV channel to 15m spatial resolution by Bi-linear interpolation method; Convert digital number of the main component channel PC1 on LANDSAT-8 image, HV channel on ALOS-2/PALSAR-2 image to the same 12-bit by Bi-linear interpolation method.

(3) - Filter images (PC1 on LANDSAT-8 image; HV on ALOS-2/PALSAR-2 image) with window size 13x13, determine forest wood volume (M) to each pixel of image by the formula (4.12).

$$M = \begin{cases} EXP[(\sqrt{HV13_{TB} * PC1_{13_{TB}}} - 12391)/731,94] & EXP(0,00020 * HV13_{TB} + 0,00094 * PC1\_13_{TB} - 9,0454) \leq 200 \\ EXP(0,00020 * HV13_{TB} + 0,00094 * PC1\_13_{TB} - 9,0454) & EXP(0,00020 * HV13_{TB} + 0,00094 * PC1\_13_{TB} - 9,0454) > 200 \end{cases} \quad (4.12)$$

(4) - Identify forest wood volume to each forest plot: Overlay the forest plot boundary on the identified forest wood volume image to each pixel and determine forest wood volume for each forest plot.

(5) - Applying the process of determining forest reserves: The total reserve of evergreen broadleaf natural forest in Dak Nong Province based on the model (4.10) is: 18,839,453 m^3 (18,800,000 m^3). According to the forest inventory results at the same time, the total reserve of the forest type is 20,500,000 m^3. Comparison of results of determination of forest wood volume by model and according to the results of the forest inventory shows that: The total reserve calculated by the model is lower than the inventory results: 1,700,000 m^3 (8.3%).

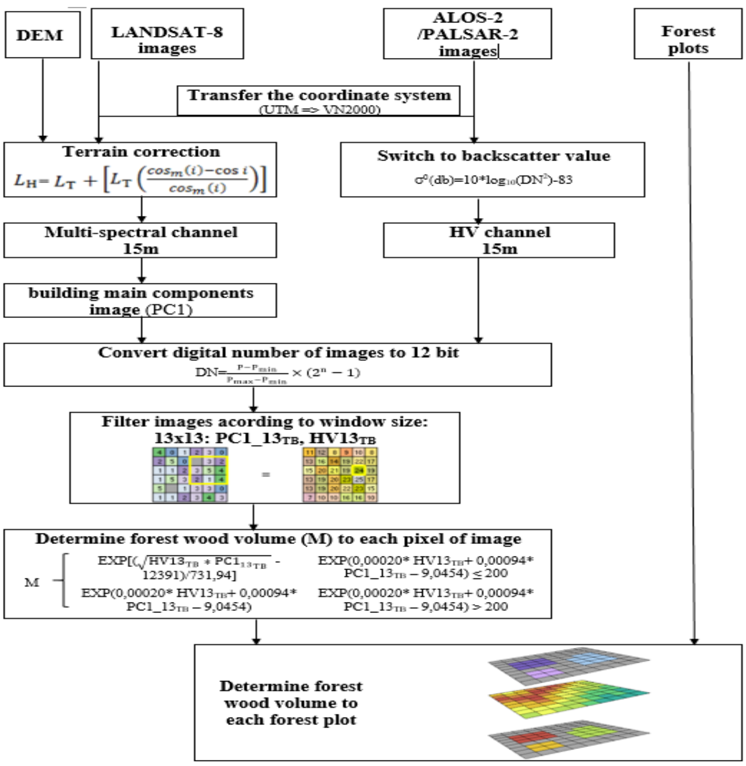

Figure 4.3. Process diagram for determining forest wood volume according to model 4.10

4.6.2. The process of determining of forest wood volume according to model 4.11

The process of determining forest wood volume according to model 4.11 is illustrated in Figure 4.4. Specifically:

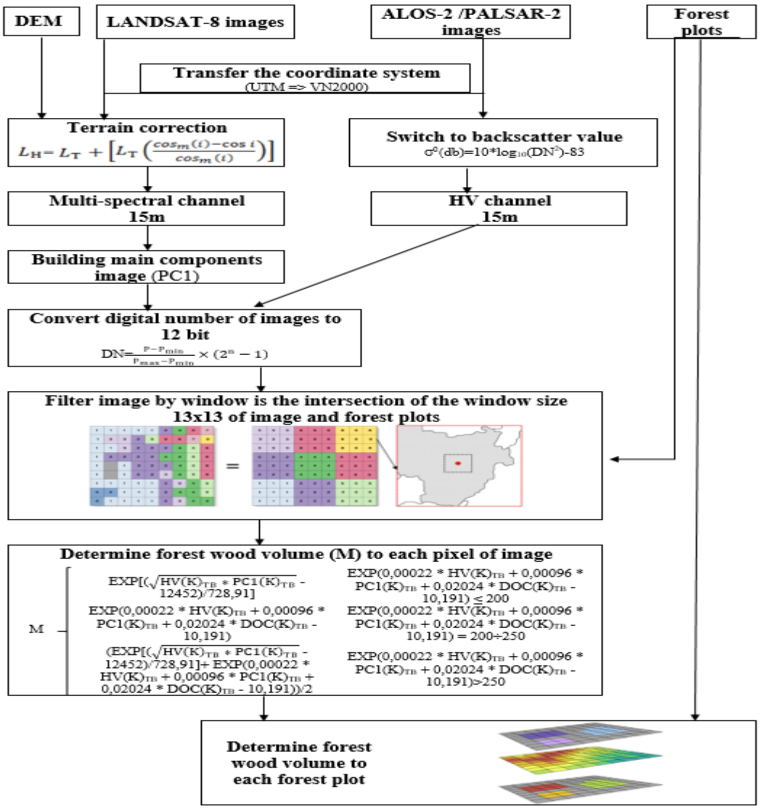

Figure 4.4. Process diagram for determining forest wood volume according to model 4.11

(1) - Input data required to perform the procedure: Similar to input data required to perform the procedure for determining forest wood volume according to the model 4.10.

(2) - Data processing: similar to the model 4.10.

(3) - Filter images, define forest wood volume to each pixel of image: Number the forest plots from 1 to n (n is the total number of forest plots) and convert the plots layer from vector form to raster form with the attribute field is order (TT), spatial resolution of 15m (called layer Ras1); Create a grid layer covering the object of the study, 195m grid size and number the grid cells from n + 1 to m (m-n-1 is the total number of grid cells) and convert the grid layer from vector form to raster format with field attribute is order (TT), spatial resolution 15m (called layer Ras2); Multiply the two layers Ras1 and Ras2 together, creating a new raster layer (called Ras); Filter PC1 channel on LANDSAT-8 image, HV channel on ALOS-2/PALSAR-2 image by Ras layer; Determine forest wood volume to each pixel by the formula (4.13).

$$\mathbf{M} = \begin{cases} EXP[(\sqrt{HV(K)_{TB} * PC1(K)_{TB}} - 12452)/728{,}91] & EXP(0{,}00022 * HV(K)_{TB} + 0{,}00096 * PC1(K)_{TB} + 0{,}02024 * DOC(K)_{TB} - 10{,}191) \leq 200 \\ EXP(0{,}00022 * HV(K)_{TB} + 0{,}00096 * PC1(K)_{TB} + 0{,}02024 * DOC(K)_{TB} - 10{,}191) & EXP(0{,}00022 * HV(K)_{TB} + 0{,}00096 * PC1(K)_{TB} + 0{,}02024 * DOC(K)_{TB} - 10{,}191) = 200 \div 250 \\ (EXP[(\sqrt{HV(K)_{TB} * PC1(K)_{TB}} - 12452)/728{,}91] + EXP(0{,}00022 * HV(K)_{TB} + 0{,}00096 * PC1(K)_{TB} + 0{,}02024 * DOC(K)_{TB} - 10{,}191))/2 & EXP(0{,}00022 * HV(K)_{TB} + 0{,}00096 * PC1(K)_{TB} + 0{,}02024 * DOC(K)_{TB} - 10{,}191) > 250 \end{cases} \quad (4.13)$$

(4) - Determining forest wood volume to each forest plot: Overlapping the forest plot boundary above the forest wood volume image and determine forest wood volume for each forest plot.

(5) Applying the determination procedure forest wood volume: Total forest wood volume of evergreen broadleaf natural forest in Dak Nong Province calculated according to the model (4.11) is: 19,899,336 m3 (19,900,000 m3). Comparison of results of determination of forest wood volume by model and forest inventory results shows that: Total forest wood volume calculated by the model is lower than the inventory results: 600,000 m3 (equivalent to 2.9%). So:

+ When applying the model (4.11) to determine wood volume for the evergreen broadleaf natural forest in the area, the total forest wood volume calculated by the model is smaller than 3% compared to the total reserve of forest inventory results;

+ Between total forest wood volume calculated according to the model (4.10) and (4.11), the total forest wood volume according to the model (4.11) is closer to the forest inventory result. This demonstrates that using model (4.11) to determine wood volume for evergreen broadleaf natural forest in Dak Nong is better than the model (4.10).

The wood volume distribution map of evergreen broadleaf natural forest was built according to the model 4.11 to each forest plot in Dak Nong province is shown in Figure 4.5.

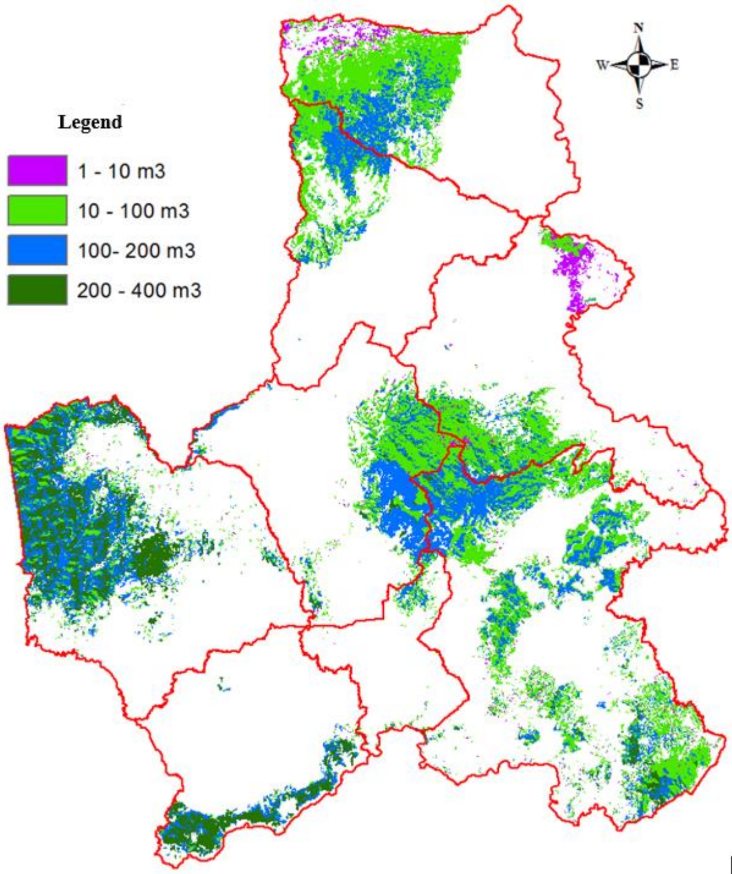

Figure 4.5. Map of distribution of wood volume of evergreen broadleaf natural forest built according to model 4.11 to each forest plot in Dak Nong province

4.7. Discuss

Technical research to determine forest wood volume from satellite images is the study: (1) The technique of selecting satellite images used to determine forest wood volume: based on the nature of the image in accordance with the specific conditions in the area research; (2) Image processing techniques: require the level of image processing provided by the manufacturer and the required processing steps of users: terrain correction, spectrum conversion, image filtering ... before using used to determine forest wood volume; (3) Techniques to select variables from images based on experience combined with verification in a specific area; (4) Field-based forest wood volume identification technique on Plots; (5) The technique of selecting algorithms to build a model to determine forest wood volume based on experience combined with verification in a specific area; (6) The technique of evaluating the errors of the models and identifying the main technical factors that affect the accuracy of the determination model forest wood volume in the relationship between the required accuracy to achieve specific financial conditions, image processing methods and equipment according to the study area; (7) Technique for identifying forest wood volume to each pixel; (8) Technique for identifying forest wood volume to each forest plot. Since then, based on the research results, in specific conditions in Vietnam in general and Dak Nong in particular, the thesis focuses on discussing the following issues:

4.7.1. Selecting satellite images used to determine forest wood volume

There are many ways to classify remote sensing images, based on wavelength, divided into: optical image, RADAR image and LIDAR image. Of which: optical data is most widely used to determine forest reserves, followed by RADAR images and finally LIDAR images. This seems to be contrary to their data nature, because: LIDAR is a relatively complete source of information about a forest; RADAR images are only capable of providing some information compared to LIDAR images; Optical Images only have information on the top of the canopy, almost no three-dimensional information about the objects on the ground. Therefore, decisions about which materials to use now take into account document availability, costs, and complexity of the analysis rather than which system will provide the most comprehensive information on a Forest.

The selection of 2 types of materials: LANDSAT-8 and ALOS-2/PALSAR-2 is optimal in both aspects: (1) Research; (2) Practical applications in terms of our country.

- From a research perspective: LANDSAT-8 is an average satellite image of space resolution, provided free of charge with information on canopy of forests; ALOS-2/PALSAR-2 is a theoretical L-channel RADAR image that can bring information on branches and stems to study and develop forest wood volume identification models for Dak Nong province. Therefore, the selection of these two types of image materials is based on the nature of their data that can provide useful information for the determination of forest wood volume.

- In terms of application: LANDSAT-8 and the next generation LANDSAT-9 (expected to be operational from 2020) will provide long-term sources of optical satellite images, average spatial resolution. ALOS-2/PALSAR-2 images are taken in all weather conditions, so it is possible to provide photos at any location in our country over time. In particular, ALOS-2/PALSAR-2 images were recently created by combining one-year sequences of images at 25m resolution to provide users with free data delays of about 1 year. Given the opportunity to use this type of material in identifying forest wood volume. Therefore, there are always available sources of images used to build the model. This is a prerequisite to expand this research in other localities and forest types in order to improve the process and build models and programs that allow the identification of forest wood volume in our country.

On LANDSAT-8 images often have clouds and cloud shadows, if used to determine forest wood volume will affect the results. This is because the values in the image in cloudy areas and cloud shadows do not reflect the right objects on the ground. Determination forest wood volume does not require images to be constantly updated as for forest monitoring. Therefore, depending on the time it is necessary to determined forest wood volume some solutions to limit the effect of clouds and cloud shadows as follows:

- Choose images that are less cloudy, then remove cloudy areas before using them to determine forest wood volume: Based on the author's multi-year track results, LANDSAT-8 images are alway selected every year there is a small cloud cover (lt 5%) in Dak Nong but these images were only taken during the dry season. The author's of statistics for the scene code coded 124_052 (the main LANDSAT-8 scene in Dak Nong province) in the period of 2014-2018 shows that: a year have 19 images, but only 2 images have cover of cloud below 5% to reach nearly 13%. Thus, the number of images with cloud ratio (<5%) is much smaller than the total number of images obtained.
- Use multiple images in a period of time: because the clouds on the image are not fixed, resulting in different scenes, the location of the cloud is different. From there it is possible to select areas without clouds on different scenes to create images without clouds.

After selecting LANDSAT-8 images, based on the time to choose suitable ALOS-2/PALSAR-2 images so that these two types of documents will be taken in a certain period of time. The purpose of this selection is to minimize the influence of seasonal factors, the difference in the time between two types of image acquisition, the difference between the image collection time and the field data collection time... affect the result of determining forest wood volume.

4.7.2. Collect and calculate forest wood volume in the field

Data on forest wood volume determined through field Plots is a very important basis for building models that allow forest wood volume ditermination from satellite images. Some methods of collecting field data to determine forest wood volume include: (1) Cutting down trees and directly measuring; (2) Identify forest wood volume through measurement information without affecting forest objects; (3) Use available models or data, combined with existing survey data to identify forest wood volume.

The thesis uses the third method to identify forest wood volume in the field. After establishing the plots, measure the diameter and height of the trees in the plot by the second method. Using the volume chart of 2 factors made nationwide for evergreen broadleaf natural forest - the shape group 3 (Formed by the first method) to determine the volume of each individual tree, then determine the total volume of the trees in plot and forest wood volume at the location of plot. As such, the collection and calculation of forest wood volume in the field are feasible and ensure the most reliable conditions in the current closure of natural forests to carry out the research contents.

Investigating to have reserve information on sample plots often takes time, effort and expense. This condition is considered to be the most difficult for building a model to identify forest wood volume from future satellite images. However, today, when looking at the database of forest inventory in our country, it is found that there are a large number of forest Plots that have been investigated by many different topics and projects, they are distributed in most forested areas, it was investigated at different times. Only 2 projects: national forest inventory census period 2013 - 2016; Assessing and monitoring national forest resources, the period of 2016-2020 has investigated in detail tens of thousands of plots including positioning plots and temporary plots. This is a valuable source of data for research, construction as well as adjustment and evaluation of the accuracy of the determination models of forest wood volume. Therefore, to save time, money, effort ... it is necessary to study the method using survey data at existing Plots to build a model ditermation of forest wood volume. This research should focus on creating "banks" of plots in the ground and technology to exploit them to ditermined models forest wood volume with any photo scenes of different image materials. The most important is to identify forest wood volume at the location of the plots at different times. The solution to this problem is as follows:

- Inherit and build the plots database on ground surveys of key programs and projects, the most important are 2 projects: National Forest Inventory and Inventory Survey in 2013 - 2016; Assessment and monitoring of national forest resources, period 2016-2020.

- Develop algorithms and software to automatically detect forest changes at plots locations over time by satellite images: LANDSAT-8, SENTINEL-1, SENTINEL-2 ... If detected at the plot location there are variables In case of abnormal fluctuations (fluctuations due to other causes not due to forest growth), plot is removed from the sample library.

- Developing growth functions based on forest type and geographic region to apply time-based reserves determination to the remaining plots in the sample library.

4.7.3. Select variables from satellite images to build a forest wood volume-defined model

In the thesis, the author first uses professional knowledge and experience in the field of research to preliminarily determine the input variables from satellite images and non-images used to build forest wood volume-defined models for each type image, then use regression method to remove non-correlated variables, retain the variables that are related to forest wood volume through the correlation coefficient r. Research results with evergreen broadleaf natural forest type in Dak Nong show:

- From 6 variables on LANDSAT-8 image put to the test, select 3 variables were channel average values: NDVI, PC1, PC2 according to different image filtering have relationship with forest wood volume;

- From 4 variables on the ALOS-2/PALSAR-2 image put to the test, 2 variables were selected are the average of the back scatter channel: HH, HV according to different image filtering have relationship with forest wood volume;

- From 3 non-image variables put to the test, 1 variable selected (slope) (DOC) has a relationship with forest wood volume.

Only variables that are actually related to forest wood volume can be used to build a model to determine forest wood volume. In principle, at each window size of image filter can put all variables on the image and non-image are related for forest wood volume to build a model to determine forest wood volume. However, if the variables are related to each other, the results of constructing models are not stable. Therefore, the author explored the relationship between the independent variables used to build a model to determine forest wood volume corresponding to each of the window size. The result is between: PC1 and NDVI on LANDSAT-8 images; HH and HV on ALOS-2/PALSAR-2 images are always very closely related. Therefore, the author has selected the input variables to model for each type of image and combine 2 types of images:

- LANDSAT-8: models are built with 2 cases of input variables: (1) PC1, PC2, DOC; (2) NDVI, PC2, DOC according to the different image filters.

-ALOS-2/PALSAR-2: models are built with 2 cases of input variables: (1) HH, DOC; (2) HV, DOC according to the different image filters.

- Combining ALOS-2/PALSAR-2 and LANDSAT-8: models are built with 4 cases of input variables: (1) NDVI, PC2, HV, DOC; (2) NDVI, PC2, HH, DOC; (3) PC1, PC2, HV, DOC; (4) PC1, PC2, HH, DOC according to the different image filters.

Thus, the results of the study have identified the image and non-image variables that exist in relation with forest wood volume. Therefore, confirming the ability to use each type of satellite image LANDSAT-8 and ALOS-2/PALSAR-2 or combine them together in determining forest wood volume for evergreen broadleaf natural forest type in Dak Nong province.

4.7.4. Select the algorithm used to identify forest wood volume from satellite images

Many algorithms have been developed for determining forest wood volume from satellite images, they can be divided into two groups: parametric and non-parametric. In particular, parametric algorithms often use the form of relatively familiar equations should be used by many studies. Non-parametric algorithms often confuse users about how they work, so they are used less often. However, if it is necessary to build automated forest wood volume-determination models on a large scale, the application of non-parametric algorithms will be more advantageous because they can automatically set forest wood volume-defined models based on forest wood volume in the field and input satellite images, while using parametric algorithm need to rebuild the model.

In the thesis, after choosing the variables on the image and non-image that are related to forest wood volume, the author used 4 different algorithms to build models to determine forest wood volume. In particular, the regression function Multivariate rules are used to build and identify the best models for each image and combine the two types of images together. Then, using the input variables of these models to build the model to determine forest wood volume by 3 algorithms: ANN, RF, K-NN. The results show that, with the same input conditions, when using different algorithms, the model determines forest wood volume for different errors. The general trend is that models built by multivariate regression functions have the lowest errors and are nearly similar to those built with ANN algorithms, followed by construction models using RF algorithm, construction models by K-NN algorithm has the highest error.

Thus, the first step asserted: different algorithms affect the accuracy of the forest wood volume-determination model. From there, the thesis selects the optimal algorithm for determining wood volume of evergreen broadleaf natural forest type in Dak Nong province is using the multivariate regression function and ANN algorithm from the current common applied algorithms.

4.7.5. Error in determining forest wood volume from satellite images

4.7.5.1. Error in determining forest wood volume from LANDSAT-8 image

When using LANDSAT-8 image to build a forest wood volume -determination model, the error of the best forest wood volume-determination model is within the error threshold compared to the results of several studies in the world using the same LANDSAT image data.

In Vietnam, there has not been any published work on the problem of forest wood volume determination from LANDSAT-8 images in particular or LANDSAT in general, but compared to other studies on forest wood volume determination from other optical satellite images, the uncertainty was determined forest wood volume for each pixel of this study falls within the limits of the published work.

LANDSAT-8 image materials are available, provided free but the errors still high, the techniques and models in this study should only be applied to quickly determine the average reserves of evergreen broadleaf natural forest differs from similar properties.

4.7.5.2. Error in determining forest wood volume from ALOS-2/PALSAR-2 image

When using ALOS-2/PALSAR-2 image to build a model to determine forest wood volume, the errors of the models: MAE: 35-41 m^3/ha; MAE%: 37-45%; RMSE: 46-54 m^3/ha; RMSE%: 59-77%. In particular, the best model has: MAE_{MH}= 35 m^3/ha; MAE_{MH} (%) = 37%; $RMSE_{MH}$ = 46 m^3/ha; $RMSE_{MH}$ (%) = 59%.

The best $RMSE_{MH}$ (%) value is 59% within the error range of studies in the world on RADAR band L image data.

The absolute errors, the global mean errors of the optimal models using ALOS-2/PALSAR-2 images are all smaller than the optimal models using LANDSAT-8 images. It proves that the use of ALOS-2/PALSAR-2 images to determine wood volume for evergreen broadleaf natural forest in the study area is better than using LANDSAT-8 images.

Although the error of forest wood volume determination from ALOS-2/PALSAR-2 image is lower than that of LANDSAT-8 image, these error values are still high when determining forest wood volume for a specific pixel. Therefore, similar to the LANDSAT-8 image, the techniques and models in this study with ALOS-2/PALSAR-2 images can be replicated to determine the overall reserves of evergreen broadleaf natural forest differs from similar properties.

4.7.5.3. Error in determining forest wood volume for combining LANDSAT-8 and ALOS-2/PALSAR-2 images

Combining LANDSAT-8 image with ALOS-2/PALSAR-2 to build forest wood volume-determination model, the best models for errors: MAE from 28-32 m^3/ha; MAE% from 27-32%; RMSE ranges from 38-42 m^3/ha and RMSE% from 39-46%. In which, the error of the best model (4.11): MAE_{MH} = 32 m^3/ha; MAE_{MH} (%) = 29%; $RMSE_{MH}$ = 44 m^3/ha; $RMSE_{MH}$ (%) = 41%. These error values are all smaller when using only one image type. It proves that the combination of 2 types of images to build a model to identify reserves of evergreen broadleaf natural forest in the study area is better than using each type of image.

When combining 2 types of images with the inventory plot boundaries to build a model to determine forest wood volume, the best model for errors is: MAE_{MH} = 25 m^3/ha; MAE_{MH}(%) = 25%; $RMSE_{MH}$ = 33 m^3/ha; $RMSE_{MH}$(%) = 35%. These error values are all smaller when only two types of images are used but are not combined with the inventory lot boundary. In essence, the combination of inventory lot boundary, ALOS-2/PALSAR-2 and LANDSAT-8 images are the combination of three types of image materials: (1) High-resolution optical images used to delineate the plot (inventory boundaries of the inventory in the study area are delineated from SPOT-6 images); (2) RADAR band L image (ALOS-2/PALSAR-2) and (3) medium resolution optical image (LANDSAT-8) to build a model for identifying forest wood volume. It’s demonstrate that the use of optical images high-resolution learning to localize areas into relatively homogeneous plots. Then, using this circumference layer to extract spectral values on medium-resolution optical images and on RADAR band L images as input variables to build forest wood volume-defined models can be an effective solution to reduce errors of models. Because the limited high resolution satellite imagery is a large fluctuation value due to the shade of the canopy and the terrain's shadow, it causes errors for the computation model forest wood volume. However, this image data it is considered to be better than RADAR image and medium resolution optical image in delineating objects with relatively homogeneous state. Therefore, this combination has combined the advantages, eliminates the limitations of each image type, leading to increased forest wood volume identification efficiency.

In specific conditions in Dak Nong province, based on the results of the inventory and the updated forest change results, there is always a forest-type boundary map evergreen broadleaf natural but there is no boundary map of forest status. The reason is that the forest type does not change over time, but the forest status depends on forest wood volume, so it always changes over time, without forest wood volume parameters, the forest status will not be known. Therefore, using the traditional method (stratification and multiplication) to determine forest wood volume would have to assign the average of plots (M_{TB} = 123.8 m^3/ha) to all positions in the forest type. In this case, the error of determining forest wood volume to each pixel when verified reaches: MAE_{KC} = 48.3 m^3/ha; MAE_{KC}(%) = 58.7%; $RMSE_{KC}$ = 59.9 m^3/ha; $RMSE_{KC}$(%) = 97.7%. Therefore, the use of satellite images to identify forest wood volume according to the research method has significantly improved the errors compared to the conventional method used previously.

Although no studies have combined: ALOS-2/PALSAR-2 and LANDSAT-8 to determine forest wood volume, but compared to studies on similar types of images: the error of forest wood volume determination in this study lies in range of published studies.

Combining LANDSAT-8 and ALOS-2/PALSAR-2 images gives the determination that forest wood volume has acceptable tolerances for practical application in forest inventory.

CONCLUSION, SHORTCOMING AND RECOMMENDATION

1. Conclusion

Based on the achieved results, the thesis draws the following conclusions:

1) Research techniques for determine forest wood volume from satellite images include a series of work steps: Selection of satellite images; Processing satellite images; Selecting variables from the satellite images; Determining forest wood volume at the scene (Plots); Selecting algorithms to build models; Assessing the error of the model and identify the main technical factors affecting the accuracy of the model to determine the forest wood volume; Determining forest wood volume to each pixel; Determining forest wood volume to each forest plot.

2) Identify variables on image (LANDSAT-8; ALOS-2/PALSAR-2) and non-image exist in relation to forest wood volume, thereby confirming the ability to use each type or combination of two image types in determine wood volume for evergreen broad-leaved natural forest in Dak Nong province.

3) Select the optimal model to use: each type of images (LANDSAT-8; ALOS-2/PALSAR-2), combine 2 types of photos, combine 2 types of photos with forest plot boundaries to determine forest wood volume by Multivariate regression function.

4) With non-parametric algorithm, the model determining of forest wood volume built by ANN algorithm gives the lowest error, followed by the model built by RF algorithm, the model built by the K-NN algorithm for the biggest errors type.

5) Selecting the optimal algorithm for determination of wood volume for evergreen broad-leaved natural forest in Dak Nong Province is using multivariate regression functions and ANN algorithms from common applied algorithms.

6) Combining LANDSAT-8 and ALOS-2/PALSAR-2 images to build a model to determine wood volume for evergreen broad-leaved natural forest in Dak Nong better than using each image type.

7) Combining 2 types of images with the boundaries of the forest plots has the best forest wood volume-defined model compared to combining only two types of images or using only each image type.

8) Developed 2 processes: i) Determine forest wood volume according to the optimal model combining 2 types of images; ii) Determine forest wood volume according to the optimal model combining the two types of images with the forest plots to determine the wood volume of for evergreen broad-leaved natural forest from satellite images in Dak Nong.

9) Errors in the best model used to determine wood volume for evergreen broad-leaved natural forest in Dak Nong province from each type and combining 2 types of images

together to achieve:

- LANDSAT-8: MAE=38m^3/ha; MAE(%)=35%;RMSE=53m^3/ha; RMSE(%)=51%.
- ALOS-2/PALSAR-2: MAE = 35 m^3/ha; MAE(%) = 37%; RMSE = 46 m^3/ha; RMSE(%) = 59%.
- Combining 2 types of images: MAE = 33 m^3/ha; MAE (%) = 31%; RMSE = 43 m^3/ha; RMSE (%) = 44%.
- Combining 2 types of images with the inventory plot boundary: MAE = 32 m^3/ha; MAE(%) = 29%; RMSE = 44 m^3/ha; RMSE(%) = 41%.

This error is within the threshold of error compared to the results of studies in the world using the same image data.

2. Shortcoming

Although the thesis has completely addressed the contents and met the research objectives, the thesis still has some shortcomings:

- ALOS-2/PALSAR-2 images were taken during the dry season in the region (from September 2014 to January 2015) but the time difference and angle factors have not been studied and processed in Thesis.
- The area of plot using only one type of 1,000m^2 may not be large enough, leading to certain errors when determining forest wood volume from satellite images, especially for LANDSAT-8 images.

3. Recommendations

Extend research to build forest wood volume-determination models with different plot sizes to select the optimal field plot size to build forest wood volume-determination models for each image type.

Extend research for the same forest type in other areas, for other forest types to complete the method