RMIT UNIVERSITY

Extrapolating forest biomass dynamics over large areas using time-series remote sensing

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Huy Trung Nguyen

Bsc (Hons), Thai Nguyen University, Vietnam

Msc Environmental Science, Thai Nguyen University, Vietnam

School of Science

College of Science, Engineering and Health

RMIT University

February 2020

Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship

Trung Nguyen

26 February 2020

Abstract

Forest biomass, accounting for over 80% of global vegetation biomass, is considered a key factor in terrestrial ecology, atmospheric processes and the water and carbon cycles. Forest biomass has been recently recognised as a Global Climate Observing System (GCOS) Essential Climate Variable (ECV), which is an important input to the United Nations' Reducing Emissions from Deforestation and forest Degradation-plus (REDD+) program and Earth system models. Reducing carbon emissions from forest changes is one of the core requirements to mitigate the impacts of climate change on Earth. Consequently, monitoring forest biomass dynamics is an international concern which has attracted attention from government (at local, regional, national and international levels), academics and the general public. According to the Global Forest Resources Assessment 2015, deforestation and forest degradation have been persisting in tropical developing countries where demand for exploiting natural resources are high and significantly increasing. Thus, these countries urgently need a robust and cost-effective national forest biomass monitoring system that can support their policy-making processes that aim to protect ecosystem integrity in forests and reduce greenhouse gas emissions while simultaneously maintaining their social-economic development needs. While improving the quality of carbon reporting is needed, it is challenging for most developing countries due to their low capacities to perform national forest inventory on a regular basis. Forest inventory data may be available in these countries, but they are often out-of-date. Using remote sensing data, such as Landsat satellite imagery, is one of the most practical and cost-effective alternatives to enable developing countries to overcome this current challenge. Landsat satellites are unique as they have been creating the longest continuously-acquired, space-based and moderate-resolution data collection since 1972. The free access and use data policy of the Landsat archive since 2008 has revolutionized the use of Landsat data for worldwide forest research and monitoring activities, especially forest biomass monitoring.

This research first comprehensively reviewed the state and improvements of current approaches using Landsat time-series (LTS) for characterising forest biomass dynamics. This literature review indicated that the use of LTS not only enables production of spatially and temporally explicit estimates of biomass but also can improve the quality and accuracy of biomass models. Many innovative approaches for estimating forest biomass across space and time from LTS have been recently demonstrated. However, most of these methods have

been developed for areas that are supported by comprehensive forest inventories and/or Lidar datasets. Therefore, it is important to demonstrate an approach that is more possible for applications in developing countries where forest inventory data are measured for a single-time step which is often out-of-date.

This research develops a robust and consistent Landsat-based framework that can support developing countries improve their capacities in monitoring and reporting forest biomass and carbon stocks and changes across large areas. The framework is developed by utilising a 30-year annual time-series of Landsat images (1988-2017) and one-off inventory data, which are commonly available in developing countries. The study area comprised over 7.1 million ha of public forests in Victoria, south-eastern Australia. Although Victoria is not a country, its size / forest inventory scenario is similar to many developing countries, making it a good case study. LTS data were processed through several steps to produce a stack of cloud-free, annual mosaic composites. This dataset was then used as a foundation input in further analyses for characterising forest disturbance and recovery and estimating forest biomass dynamics across space and time.

In the first stage, LTS data were utilised for developing a robust approach for mapping forest disturbance and recovery at a landscape scale. Forest changes were detected through pixel-based change detection process using the LandTrendr temporal segmentation algorithm. A two-phase classification process was then developed using the Random Forest (RF) algorithm to predictively map disturbance and recovery levels (high, medium and low) and disturbance causal agents (including wildfire, planned burns, clear-fell logging, selective logging) for multiple detected disturbance events (both primary and secondary events). Model explanatory data included a range of trajectory-based change metrics derived from the LandTrendr analysis, while model training and validation data were derived from a human-interpreted reference dataset. In addition, a space-time data cube concept was introduced to simultaneously report on both newly detected disturbance events (detected disturbances) as well as events that have previously occurred but are ongoing (ongoing disturbances), which has been often under-reported. RF classification models obtained high overall accuracies (73-81%). The data cube analysis revealed that although annual disturbance area was dominated by newly detected disturbances, ongoing disturbances accounted for a considerable area (over 50% of newly detected disturbances). These results indicate the utility of LTS in accurately capturing and mapping forest disturbance and recovery, facilitating further analyses on biomass estimates.

The second stage of this research tested and compared different modelling approaches for estimating forest biomass using Landsat time-series and inventory data. This analysis used the outputs from the first stage (i.e., spectral change metrics, predicted disturbance and recovery levels and causal agents) in combination with data extracted from forest inventory field plots. In particular, 18 k-nearest neighbour (kNN) imputation models were tested to predict three aboveground biomass (AGB) variables (total AGB, AGB of live trees and AGB of dead trees). These models were developed using different distance techniques (RF, Gradient Nearest Neighbour (GNN), and Most Similar Neighbour (MSN)) and different combinations of response variables (model scenarios). Direct biomass imputation models were trained according to the biomass variables while indirect biomass imputation models were trained according to combinations of forest structure variables (e.g., basal area, stem density and stem volume of live and dead-standing trees). The results show that RF consistently outperformed MSN and GNN distance techniques across different model scenarios and biomass variables. The indirect imputation method generally achieved better biomass predictions than the direct imputation method. In particular, the RF-based kNN model trained with the combination of basal area and stem density variables was the most robust for estimating forest biomass. As the kNN imputation method is increasingly being used by land managers and researchers to map forest biomass, this analysis helps those using these methods to ensure their modelling and mapping practices are optimized.

The last stage presented a consistent approach for estimating forest AGB dynamics across space and time using LTS and single-date inventory data. This approach consisted of three components: (1) a modelling method for creating annual forest AGB maps from Landsat time-series and one-off inventory data; (2) evaluation of the robustness and transferability of applying a single model through time to estimate AGB dynamics; (3) a spatial and temporal analysis of AGB dynamics according to forest disturbance and recovery histories, from which to inform jurisdictions as to how these ecological changes impact AGB dynamics. These analyses were based on the findings of the first two stages. A RF-based kNN imputation model, which was defined as the most accurate method in the second stage, was developed to produce annual maps of AGB for 30 years (from 1988 to 2017 over 7.2 million ha of forests in Victoria, Australia). Annual predictions of AGB and its change were independently evaluated using multi-temporal Lidar data. These obtained relatively high accuracies, indicating the robustness and transferability over time of the developed modelling method. Temporal trends of AGB were analysed according to forest disturbance and recovery levels and causal agents (derived in the first stage) in order to understand how AGB responds to both natural and anthropogenic processes. Specifically, change metrics (e.g., AGB loss and gain, Years to Recovery - Y2R) were calculated at the pixel level to characterise the patterns of AGB dynamics resulting from forest changes. AGB change metrics showed that changes in AGB values associated with forest disturbance and recovery (decrease and increase, respectively) were captured by predicted maps. Results also indicated that AGB loss and Y2R varied across the states' biogeographic regions and were highly dependent on the level of disturbance severity (i.e., a greater loss and longer recovery time were associated with a higher severity disturbance).

The framework presented in this research has potential for application in different forest areas to support forest managers and policy makers to measure and report on forest biomass changes. This research focuses on providing a solution for developing countries, where only single-date (often out-of-date) and sparse inventory data are available, to improve their capacities in monitoring and reporting forest carbon stocks and changes. The findings from this research also demonstrate the utility of Earth Observation satellite data in monitoring forests across large areas (a difficult task when only reliant on field-based methods). Furthermore, regular and consistent observations acquired through LTS can provide us with a better understanding of the complexity and dynamic nature of forested systems and help us meet forest related sustainable management and development goals.

Acknowledgement

I would like to take the opportunity to specifically thank those who have contributed to this research and support me throughout my PhD. Without your help, it could not be completed. My first gratitude goes to my panel of supervisors Prof Simon Jones and Dr Mariela Soto-Berelov from RMIT, and Dr Andrew Haywood from the European Forest Institute. For all of you, I would like to thank for your patience and understanding my strengths and weakness. Your supports through the last four years are unwavering and invaluable. Also, I would like to thank you for adding me in the LandFor project team that allowed me to conduct me PhD research in a collaborative approach and to achieve high quality outputs. Simon, thank you for accepting me onto this PhD from a very early date (nearly five years ago) and for your on-going support and encouragement since then. Mariela, thank you for being not only my supervisor but also one of my best friends in Australia. Your advice has been always invaluable. Andrew, your industry perspective and high-level strategic advice have been of great benefits to my PhD research. I also thank to my PhD companion, Samuel Hislop, for his support and contribution throughout our shared PhD journeys.

I would like to extend my gratitude to my RMIT fellow PhD and postdocs: Sam (Hislop), Chithra, Ahmad, Nenad, Luke, Sam (Hillman), Bryant, Daisy, Chats, Shirley, Eloise, Jenna, Fiona and Jing. I appreciate your friendship and support for the last four years. I was not alone on my PhD journey as we were always together. I would like to acknowledge the Victorian Forest Monitoring Program team (Salahuddin Ahmad and Liam Costello) at the Department of Environment, Land, Water and Planning, who provided forest inventory data and support for this research.

I would like to acknowledge the Australian Award Scholarship (AAS) for providing the funding that made my PhD in Australia possible. My thank goes also to Jamie Low, AAS coordinator at RMIT, for her assistance in various matters. I also thank FrontierSI (formally CRCSI) for providing me a top-up scholarship to improve the quality of this research.

I greatly appreciate the constant support of my friends (in both Vietnam and Australia) during the last four years. Finally, to my family (bố Quang, mẹ Lan, bố Mẫn, mẹ Xuân, chị Hiền, Trang, và Hiếu Hạnh), without you I was not able to achieve this PhD. Mom and Dad, I know you will never read and understand what I am writing here (and I will also never tell you) but you are always my greatest motivation. Most importantly I would like to thank my wife, Hòa, and my two daughters, Chi and little Cherry; the reasons I get out of bed in the morning and come back home in the evening! Thank you for always with me, for your unwavering love and patience. Chi, you had obtained your first master's with your mom, and now your first PhD with me. We are so proud of you!

Thank you, everyone.

Contents

List of figures