Multi-source Remote Sensing for Forest Characterization and Monitoring

By

Qi Zhang

A thesis submitted to The University of New South Wales in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Geoscience and Earth Observing Systems Group (GEOS)
School of Civil and Environmental Engineering
Faculty of Engineering
The University of New South Wales
Sydney, NSW 2052, Australia

February 2022

Thesis/Dissertation Sheet

Surname/Family Name	:	**Zhang**
Given Name/s	:	**Qi**
Abbreviation for degree as give in the University calendar	:	**Doctor of Philosophy**
Faculty	:	**Faculty of Engineering**
School	:	**School of Civil and Environmental Engineering**
Thesis Title	:	**Multi-source Remote Sensing for Forest Characterization and Monitoring**

Abstract 350 words maximum:

As a dominant terrestrial ecosystem of the Earth, forest environments play profound roles in resource utilization, ecology, and biodiversity applications, which highlights the significance of forest characterization and monitoring. Related forest parameters can help track climate change and quantify the global carbon cycle and hence attract growing attention from various research communities. Compared with traditional in-situ methods with expensive and time-consuming field works involved, airborne and spaceborne remote sensors collect cost-efficient and consistent observations at global or regional scales and have been proven to be an effective way for forest measuring. With the looming paradigm shift toward data-intensive science and the development of remote sensors, remote sensing data with higher resolution and diversity have been the mainstream in data analysis and processing. However, significant heterogeneities in the multi-source remote sensing data largely restrain their synergy on forest applications urging the research community to come up with effective fusion strategies.

The work presented in this thesis contributes to the field by exploring the potential of the Synthetic Aperture Radar (SAR), SAR Polarimetry (PolSAR), SAR Interferometry (InSAR), Polarimetric SAR Interferometry (PolInSAR), Light Detection and Ranging (LiDAR), and multispectral remote sensing in forest characterization and monitoring from three main aspects including forest height estimation, active fire detection, and burned area mapping. Forest height estimation is first carried out on PolInSAR data based on model-based methods. LiDAR data is further employed to boost the accuracy of forest height estimation through the synergy with PolSAR and PolInSAR data. Active fire detection is performed on the multispectral acquisitions and the performance of burned area mapping is improved through the synergy of SAR, InSAR, and multispectral data. Since machine learning, especially deep learning methods have achieved great success in the field of data fusion, the fusion of the multi-source remote sensing data in this thesis is realized by Deep Neural Network (DNN) architectures like Convolutional Neural Network (CNN), Generative Adversarial Network (GAN), and self-attention mechanism, etc.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

ORIGINALITY STATEMENT

‘I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.’

Signed …….....

Date …….....

COPYRIGHT STATEMENT

Signed …….....

Date …….....

AUTHENTICITY STATEMENT

Signed …….....

Date …….....

INCLUSION OF PUBLICATIONS STATEMENT

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in their thesis in lieu of a Chapter if:

- The candidate contributed greater than 50% of the content in the publication and is the "primary author", ie. the candidate was responsible primarily for the planning, execution and preparation of the work for publication
- The candidate has approval to include the publication in their thesis in lieu of a Chapter from their supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis

Please indicate whether this thesis contains published material or not:

☐ This thesis contains no publications, either published or submitted for publication

☐ Some of the work described in this thesis has been published and it has been documented in the relevant Chapters with acknowledgement

☒ This thesis has publications (either published or submitted for publication) incorporated into it in lieu of a chapter and the details are presented below

CANDIDATE'S DECLARATION

I declare that:

- I have complied with the UNSW Thesis Examination Procedure
- where I have used a publication in lieu of a Chapter, the listed publication(s) below meet(s) the requirements to be included in the thesis.

Candidate's Name	**Signature**	**Date (dd/mm/yy)**
Qi Zhang		

POSTGRADUATE COORDINATOR'S DECLARATION

I declare that:

- the information below is accurate
- where listed publication(s) have been used in lieu of Chapter(s), their use complies with the UNSW Thesis Examination Procedure
- the minimum requirements for the format of the thesis have been met.

PGC's Name	**PGC's Signature**	**Date (dd/mm/yy)**
Ehab Hamed		

For each publication incorporated into the thesis in lieu of a Chapter, provide all of the requested details and signatures required

Details of publication #1:						
Full title: Deep-learning-based burned area mapping using the synergy of Sentinel-1&2 data *Authors:* Zhang, Q., Ge, L., Zhang, R., Metternicht, G. I., Du, Z., Kuang, J. and Xu, M. *Journal or book name:* Remote Sensing of Environment *Volume/page numbers:* 264/112575 *Date accepted/ published:* June 2021/October 2021						
Status	*Published*	☒	*Accepted and In press*		*In progress (submitted)*	
The Candidate's Contribution to the Work The first author of the article: Conceptualization, Methodology, Software, Validation, Writing original draft, Writing review and editing.						
Location of the work in the thesis and/or how the work is incorporated in the thesis: This work is located in Chapter 6. Related methods, experiments, results, and analysis in this work are presented in Chapter 6.						
PRIMARY SUPERVISOR'S DECLARATION I declare that: • the information above is accurate • this has been discussed with the PGC and it is agreed that this publication can be included in this thesis in lieu of a Chapter • All of the co-authors of the publication have reviewed the above information and have agreed to its veracity by signing a 'Co-Author Authorisation' form.						
Primary Supervisor's name **Linlin Ge**	**Primary Supervisor's signature**			**Date (dd/mm/yy)**		

Details of publication #2:						
Full title: Towards a Deep-Learning-Based Framework of Sentinel-2 Imagery for Automated Active Fire Detection *Authors:* Zhang, Q., Ge, L., Zhang, R., Metternicht, G. I., Liu, C., and Du, Z. *Journal or book name:* Remote Sensing *Volume/page numbers:* 13/4790 *Date accepted/ published:* November 2021/November 2021						
Status	*Published*	☒	*Accepted and In press*		*In progress (submitted)*	
The Candidate's Contribution to the Work						

The first author of the article: Conceptualization, Methodology, Software, Validation, Data curation, Writing-original draft preparation, Visualization.

Location of the work in the thesis and/or how the work is incorporated in the thesis:

This work is located in Chapter 5. Related methods, experiments, results, and analysis in this work are presented in Chapter 5.

PRIMARY SUPERVISOR'S DECLARATION

I declare that:

- the information above is accurate
- this has been discussed with the PGC and it is agreed that this publication can be included in this thesis in lieu of a Chapter
- All of the co-authors of the publication have reviewed the above information and have agreed to its veracity by signing a 'Co-Author Authorisation' form.

Primary Supervisor's name	**Primary Supervisor's signature**	**Date (dd/mm/yy)**
Linlin Ge		

Details of publication #3:

Full title: PolGAN: A deep-learning-based unsupervised forest height estimation based on the synergy of PolInSAR and LiDAR data

Authors: Zhang, Q., Ge, L., Hensley, S., Metternicht, G. I., Liu, C. & Zhang, R.

Journal or book name: ISPRS Journal of Photogrammetry and Remote Sensing

Volume/page numbers: 186/123-139

Date accepted/ published: February 2022/April 2022

Status	*Published*		*Accepted and In press*	☒	*In progress (submitted)*	

The Candidate's Contribution to the Work

The first author of the article: Conceptualization, Methodology, Software, Validation, Data curation, Writing-original draft preparation, Visualization.

Location of the work in the thesis and/or how the work is incorporated in the thesis:

This work is located in Chapter 4. Related methods, experiments, results, and analysis in this work are presented in Chapter 4.

PRIMARY SUPERVISOR'S DECLARATION

I declare that:

- the information above is accurate
- this has been discussed with the PGC and it is agreed that this publication can be included in this thesis in lieu of a Chapter

• All of the co-authors of the publication have reviewed the above information and have agreed to its veracity by signing a 'Co-Author Authorisation' form.		
Primary Supervisor's name **Linlin Ge**	**Primary Supervisor's signature**	**Date (dd/mm/yy)**

Abstract

As a dominant terrestrial ecosystem of the Earth, forest environments play profound roles in ecology, biodiversity, resource utilization, and management, which highlights the significance of forest characterization and monitoring. Some forest parameters can help track climate change and quantify the global carbon cycle and therefore attract growing attention from various research communities. Compared with traditional in-situ methods with expensive and time-consuming field works involved, airborne and spaceborne remote sensors collect cost-efficient and consistent observations at global or regional scales and have been proven to be an effective way for forest monitoring. With the looming paradigm shift toward data-intensive science and the development of remote sensors, remote sensing data with higher resolution and diversity have been the mainstream in data analysis and processing. However, significant heterogeneities in the multi-source remote sensing data largely restrain its forest applications urging the research community to come up with effective synergistic strategies.

The work presented in this thesis contributes to the field by exploring the potential of the Synthetic Aperture Radar (SAR), SAR Polarimetry (PolSAR), SAR Interferometry (InSAR), Polarimetric SAR Interferometry (PolInSAR), Light Detection and Ranging (LiDAR), and multispectral remote sensing in forest characterization and monitoring from three main aspects including forest height estimation, active fire detection, and burned area mapping.

First, the forest height inversion is demonstrated using airborne L-band dual-baseline repeat-pass PolInSAR data based on modified versions of the Random Motion over Ground (RMoG) model, where the scattering attenuation and wind-derived random motion are described in conditions of homogeneous and heterogeneous volume layer, respectively. A boreal and a tropical forest test site are involved in the experiment to explore the flexibility of different models over different forest types and based on that, a leveraging strategy is proposed to boost the accuracy of forest height estimation.

The accuracy of the model-based forest height inversion is limited by the discrepancy between the theoretical models and actual scenarios and exhibits a strong dependency on the system and scenario parameters. Hence, high vertical accuracy LiDAR samples are employed to assist the PolInSAR-based forest height estimation. This multi-source forest height estimation is reformulated as a pan-sharpening task aiming to generate forest heights with high spatial resolution and vertical accuracy based on the synergy of the sparse LiDAR-derived heights and the information embedded in the PolInSAR data. This process is realized by a specifically designed generative adversarial network (GAN) allowing high accuracy forest height estimation less limited by theoretical models and system parameters. Related experiments are carried out over a boreal and a tropical forest to validate the flexibility of the method.

An automated active fire detection framework is proposed for the medium resolution multispectral remote sensing data. The basic part of this framework is a deep-learning-based semantic segmentation model specifically designed for active fire detection. A dataset is constructed with open-access Sentinel-2 imagery for the training and testing

of the deep-learning model. The developed framework allows an automated Sentinel-2 data download, processing, and generation of the active fire detection results through time and location information provided by the user. Related performance is evaluated in terms of detection accuracy and processing efficiency.

The last part of this thesis explored whether the coarse burned area products can be further improved through the synergy of multispectral, SAR, and InSAR features with higher spatial resolutions. A Siamese Self-Attention (SSA) classification is proposed for the multi-sensor burned area mapping and a multi-source dataset is constructed at the object level for the training and testing. Results are analyzed by different test sites, feature sources, and classification methods to assess the improvements achieved by the proposed method.

All developed methods are validated with extensive processing of multi-source data acquired by Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR), Land, Vegetation, and Ice Sensor (LVIS), PolSARproSim+, Sentinel-1, and Sentinel-2. I hope these studies constitute a substantial contribution to the forest applications of multi-source remote sensing.

List of Abbreviations

ABI	Advanced Baseline Imager
ABoVE	Arctic Boreal Vulnerability Experiment
ACPE	Striped maple
ACRU	Red maple
ACSA	Sugar maple
ACT	Australia Capital Territory
AMSL	Above Mean Sea Level
ASI	Italian Space Agency
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVHRR	Advanced Very High Resolution Radiometer
BA	Burned Area
BEAL	Yellow birch
BEEX	Brazil nut
BELE	Sweet birch
BEPA	Paper birch
BEPO	Grey birch
BERMS	Ecosystem Research and Monitoring Sites
BN	Batch Normalization
BOA	Bottom of Atmosphere
BS	Backscatter Coefficient
CE	Commission Error
CNN	Convolutional Neural Network
COR	Interferometric Coherence
CSA	Canadian Space Agency
DC	Dice Coefficient
DCNN	Deep Convolutional Neural Network
DCPA	Dual-domain Channel-Position Attention
DL	Deep Learning
DLR	German Aerospace Centre

dNDVI	Difference of Normalized Difference Vegetation Index
DNN	Deep Neural Network
DOF	Degree of Freedom
EFFIS	European Forest Fire Information System
ESA	European Space Agency
ETM+	Enhanced Thematic Mapper
FAGR	American beech
FCN	Fully Convolutional Network
FIRMS	Fire Information for Resource Management System
FLS	Full Lambda-schedule
FN	False Negative
FP	False Positive
GA	Geoscience Australia
GAN	Generative Adversarial Network
GEDI	Global Ecosystem Dynamics Initiative
GIS	Geographic Information System
GRD	Ground Range Detected
HFT	Hypercomplex Fourier Transform
HTA	High-temperature Anomalies
InSAR	Synthetic Aperture Radar Interferometry
ICESat	Ice, Cloud, and land Elevation Satellite
IFOV	Instantaneous Field of View
IoU	Intersection over Union
ISA	Italian Space Agency
JAXA	Japan Aerospace Exploration Agency
JPL	Jet Propulsion Laboratory
LiDAR	Light Detection and Ranging
LR	Linear Regression
LSTM	Long Shot-Term Memory
LVA	Linear Volume Attenuation
LVIS	Land, Vegetation, and Ice Sensor
LVM	Linear Volume Motion

MCD12Q1	MODIS Land Cover Type Product
MGRS	Military Grid Reference System
ML	Machine Learning
MLP	Multilayer Perceptron
MLR	Multi-layer Perceptron
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	Multispectral Instrument
MSS	Multispectral Scanner
NASA	National Aeronautics and Space Administration
NHI	Normalized Hotspot Index
NIR	Near Infrared
NISAR	NASA-ISRO Synthetic Aperture Radar
NSW	New South Wales
OE	Ommission Error
OLI	Operational Land Imager
OSVI	American hophornbeam
PCA	Principal Component Analysis
PIRE	Red pine
PIRU	Red spruce
PIST	White pine
PNN	CNN-based Pan-sharpening
PNP	Pongara National Park
PolInSAR	Polarimetric Synthetic Aperture Radar Interferometry
PolSAR	Synthetic Aperture Radar Polarimetry
PRSE	Black cherry
PSGAN	Pan-sharpening GAN
QUAL	White Oak
QURU	Red Oak
QVA	Quadratic Volume Attenuation
QVM	Quadraric Volume Motion
RF	Random Forest
RHMA	Red Mangrove

RH100	Relative Height 100
RMoG	Random Motion over Ground
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RVoG	Random Volume over Ground
SAR	Synthetic Aperture Radar
SE	Squeeze-and-excitation
SEDI	Space Environment Data Acquisition monitor
SEVIRI	Spinning Enhanced Visible and Infra-Red Imager
SLC	Single-look Complex
SLSTR	Sea and Land Surface Temperature Radiometer
SNR	Signal-to-Noise Ratio
SOTA	State-of-the-art
SR	Surface Reflectance
SSA	Siamese Self-attention
SVM	Support Vector Machine
SWIR	Shortwave Infrared
TIR	Thermal Infrared
TIRS	Thermal Infrared Sensor
TM	Thermal Mapper
TN	True Negative
TOA	Top of Atmosphere
TomoSAR	Synthetic Aperture Radar Tomography
TP	True Positive
TS	Time-series
TSCA	Eastern hemlock
UAV	Unmanned Aerial Vehicle
UAVSAR	Uninhabited Aerial Vehicle Synthetic Aperture Radar
VIC	Victoria
VIIRS	Visible Infrared Imaging Radiometer Suite
VIRR	Visible and Infra-Red Radiometer
VNIR	Visible and Near Infrared

3D	Three Dimensional

Acknowledgments

This research was conducted in the Geoscience and Earth Observing Systems Group (GEOS), School of Civil and Environmental Engineering at the University of New South Wales (UNSW), Sydney, Australia from October 2017 to October 2021 under the supervision of Professor Linlin Ge and Professor Graciela Isabel Metternicht from UNSW, and Dr. Scott Hensley from NASA's Jet Propulsion Laboratory (JPL) at the California Institute of Technology (Caltech). First and foremost, I would like to thank my supervisor, Professor Linlin Ge, for supervising and supporting my Ph.D. research. Meanwhile, I would like to express gratitude to my co-supervisors, Professor Graciela Isabel Metternicht and Dr. Scott Hensley, for their gentle guidance and reassurance in supporting this work.

Thanks also go to all GEOS group members for numerous discussions, particularly regarding radar and multi-spectral remote sensing, image processing, deep learning, and the philosophy of life. Especially, I am so grateful to our current and former GEOS group members including Zheyuan Du, Chang Liu, Jianming Kuang, Tony Sleigh who is also the director of NSW Department of Land, and Ziheng Sheng for their support and friendship. I would also like to acknowledge the assistance from Dr. Ahmed Razi at JPL, Dr. Qiaolin Shi at Huawei, and Dr. Ruiheng Zhang at the Beijing Institute of Technology for their precious suggestions on PolInSAR data, digital signal, and deep-learning-based image processing.

NASA’s Jet Propulsion Laboratory is acknowledged for providing the UAVSAR data and PolSARproSim+ SAR simulation software, the Laser Vegetation and Ice Sensor team from NASA’s Goddard Space Flight Center is acknowledged for providing the LVIS data. The European Space Agency is acknowledged for offering multi-spectral and SAR data from Sentinel satellites. Planet Labs, Inc. is acknowledged for offering high-resolution PlanetScope imagery.

Last but most important, my deepest gratitude to my parents, Xingya Zhang and Qiang Liu, for their unconditional love and emotional support. My best friends, Qiaolin Shi, Huiyang Li, and Chang Liu, for their company and support.

List of Publications

Several journals and conference papers have been accepted for publishing, published, or submitted during the journey of this study. The publication list is shown as follows:

Journal Papers:

Zhang, Q., Ge, L., Zhang, R., Metternicht, G. I., Du, Z., Kuang, J. & Xu, M., 2021. Deep-learning-based burned area mapping using the synergy of Sentinel-1&2 data. Remote Sensing of Environment, 264, 112575.

Zhang, Q., Ge, L., Zhang, R., Metternicht, G. I., Liu, C., & Du, Z., 2021. Towards a Deep-Learning-Based Framework of Sentinel-2 Imagery for Automated Active Fire Detection. Remote Sensing, 13(23), 4790.

Zhang, Q., Ge, L., Hensley, S., Metternicht, G. I., Liu, C. & Zhang, R., 2022. PolGAN: A deep-learning-based unsupervised forest height estimation method based on the synergy of PolInSAR and LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing. (Accepted)

Zhang, R., Yang, S., **Zhang, Q.***, Xu L., He, Y., & Zhang, F., 2021.Graph-based Few-Shot Learning with Transformed Feature Propagation and Optimal Class Allocation. Neurocomputing, 470, 247-256.

Du, Z., Ge, L.*, Ng, A.H.M. & **Zhang, Q.**, 2020. Towards a revised framework of modified time series InSAR for mapping land deformation. Journal of Geodesy, 94(9), pp.1-13.

Du, Z., Ge, L.*, Ng, A.H.M., **Zhang, Q.** & Alamdari, M.M., 2018. Assessment of the accuracy among the common persistent scatterer and distributed scatterer based

on SqueeSAR method. IEEE Geoscience and Remote Sensing Letters, 15(12), pp.1877-1881.

Du, Z., Ge, L.*, Ng, A.H.M., Zhu, Q., **Zhang, Q.**, Kuang, J. & Dong, Y., 2019. Long-term Subsidence in Mexico City from 2004 to 2018 Revealed by Five Synthetic Aperture Radar Sensors. Land Degradation & Development, 30(15), pp.1785-1801.

Du, Z., Ge, L.*, Ng, A.H.M., Zhu, Q., Horgan, F.G. & **Zhang, Q.**, 2020. Risk assessment for tailings dams in Brumadinho of Brazil using InSAR time series approach. Science of The Total Environment, 717, p.137125.

Du, Z., Ge, L., Ng, A. H. M., Lian, X., Zhu, Q., Horgan, F. G., & **Zhang, Q.**, 2021. Analysis of the impact of the South-to-North water diversion project on water balance and land subsidence in Beijing, China between 2007 and 2020. Journal of Hydrology, 603, 126990.

Conference Papers:

Zhang, Q., Ge, L., & Du, Z. (2019, July). A Modified RMoG Model for Forest Height Inversion Using L-Band Repeat-Pass Pol-InSAR Data. In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium (pp. 3217-3220). IEEE.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Scientific Relevance and Motivation

Forest are distributed throughout the world as the dominant terrestrial ecosystem of the Earth. The State of the World's Forests 2020 reported that almost 31% of the global land area is covered by forests providing habitats for 80% of amphibian species, 75% of bird species, and 68% of mammal species (FAO, 2020). Besides that, by converting carbon dioxide into oxygen and biomass, forests are acting as a carbon sink to mitigate climate change (Kiat et al., 2020, Xu et al., 2020). In the past decades, deforestation and forest degradation continue to take place at alarming rates because of human activities such as agriculture, mining, infrastructure, and urban expansion (Van Khuc et al., 2018), which have caused significant decreases in biodiversity and negative carbon emission (Binzaid, 2020). On the one hand, human-wildlife conflicts will occur when the natural balance is disrupted by habitat losses. The interaction between humans and wildlife increases the exposure and risk of people to zoonotic diseases such as malaria, Chagas disease, African trypanosomiasis, leishmaniasis, Lyme disease, HIV, Ebola, and also the recent COVID-19 pandemic (Austin, 2020, Tollefson, 2020). On the other hand, deforestation and forest degradation weaken the ability of forests in capturing and storing carbon and thus increase the potential of climate change (Withey et al., 2019). The duration of fire season and the occurrence of catastrophic fires are significantly increased by the extreme weather, which turns into a huge threat to life and property (Jones et al., 2020). The large amounts of smoke and ash produced by the wildfire can impact human health by triggering lung disease

and the burned areas left behind are susceptible to extreme rainfalls and more likely to cause landslides and mudslides (Xu et al., 2020). In a word, a sustainable forest ecology is critical to human society and the natural environment, and therefore the global monitoring and management attract a lot of research.

As an effective approach in pressing environmental challenges and driving solutions, remote sensing techniques have been widely applied in forest ecology (Goswami et al., 2020, Lechner et al., 2020). The physical parameters, three-dimensional structures, and spatial distributions of forest ecosystems can be efficiently estimated through the reflected and released electromagnetic radiation collected by sensors onboard airborne and spaceborne platforms without direct contact with the Earth's surface (Luo et al., 2019, Yao et al., 2019). Compared with traditional in-situ methods with expensive and time-consuming field works involved, the spaceborne and airborne remote sensors can provide consistent and cost-efficient observations at the regional or/and global scales. With the rapid advancement in the associated sensors, remote sensing data with better spatial, temporal, spectral, or radiance resolution are continuously captured to support applications in weather, agriculture, forest ecosystems, water resources, and marine sciences (Cui et al., 2019, Shanmugapriya et al., 2019, Watt et al., 2019). The continuous upgrade of remote sensing systems inevitably benefits the development of related data processing algorithms. At the earlier stage, researchers are dedicated to extracting environmental parameters based on physical models, which largely rely on prior knowledge and exhibit great uncertainties in the various physical processes and imaging scenarios (Yuan et al., 2020). Hence, the looming paradigm shift toward data-intensive science like Machine Learning (ML) especially Deep Learning (DL) (Zhu et

al., 2017). In the past few years, with the increasing availability of Earth's big data, various ML or/and DL methods (Zhang et al., 2016a, Yuan et al., 2020) have achieved significant success in various remote sensing tasks. However, new challenges emerge when applying the ML or DL methods to remote sensing data because they are often multimodal, geolocated, quality controlled, and time-varying. These characteristics of remote sensing data urge scientists to further explore the potential and tackle the challenge of applying ML and DL methods for different scientific purposes.

1.2 Research Objective and Thesis Outline

1.2.1 Research Object

The general goal of this thesis is the optimum exploitation of SAR, InSAR, PolSAR, PolInSAR, LiDAR, and multi-spectral remote sensing for forest characterization and monitoring from three main aspects including the forest height estimation, active fire detection, and burned area mapping. In normal conditions, the focuses are mainly on forest height estimation using PolSAR, PolInSAR, or/and LiDAR products to provide a key indicator of forest productivity and biomass level. During a wildfire, research is dedicated to active fire detection based on multi-spectral acquisitions to extract timely and effective geolocation information of the burning fires. After the fire event, studies on the burned area mapping with the synergy of SAR, InSAR, and multi-spectral data can help to evaluate biomass losses, manage post-fire policies, and prevent secondary disasters. Since different sources of remote sensing data are involved in different forest applications, multiple data processing strategies need to be deployed to boost accuracy.

First of all, prior studies have been dedicated to estimating the canopy height through a series of model-based methods using PolInSAR data. Although physical models can effectively express the formation process from forest parameters to related PolInSAR observations, the forest parameter inversion performance exhibits a strong dependency on the prior knowledge of model parameters, imaging scenarios, and data acquisition configurations. This dependency brings uncertainties into the large-scale or long-term forest height inversion because the physical and imaging process varies with different regions and times. Therefore, the forest height inversion performance of models with different structural and physical characteristics needs to be evaluated in conditions of different imaging scenarios and data acquisition configurations. On basis of that, the ultimate goal is to achieve confidence forest height measurements for PolInSAR data by taking advantage of different physical models.

Though efforts have been placed on improving the performance of model-based forest height inversion using PolInSAR data. Limitations still exist because, in a SAR image, 3D radar reflectivity is projected into an azimuth-range domain where the forest height information is compressed in the complex pixel values. Though physical models offer prior knowledge for the forest height inversion, it is still hard for them to fully describe the physical process with high complexity, which will inevitably limit the accuracy of PolInSAR-based forest height inversion. Compared with PolInSAR, LiDAR measures the forest height more intuitively and thus provides height metrics with higher vertical accuracy. However, LiDAR samples are normally with lower spatial densities and thus unable to reveal the spatial details of height. Therefore, the goal is to develop a fusion algorithm to generate forest height estimates with higher spatial resolution and vertical

accuracy using PolSAR, PolInSAR, and LiDAR data. As a natural candidate to tackle the challenge in remote sensing data fusion, a deep neural network (DNN) is employed to complete this fusion task.

As to active fire detection, most prior studies take advantage of thermal infrared and mid-infrared bands with coarse spatial resolutions. However, the background radiance in coarse resolution pixels tends to overwhelm signals from subtle fires making it hard to reveal active fires with more spatial details. Hence, further efforts have been placed on active fire detection using medium-resolution multispectral data, where active fires exhibit sharp contrast with the background in the shortwave infrared (SWIR) and red false-color composite. By formulating active fire detection as a semantic segmentation task, different deep-learning-based models can be applied. The goal is to compare the detection performance of different segmentation models and adjust the architecture of DNN to make it more suitable for fire detection. Besides that, an automated active fire detection framework is constructed on the developed model for its further application on global monitoring.

The capability of medium-resolution multispectral data on burned area mapping has been evidenced in prior studies. However, the exclusive use of multispectral data for burned area mapping will be limited by solar illumination, weather condition, and spectral confusion. In this regard, the potentials of SAR and InSAR techniques on burned area mapping have also been explored because the active microwave has better penetration through the cloud coverage and less dependency on the weather condition. However, marginal improvements have been achieved by the synergy of the radar

backscatter coefficient (BS) and surface reflectance (SR) over cloud-free areas. The goal is to explore whether improvements in burned area mapping can be achieved through the synergy of SR, BS, and interferometric coherence (COR). As an effective tool to address challenges in multi-source feature extraction and fusion, a DNN is used to complete this task.

1.2.2 Thesis Outline

Chapter 3 describes a dual-baseline model-based method to extract forest height from airborne L-band repeat-pass PolInSAR data. Volumetric and temporal decorrelations with homogeneous and heterogenous descriptions are employed in the physical model to explore the potential of different structural and physical characteristics in the forest height inversion. PolInSAR data collected in repeat-pass dual-baseline configurations are used to address the underdetermined problem between coherence observations and forest parameters and a leveraging strategy is used to effectively utilize the advantages of different descriptions in the physical model.

Chapter 4 depicts a deep-learning-based unsupervised forest height estimation method based on the synergy of L-band repeat-pass PolInSAR and sparse large-footprint full-waveform LiDAR data. By reformulating the forest height inversion process as a pan-pansharpening task between the sparse LiDAR height and the high-resolution PolSAR and PolInSAR features, forest height estimates with high spatial resolution and vertical accuracy can be generated by a tailored PolGAN model with one generator and two discriminators. Shape, texture, and other spatial information embedded in high-spatial-resolution features are effectively utilized by the Convolutional Neural Network (CNN)

in the proposed architecture and the dependency of forest height inversion on physical models is significantly reduced. Contents in this chapter have been summarized and submitted to the *ISPRS Journal of Photogrammetry and Remote Sensing.*

Chapter 5 develops an automated active fire detection framework for the Sentinel-2 imagery. The basic part of the framework is an active fire detection module built on a DCPA (Dual-domain Channel-Position Attention)+HRNetV2 model and a dataset with semi-manually annotated active fire samples are constructed for the training and testing. Concepts in this framework can be further applied to remote sensing data in SWIR and Red ranges and serve as a cost-efficient resource in support of governments and fire service agencies that need timely firefighting plans. Contents in this chapter have been published to the *Remote Sensing.*

Chapter 6 investigates if the existing coarse burned area mapping products derived from the exclusive use of SR acquisitions can be further improved by the synergy of SR, BS, and/or COR data with higher spatial resolutions. An SSA classification is proposed for the multi-sensor burned area mapping and a multi-source dataset is constructed at the object level for training and testing. Contents in this chapter have been published in the *Remote Sensing of Environment (Zhang et al., 2021a).*

Chapter 2 Background

This chapter presents an overview of the multi-source remote sensing techniques used in the study (SAR and optical) along with a brief introduction of related deep-learning-based algorithms.

2.1 Radar Remote Sensing

2.1.1 Synthetic Aperture Radar

As an active imaging sensor working in the microwave portion of the electromagnetic spectrum, SAR maps scattering properties of the Earth's surface almost independently on natural illuminations and weather conditions (Curlander and McDonough, 1991). SAR sensors transmit frequency-modulated pulsed waveforms (the chirp signal with a constant amplitude and an instantaneous frequency linearly varying over time) and receive the scattered echoes from the imaging target. This transmission and reception process is repeated every pulse repetition interval, which is equivalent to the reciprocal of the pulse repetition frequency (Richards, 2009). The typical SAR imaging geometry is presented in Figure 2.1, where the platform moves along the azimuth direction and views the ground target in a slant range direction perpendicular to the flight path. The swath width gives the ground-range extent of the radar scene.

The received raw SAR data need to be further processed to present useful information on the scenario. This is accomplished by compressing the two-dimensional complex echo signals with two separate matched filtering operations in the range and azimuth directions. Other processings like thermal noise removal, speckle reduction, geometric

and radiometric calibrations are also needed for the further applications of SAR data, which will not be presented in detail in this thesis but readers can refer to prior studies in (Curlander and McDonough, 1991, Fitch, 2012, Moreira et al., 2013, Ferro-Famil and Pottier, 2016) for more information.

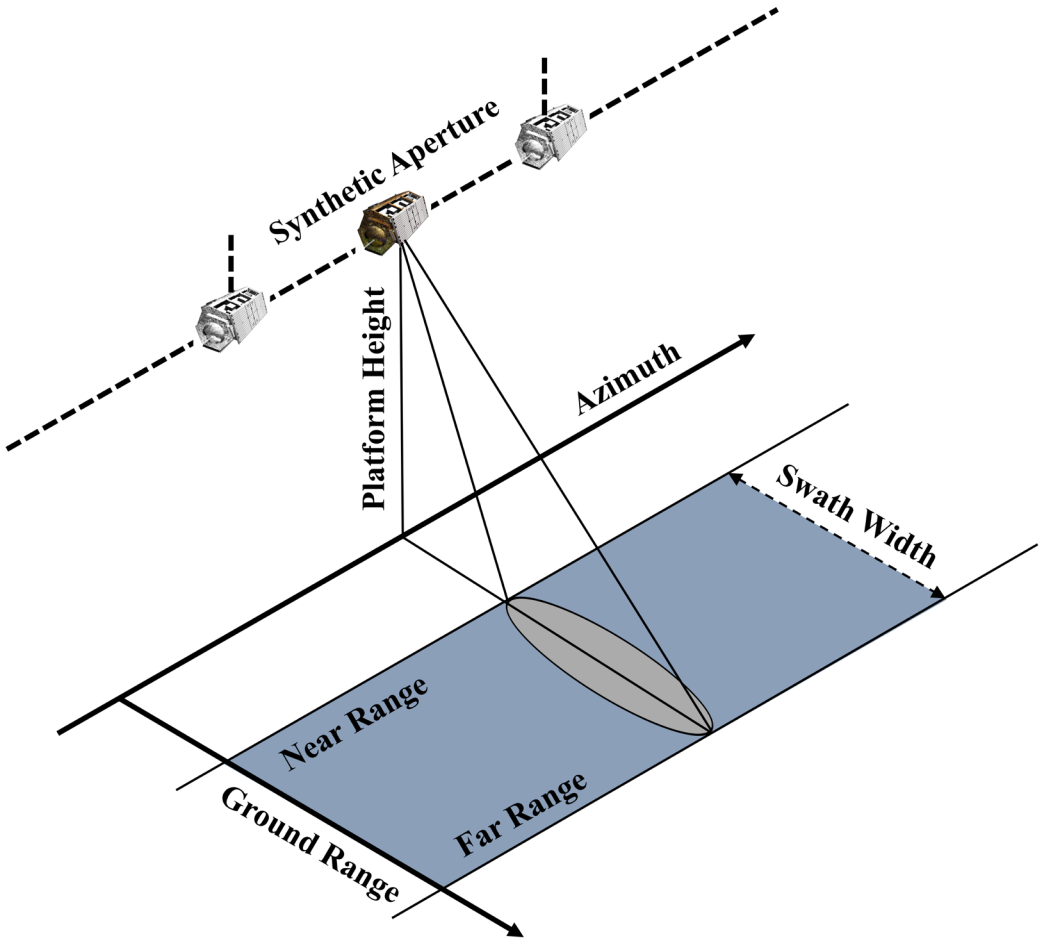


Figure 2.1 SAR imaging geometry.

The history of SAR started in 1951 when the Goodyear Aircraft company first brought up the concept. In 1978, the launch of the world's first spaceborne SAR called SeaSAT was conducted by the Jet Propulsion Laboratory (JPL) marking a successful entry of SAR into space. Several notable missions have been conducted since then including

ERS-1/2 satellites launched by the European Space Agency (ESA) in 1991 and 1995, JERS-1 launched by Japan Aerospace Exploration Agency (JAXA) in 1992, Radarsat-1/2 satellites launched by the Canadian Space Agency (CSA) in 1995 and 2007, Envisat satellite launched by ESA in 2002, ALOS satellite launched by JAXA in 2006, COSMO-SkyMed constellation with four satellites launched by the Italian Space Agency (ISA) from 2007 to 2010, TerraSAR-X/TanDEM-X satellites launched by the German Aerospace Centre (DLR) in 2007 and 2010, ALOS-2 satellite launched by JAXA in 2014, Sentinel-1A/B satellites launched by ESA in 2014 and 2016, Radarsat constellation with three satellites launched by CSA in 2019 (Lavalle, 2009). Recently, small satellite SAR technologies have been developed at a dramatic pace since the launch of the world's first SAR satellite under 100 kg (ICEYE-X1) in 2018 attracting several commercial companies such as the ICEYE, Capella Space, and UrtheCast to deploy and design small satellite constellation for high-quality, informative, and near-real-time earth observation in the future. Besides that, large SAR satellite missions are also planned by space agencies and companies, like TanDEM-L, ROSE-L, ALOS-4, etc. The more and more advanced techniques extended SAR acquisitions into multi-polarization, multi-frequency, and multi-baseline to discern the geophysical properties of the target, which boosted the development of successive techniques like PolSAR, InSAR, PolInSAR, and SAR Tomography (TomoSAR).

2.1.2 Synthetic Aperture Radar Polarimetry

The introduction of electromagnetic wave polarization provides SAR systems with the capability to explore a complete description of propagation and scattering phenomena. Instead of the scattering scalar in single-polarimetric SAR, PolSAR offers a scattering

matrix observation $\boldsymbol{S}$ allowing to construct a powerful feature space that is sensitive to shape, orientation, and dielectric properties of the scatterers and to develop physical models for the identification and separation of different scattering mechanisms in the same resolution cell (Moreira et al., 2013). $\boldsymbol{S}$ is acquired by transmitting microwave energy interleaved in the horizontal (H) and vertical (V) polarizations and recording scattered waves received in the two polarizations simultaneously. Other orthogonal polarization bases can also be used to measure the scattering matrix indicating that information embedded in $\boldsymbol{S}$ is independent of the orthogonal polarization bases. Accordingly, the scattering matrix of arbitrary polarization bases can be reconstructed through the linear combination of the elements in $\boldsymbol{S}$ (van Zyl, 2011).

$$\boldsymbol{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \tag{2.1}$$

In monostatic configurations with co-located receiver and transmitter, $\boldsymbol{S}$ is symmetric ($S_{HV} = S_{VH}$) for all reciprocal scattering media. Although the deterministic (point-like) scatterers can be fully described by $\boldsymbol{S}$, it does not apply to distributed scatterers made up of randomly distributed scattering elements. Therefore, second-order statistical formalisms like coherency $\boldsymbol{T}$ and covariance $\boldsymbol{C}$ matrices are used to fully characterize the polarimetric scattering behavior of distributed scatterers.

$$\boldsymbol{T} = \langle \vec{k}_P \vec{k}_P^{\dagger} \rangle = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{12}^* & T_{22} & T_{23} \\ T_{13}^* & T_{23}^* & T_{33} \end{bmatrix}, \text{with } \vec{k}_P = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ 2S_{HV} \end{bmatrix} \tag{2.2}$$

$$\boldsymbol{C} = \langle \vec{k}_L \vec{k}_L^{\dagger} \rangle = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{12}^* & C_{22} & C_{23} \\ C_{13}^* & C_{23}^* & C_{33} \end{bmatrix}, \text{with } \vec{k}_L = \begin{bmatrix} S_{HH} \\ \sqrt{2} S_{HV} \\ S_{VV} \end{bmatrix} \tag{2.3}$$

$\vec{k}_P$ and $\vec{k}_L$ are 3×1 scattering vectors in the Pauli and Lexicographic reformulations.

These polarimetric statistical characteristics have been widely used for terrain and land classification and segmentation based on scattering features extracted by polarimetric decomposition algorithms. For example, the scattering matrix is divided into dihedral, sphere, and helix scattering components in the Krogager algorithm (Krogager, 1990), the coherency matrix is decomposed using a four-component strategy (surface, double bounce, volume, and helix) in the Yamaguchi algorithm (Yamaguchi et al., 2011), and the covariance matrix is decomposed as a combination of the first-order Bragg surface, the canopy with randomly oriented dipoles, and the double-bounce scatter in Freeman decomposition (Freeman and Durden, 1998). Other methods including Cloude, Vanzyl, Huynen, H/A/Alpha, etc., have also been proposed for the polarimetric decomposition with their unique pros and cons (Lee and Pottier, 2017).

2.1.3 Synthetic Aperture Radar Interferometry

SAR interferometry is an imaging technique developed based on SAR for measuring the topography of a surface and/or its changes over time (Rosen et al., 2000). The basic principle of InSAR is to measure the phase difference between two or more complex radar images for a given scenario obtained by antennas at slightly different positions or/and at different times. Since this range information embedded in the phase of each SAR pixel is accurate to a small fraction of the radar wavelength, InSAR is possible to measure the surface topography and to detect the tiny path changes with centimetric or millimetric accuracy independent of the distance between the sensor and the scene. Principles of across-track InSAR for topographic mapping are similar to the traditional stereoscopic approach, where echos received by two mutually displaced antennas are coherently combined to extract the interferometric phase. The interferometric phase is

directly related to the difference between two geometric path lengths from the antenna to the target which can be further converted into the altitude of the target. In essence, the interferometric phase provides measurement in a third dimension in addition to the along- and cross-track location of the target allowing for a three-dimensional (3D) reconstruction of the topography (Bamler and Hartl, 1998). Changes of the target can be measured by the along-track or differential InSAR (Goldstein and Zebker, 1987, Ferretti et al., 2000, Reigber and Scheiber, 2003). Along-track InSAR observations are acquired in the same flight trajectory and imaging geometry but at different times. In this condition, changes of the SAR system clock, variable propagation delay, or/and surface motion in the direction of the radar line of sight can be measured by the phase differences of SAR pixels. SAR images used for differential InSAR are acquired with temporal and across-track separations at the same time, and thus contain information on both the topography and temporal changes (Moreira et al., 2013). The topographic component can be subtracted from the interferometric phase using an external DEM leading to a differential InSAR measurement sensitive to subtle changes of the range distance.

2.1.4 Polarimetric Synthetic Aperture Radar Interferometry

PolInSAR is developed by a coherent combination of InSAR and PolSAR supporting radar acquisitions sensitive to both physical and vertical properties of the scattering medium (Papathanassiou et al., 1998, Papathanassiou and Cloude, 2001). When a fully polarimetric coherent radar system observes the natural medium from two slightly different look angles at separated times as shown in Figure 2.2 PolInSAR imaging

geometry., two complex scattering matrices, $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$, related to the backscattered energy from the imaged scene will be obtained.

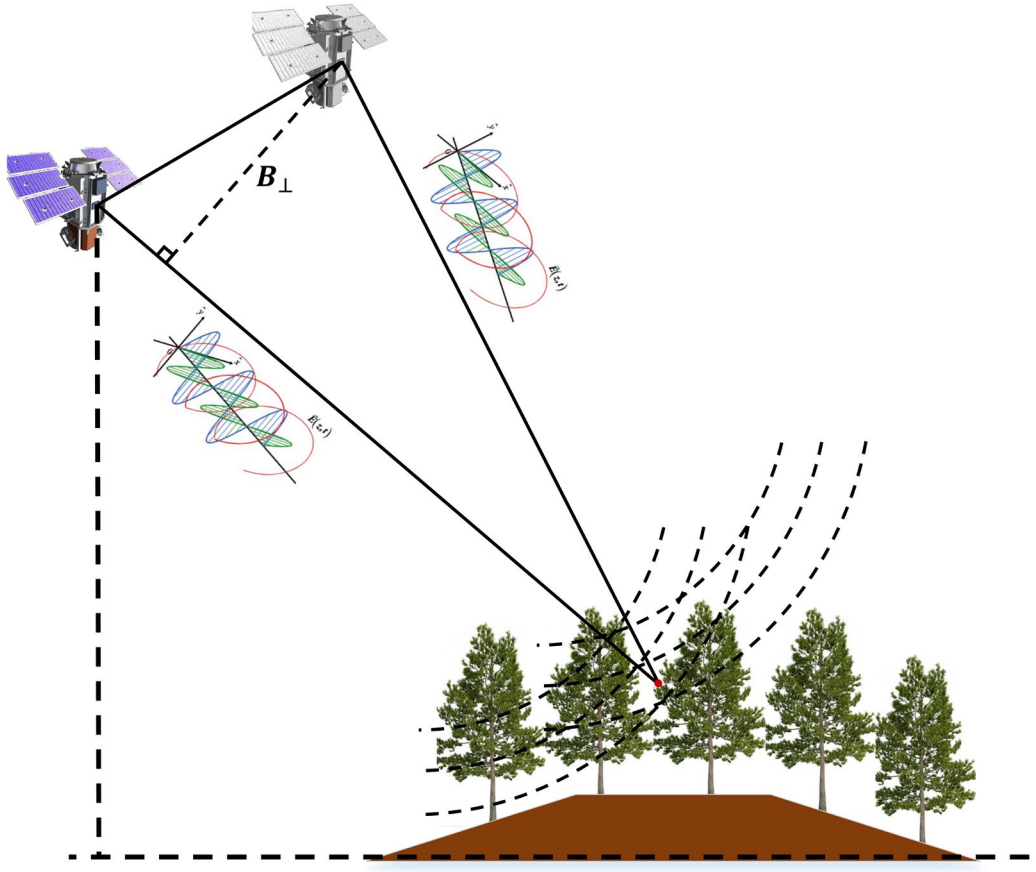


Figure 2.2 PolInSAR imaging geometry.

The scattering matrices are then vectorized into two Pauli vectors $\vec{k}_{P_i}$ to further obtain the second-order statistical matrices of PolInSAR.

$$\vec{k}_{P_i} = \frac{1}{\sqrt{2}}\begin{bmatrix} S_{i_{HH}} + S_{i_{VV}} \\ S_{i_{HH}} - S_{i_{VV}} \\ 2S_{i_{HV}} \end{bmatrix}, \text{with } i = 1,2 \tag{2.4}$$

$$\boldsymbol{T_{11}} = \langle \vec{k}_{P_1} \vec{k}_{P_1}^{\dagger} \rangle, \boldsymbol{T_{22}} = \langle \vec{k}_{P_2} \vec{k}_{P_2}^{\dagger} \rangle, \boldsymbol{\Omega_{12}} = \langle \vec{k}_{P_1} \vec{k}_{P_2}^{\dagger} \rangle \tag{2.5}$$

where $\boldsymbol{T_{11}}$ and $\boldsymbol{T_{22}}$ are 3×3 hermitian coherency matrices describing the polarimetric properties of each image and $\boldsymbol{\Omega_{12}}$ is a 3×3 non-hermitian complex matrix containing both polarimetric and interferometric information.

Finally, a general expression for the estimation of complex interferometric coherence can be derived from the second-order statistical matrices above.

$$\gamma = \frac{\langle \vec{\omega}_1^{\dagger} \boldsymbol{\Omega_1} \vec{\omega}_2 \rangle}{\langle \vec{\omega}_1^{\dagger} \boldsymbol{T_{11}} \vec{\omega}_1 \rangle \langle \vec{\omega}_2^{\dagger} \boldsymbol{T_{22}} \vec{\omega}_2 \rangle} = |\gamma| \exp(j\phi) \tag{2.6}$$

where $\vec{\omega}_1$ and $\vec{\omega}_2$ are the complex unitary vectors for an arbitrary choice of scattering mechanism, ϕ is the expected interferometric phase, $|\gamma|$ is the coherence magnitude. The estimated PolInSAR coherence mainly depends on decorrelations from thermal noise γ_{SNR}, baseline γ_b, volume γ_{vol}, and temporal changes γ_t.

$$\gamma = \gamma_{SNR} \gamma_b \gamma_{vol} \gamma_t \tag{2.7}$$

A thermal noise decorrelation is a real number only affecting the coherence magnitude and can be calculated by the signal-to-noise ratio (SNR) of the radar system (Fore et al., 2015).

$$\gamma_{SNR} = \frac{1}{1 - SNR^{-1}}, \text{with } SNR = \frac{\sigma_0}{NESZ} \tag{2.8}$$

where, σ_0 is the normalized measure of the radar return from a distributed target and $NESZ$ represents the sensitivity of the radar system to areas with low radar backscatter.

Baseline decorrelation contains a range and an azimuth spectral decorrelation derived from the slightly different incidence angles and Doppler centroids. Both contributions can be compensated by the so-called common spectral band filtering (Moreira et al.,

2013). The volumetric and temporal decorrelations are directly related to the vertical structure, temporal stability, and physical property of the scattering medium and have been widely applied to extract the vertical scattering reflectivity and temporal change profiles, especially over the forest area. Further details on the PolInSAR-based forest characterization and its state-of-art are provided in Chapter 3.

2.2 Optical Remote Sensing

2.2.1 LiDAR

Besides PolInSAR, LiDAR is also an effective active remote sensing technique well suited for forest characterization as it provides high-resolution information on vertical and horizontal vegetation structures and ground topography. LiDAR systems work at the near-infrared region of the electromagnetic spectrum measuring distances between the sensor and targets based on half the elapsed time between the emission of a pulse and the detection of the reflected returns (Wulder et al., 2012). LiDAR systems can be distinguished by the platform type (terrestrial, airborne, or spaceborne), footprint size (small or large), signal recording mode (discrete return or full-waveform), and sample scanning pattern (profiling or scanning). Large footprint profilers illuminate the target with circles in diameters from 10 to 25 m, whereas the diameters of the small footprint systems are much smaller (<1 m). Discrete systems record a single or multiple returned peaks, but full-waveform systems digitize the entire reflected energy. Profiling LiDAR instruments are usually used to measure the static or dynamic attributes of the forest at regional or/and global scales through sampling the vertical information over a swath width equal to the antenna footprint diameter, whereas the scanning LiDAR distributes

multiple pulses across a swath width providing vertical structures at stand level (Petrie and Toth, 2018).

The most commonly used LiDAR systems in forestry applications are small-footprint discrete return and large-footprint full-waveform LiDAR (Silva et al., 2018). The high sampling density of the small-footprint system provides accurate vertical and spatial information on individual trees, which are further used to estimate the stem density and timber volume, model wildfire behavior, and identify tree species (Nayegandhi et al., 2006). However, small-footprint systems require extensive flights to cover large areas and the small-diameter beams are more likely to miss the tops of trees. Therefore, large-footprint LiDAR increases the footprint size to the approximate crown diameter of a canopy-forming tree so that the laser energy consistently reaches the ground even in dense forests. The higher flying altitudes of large-footprint systems also enable a wider imaging swath and hence reduce the cost of mapping large areas.

Early LiDAR measurements were acquired by ground-based platforms in the 1960s following by the first flight on small aircraft in 1969 and on large aircrafts capable of the long-range measurements in the late 1970s. However, airborne LiDAR is usually limited by spatial coverage and high cost and hence is hard to apply to the regional or global scales. The first spaceborne LiDAR experiment was conducted using a shuttle in 1994. After that, long-duration spaceborne low-earth-orbit missions including the Ice, Cloud, and land Elevation Satellite (ICESat), ICESat-2, and the Global Ecosystem Dynamics Investigation (GEDI) were successively deployed in 2003 and 2018 (Schutz et al., 2005, Abdalati et al., 2010, Hancock et al., 2019), respectively. The development

of spaceborne LiDAR systems aims to generate a grid of the world's surface profiles by reducing the spatial sampling density.

2.2.2 Multispectral

Multispectral techniques have been widely applied in forest characterization since the launch of the first Landsat Multispectral Scanner (MSS) System in 1972. Unlike active SAR and LiDAR, multispectral remote sensors normally operate in the passive mode imaging the Earth's surface by detecting radiations reflected or/and emitted from the ground target. Passive multispectral remote sensing relies on two sources of radiation, visible, near-infrared, and shortwave infrared (VNIR-SWIR) radiation originated from the sun and thermal infrared (TIR) radiation emitted by materials on the Earth. This thesis mainly focuses on the applications of VNIR-SWIR radiation and hence a brief introduction will be carried out in this aspect. The VNIR-SWIR radiations at different locations of the electromagnetic spectrum are measured by the average reflected energy in a small spectral interval captured by multispectral sensors. Distinct spectral bands related to specific characteristics of the target are recorded simultaneously to form the multispectral observation. For instance, reflection characteristics in the range of 2 μm to 2.5 μm may give information about the mineral composition of the soil and the combined reflection characteristics of the red and near-infrared bands may indicate the biomass and health of vegetation. Inevitably, VNIR-SWIR bands are with a weaker penetration through the thick cloud and a stronger dependency on the solar illustration and weather conditions compared with the TIR bands (Schowengerdt, 2006, Xie et al., 2008).

Landsat has the longest history and widest use in the multispectral Earth's observation from space. Besides the first Landsat 1 launched in 1972, a series of Landsat missions from the Landsat 2 in 1975 to the Landsat 8 in 2013 has been successfully deployed and continuously delivered high-quality global observations for almost 50 years. The upcoming Landsat 9 satellite is being developed to replicate its predecessor Landsat 8 carrying more sophisticated Operational Land Imager 2 (OLI–2) and Thermal Infrared Sensor 2 (TIRS-2) (Masek et al., 2020). Landsat data collected by the MSS, Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), OLI, and TIRS cover a wide range on the electromagnetic spectrum in a medium spatial resolution from 15m to 100m and thus able to provide various spectral characteristics of the Earth's surface. Besides Landsat, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) on Terra and Multispectral Instrument (MSI) on Sentinel-2A/B also capture multispectral observations in a medium spatial resolution. Terra has a repeat cycle of 16 days (same as Landsat satellites) providing ASTER imagery of 14 bands with spatial resolutions of 15m in VNIR, 30m in SWIR, and 90m in TIR. The Sentinel-2 constellation operates in a shared sun-synchronous orbit with a repeat cycle of 10 days for one satellite and 5 days for two offering MSI imagery in 13 bands with spatial resolutions ranging from 10m to 60m. Medium spatial resolution acquisitions have been widely used for forest characterization at local and/or regional scales. Multispectral data acquired by coarse resolution sensors such as the Advanced Very High-Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) have been widely applied to study regional or/and global vegetation distributions and conditions in ecosystems. MODIS sensors onboard Terra and Aqua together can map the entire Earth every 1 to 2 days with observations in 36

spectral bands. AVHRR has an image archive of 43 years and therefore is very useful to study long-term changes in the vegetation (Xie et al., 2008). Observations with higher spatial resolutions of less than 5m are usually acquired by commercial satellites such as IKONOS, QuickBird, Worldview, Pleiades, Planetscope, and KOMPSAT, etc. Compared with medium and coarse resolution imagery, much more details of the Earth are revealed in these high-resolution observations and are widely used to characterize individual trees or vegetations in a small area.

2.3 Deep-learning-based Image Processing

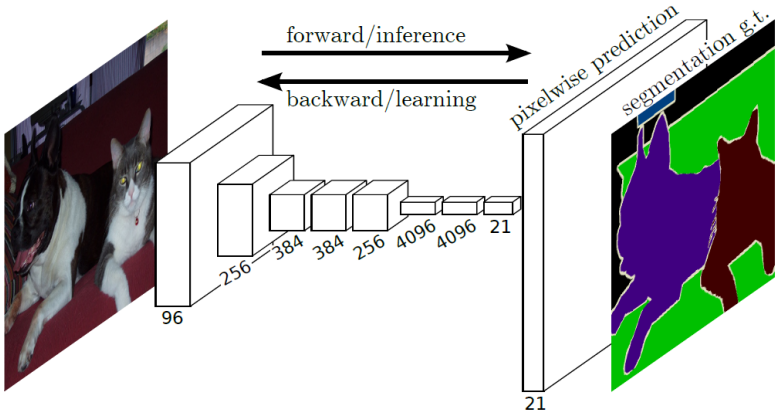
2.3.1 Segmentation

By breaking down the digital image into segments with simplified representations, the segmentation process aims to provide a reformulated image easier for further analysis and interpretation. There are two groups of image segmentation, semantic and instance. Semantic segmentation associates every pixel of an image with a class label and treats multiple objects of the same class as a single entity. Instance segmentation also assigns a class label to each pixel of an image except that it treats multiple objects of the same class as individual entities. As a fundamental problem in computer vision, numerous segmentation algorithms have been proposed since the earliest days of the field, which includes the histogram-based method (Ohlander et al., 1978), thresholding (Otsu, 1979), watershed (Najman and Schmitt, 1994), compression (Mobahi et al., 2011), k-means clustering (Dhanachandra et al., 2015), region growing (Nock and Nielsen, 2004), active contour (Kass et al., 1988), graph cuts (Boykov et al., 2001), Markov random fields (Plath et al., 2009), and the sparsity-based methods (Starck et al., 2005). Until recently, the success of deep-learning-based models in computer vision has

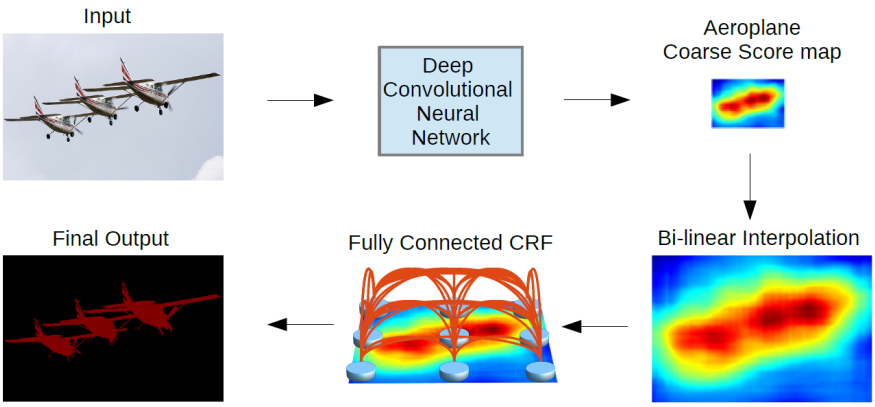
brought new momentum to this field. Since all forest applications in this thesis treat the same-class objects as a single entity, a brief introduction of the deep-learning-based semantic segmentation methods is included in this section.

As the milestone in deep-learning-based semantic segmentation, Fully Convolutional Network (FCN) (Long et al., 2015) includes only convolutional layers as presented in Figure **2.3**(a) allowing the network to take images of arbitrary size as the input. Feature maps from final layers are upsampled and fused with those of the earlier layers through skip connections for the multiscale inference. To efficiently account for the potentially useful scene-level semantic contexts, probabilistic graphical models like Conditional Random Field and Markov Random Field are incorporated into the deep convolutional neural network (DCNN) (Chen et al., 2014, Liu et al., 2015) as indicated in Figure **2.3**(b). Besides that, several popular DCNN-based segmentation models are designed based on the encoder-decoder architecture, which learns to map data points from an input domain to an output domain via a two-stage network as shown in Figure **2.3**(c). The encoder compresses high-dimensional inputs into low-dimensional latent-space representations to capture useful semantic contents, while the decoder predicts pixel-wise class probabilities based on these low-dimensional features. Examples of the encoder-decoder-based segmentation include DeConvNet (Noh et al., 2015), U-Net (Ronneberger et al., 2015), and SegNet (Badrinarayanan et al., 2017). One limitation in the encoder-decoder-based methods is the loss of fine-grained information at the encoding stage. HRNet (Wang et al., 2020) is proposed to address this problem, where high-resolution representations are maintained by repeated information exchanging among high and low-resolution convolution streams. Multiscale analysis has also been

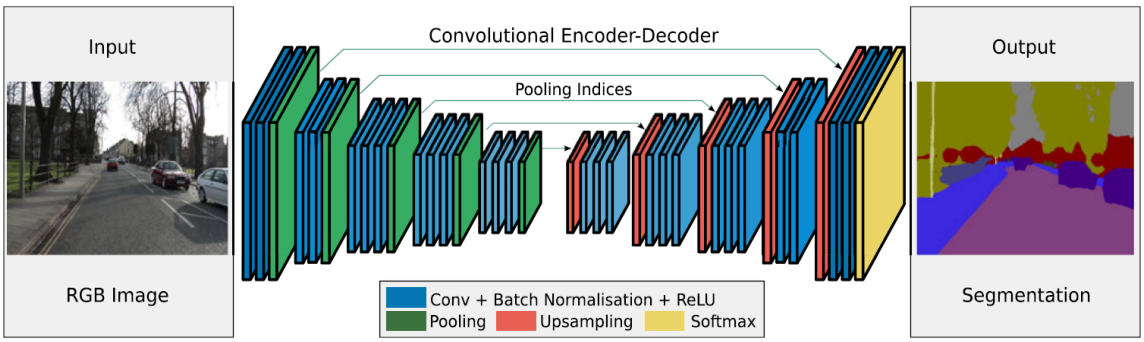
used in the DCNN-based semantic segmentation. Typical examples are Laplacian Pyramid Reconstruction and Refinement (Ghiasi and Fowlkes, 2016), Pyramid Scene Parsing Network (Zhao et al., 2017), and Feature Pyramid Network (Lin et al., 2017). Multiscale features are better integrated for semantic segmentation with the feature pyramid originated from the inherent multiscale and pyramidal hierarchy of DCNNs. Figure **2.3**(d) presents the architecture of PSPN. By introducing a dilation rate into the convolution layer, the receptive field of a kernel can be enlarged without introducing additional parameters and computational losses. Hence, dilation convolution models are widely applied to real-time segmentation including DeepLab family (Chen et al., 2017) and densely connected Atrous Spatial Pyramid Pooling (Yang et al., 2018). The architecture of the mostly used DeepLabV3+ model is shown in Figure **2.3**(e). The attention mechanisms allow models to evaluate the importance of features from different positions and scales and to capture rich contextual dependencies by assigning different weights. It has also been deployed in DCNN-based models such as Object Context Network (Yuan et al., 2018b), Pyramid Attention Network (Li et al., 2018), and Dual Attention Network (Fu et al., 2019) for the semantic segmentation. Besides CNN, Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Transformer architectures commonly used in sequential data processing have also been applied in semantic segmentation (Byeon et al., 2015, Zheng et al., 2021).


forward/inference
backward/learning
pixelwise prediction
segmentation g.t.
96
256
384
384
256
4096
4096
21
21

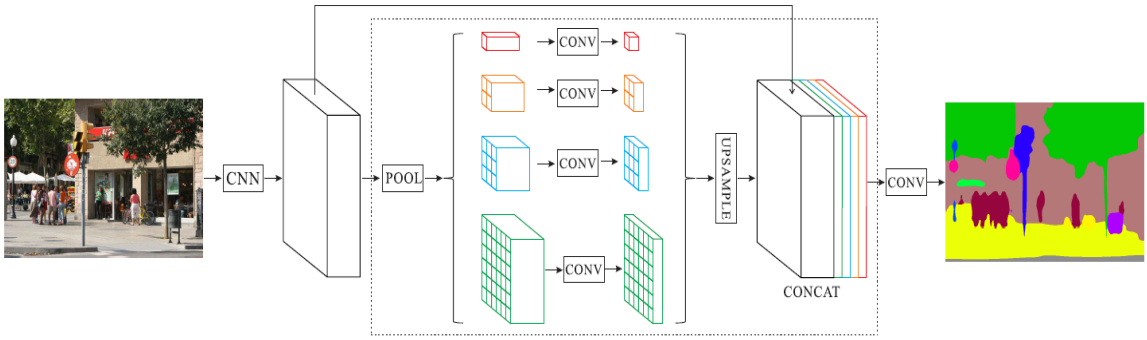
(a)


Input
Deep Convolutional Neural Network
Aeroplane Coarse Score map
Bi-linear Interpolation
Fully Connected CRF
Final Output

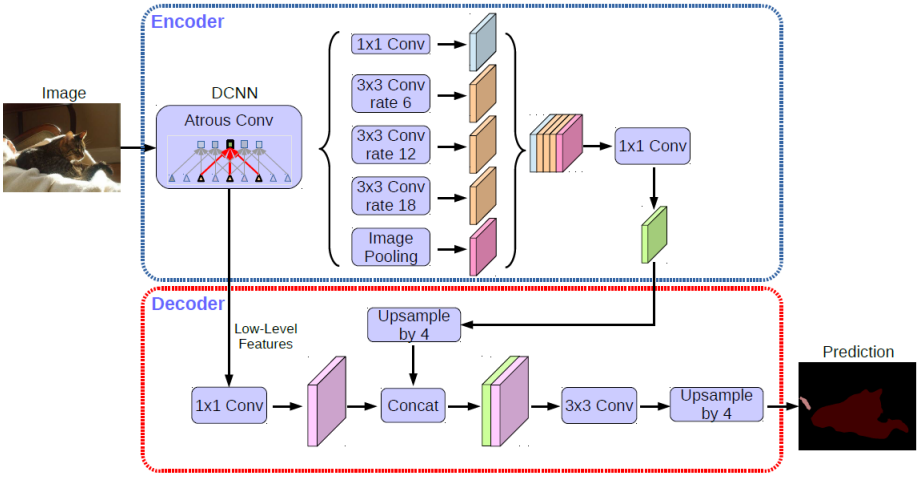
(b)


Input
RGB Image
Convolutional Encoder-Decoder
Pooling Indices
Conv + Batch Normalisation + ReLU
Pooling
Upsampling
Softmax
Output
Segmentation

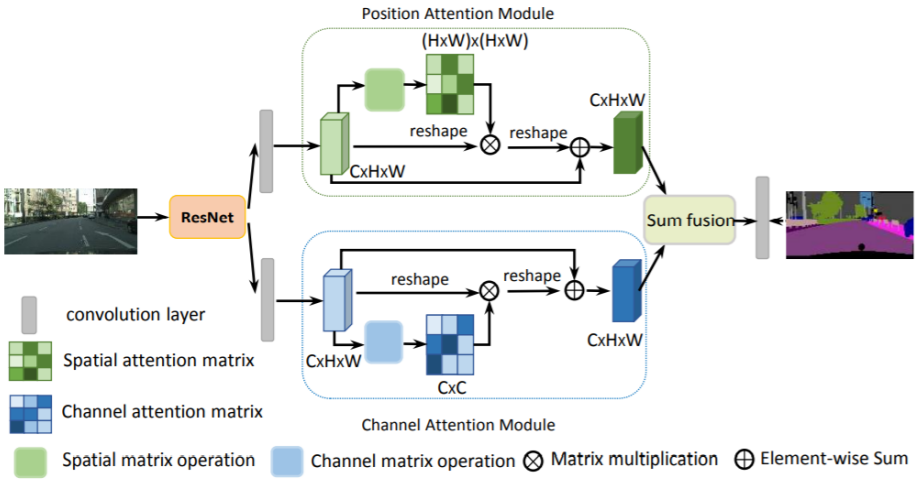
(c)


CNN
POOL
CONV
CONV
CONV
CONV
UPSAMPLE
CONCAT
CONV

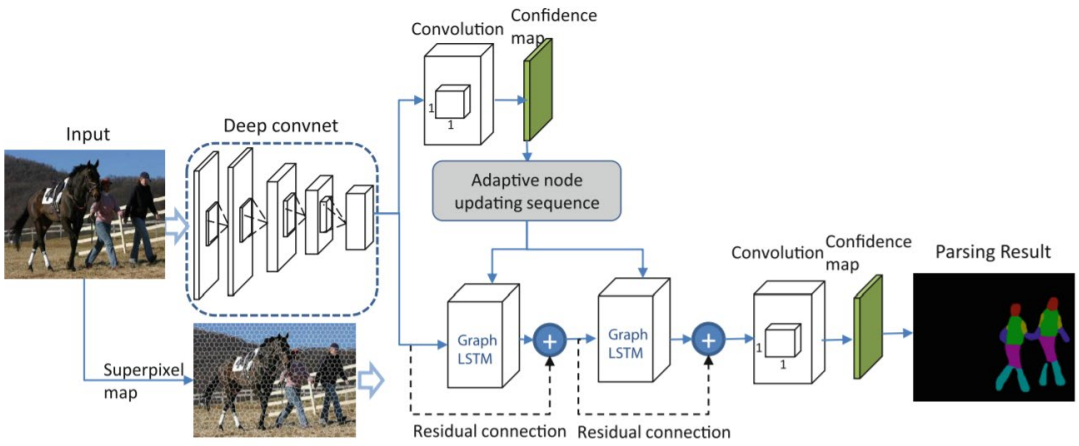
(d)


Encoder
Image
DCNN
Atrous Conv
1x1 Conv
3x3 Conv rate 6
3x3 Conv rate 12
3x3 Conv rate 18
Image Pooling
1x1 Conv
Decoder
Low-Level Features
Upsample by 4
1x1 Conv
Concat
3x3 Conv
Upsample by 4
Prediction

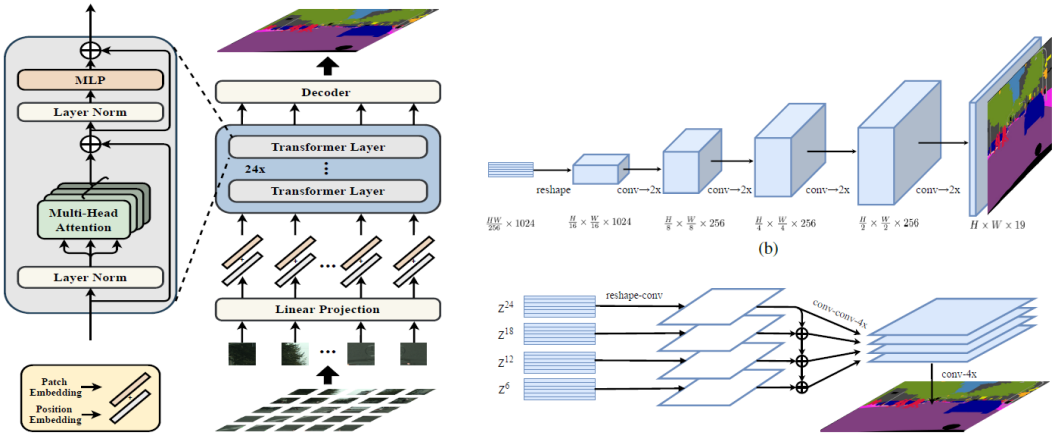
(e)


Position Attention Module
(HxW)x(HxW)
reshape
reshape
CxHxW
CxHxW
ResNet
Sum fusion
reshape
reshape
CxHxW
CxHxW
CxC
Channel Attention Module
convolution layer
Spatial attention matrix
Channel attention matrix
Spatial matrix operation
Channel matrix operation
Matrix multiplication
Element-wise Sum

(f)



(g)



(h)

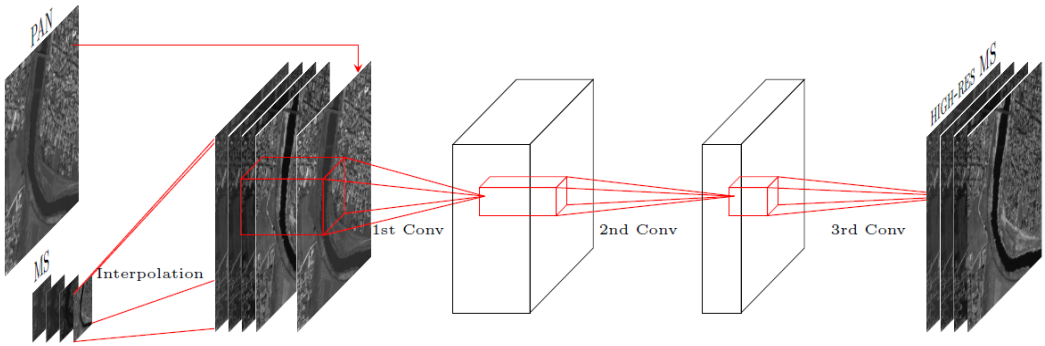
Figure 2.3 Deep-learning models for images segmentation: (a) FCN (Long et al., 2015); (b) CNN+Conditional Random Field (Chen et al., 2014); (c) SegNet (Badrinarayanan et al., 2017); (d) Pyramid Scene Parsing Network (Zhao et al., 2017); (e) DeepLabV3+ (Chen et al., 2018); (f) Dual Attention Network (Fu et al., 2019); (g) Graph-LSTM (Liang et al., 2016); (h) Segmentation Transformer (Zheng et al., 2021).

2.3.2 Pan-sharpening

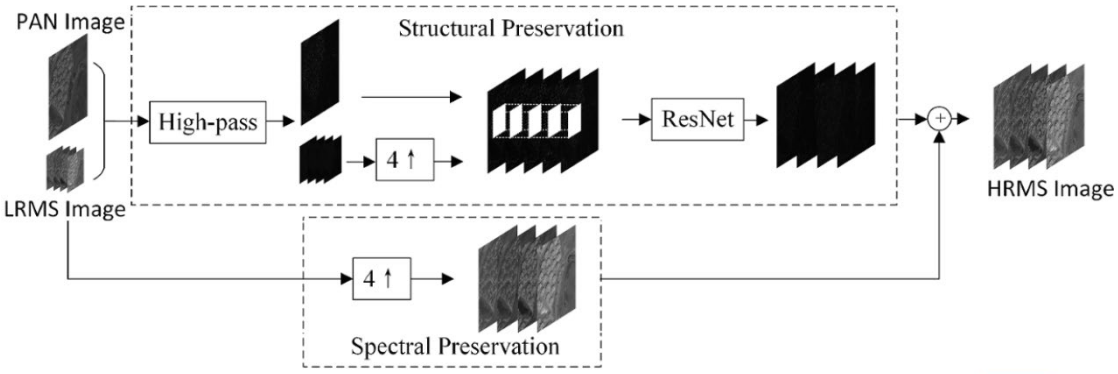
As a special instance of the more general problem of super-resolution, pan-sharpening aims to improve the low-spatial-resolution image by fusing with the data characterized by sharper spatial information. This field is dominated by works on the multispectral, thermal, or hyperspectral data with a panchromatic band. The earlier pan-sharpening techniques include component substitution, multi-resolution analysis, and model-based methods. Component substitution methods transform the low-spatial-resolution image into a suitable domain, where one of the components is replaced by the high-spatial-resolution panchromatic image (Dou et al., 2007). After upsampling the other components, the whole set is back-transformed into the original domain. Component substitution methods such as principal component analysis (Kwarteng and Chavez, 1989) and intensity-hue-saturation transform (Tu et al., 2001) are simple and easy to be implemented. Multi-resolution analysis methods like Laplacian pyramid (Aiazzi et al., 2002), discrete wavelet transform (Pradhan et al., 2006), curvelets (Ghahremani and Ghassemian, 2015) extract high-frequency details from the panchromatic images and inject them into the upsampled low-spatial-resolution images. These methods have better spectral fidelity than the projection substitution ones, while may exhibit spatial distortion and lower visual quality in the fused data. The model-based methods reformulate the pan-sharpening as an inverse problem and restore the high-spatial-resolution image from the degraded panchromatic and low-spatial-resolution images by resorting to some optimization procedures. Typical examples are the Bayesian method (Fasbender et al., 2008, Zhang et al., 2009) and the compressed sensing method (Zhu and Bamler, 2012, Zhu et al., 2015).

Recently, deep-learning has been adopted to address the challenges in pan-sharpening inspired by the great success achieved in image super-resolution. A CNN-based Pan-sharpening (PNN) model with three convolutional layers is first proposed based on the well-known Super-Resolution Convolutional Neural Network. PanNet used the skip connection in ResNet (He et al., 2016) and built a deeper network to learn the residual high-frequency features between the upsampled low-spatial-resolution image and the desired high-spatial-resolution image (Yang et al., 2017). By exploring the multiscale features extracted by different sizes of filters and multi-depth features from a shallow and a deep network, a multiscale and multi-depth convolutional neural network is further proposed to boost the performance (Yuan et al., 2018a). Super-resolution guided progressive pansharpening neural network uses a progressive pan-sharpening strategy to separate the spatial-resolution improvement process for a gradual and stable pan-sharpening (Cai and Huang, 2020). Two-stream Fusion Network fuses the panchromatic and low-spatial-resolution images in the feature domain and rebuilds the pan-sharpened image from the fused feature using a variant of U-Net (Ronneberger et al., 2015, Liu et al., 2020b). Pan-sharpening GAN (PSGAN) further improves TFNet with a generative and adversarial architecture (Liu et al., 2020a). All the methods are designed in a supervised fashion, which strongly depends on the training dataset and the availability of ground truth and hence restrains its application in the real scenario. Unsupervised pan-sharpening models are proposed to address this problem, where both Pan-GAN (Ma et al., 2020) and PGMAN (Zhou et al., 2021) are constructed based on GAN with a spatial and a spectral discriminator to refine the pan-sharpened image derived from the generator. The main difference is the generator of Pan-GAN uses a dense connection as in PNN, whereas PGMAN uses the rediual architecture in

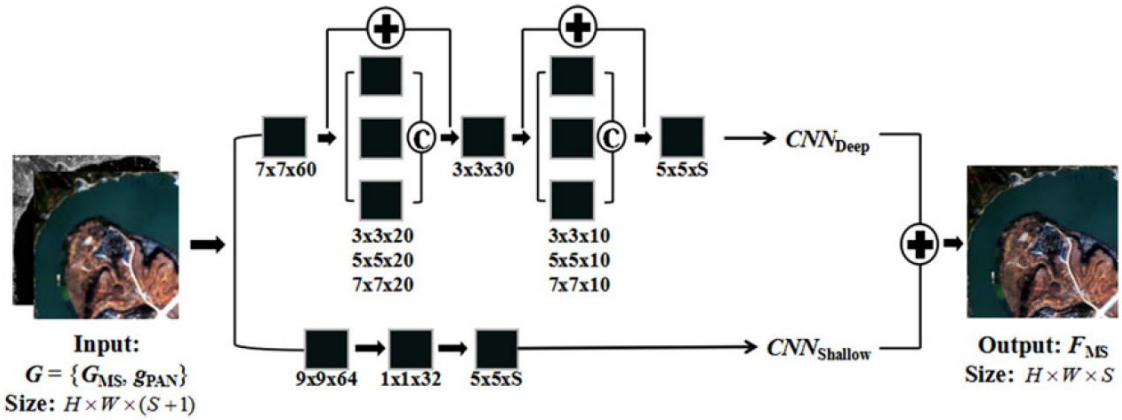
PanNet. Besides CNN, the self-attention mechanism (Vaswani et al., 2017) has also been applied in the unsupervised pansharpening (Qu et al., 2020).



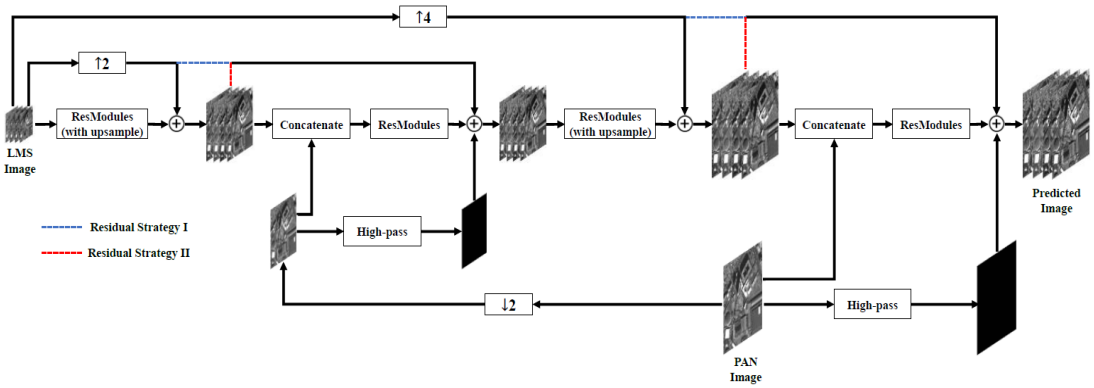
(a)



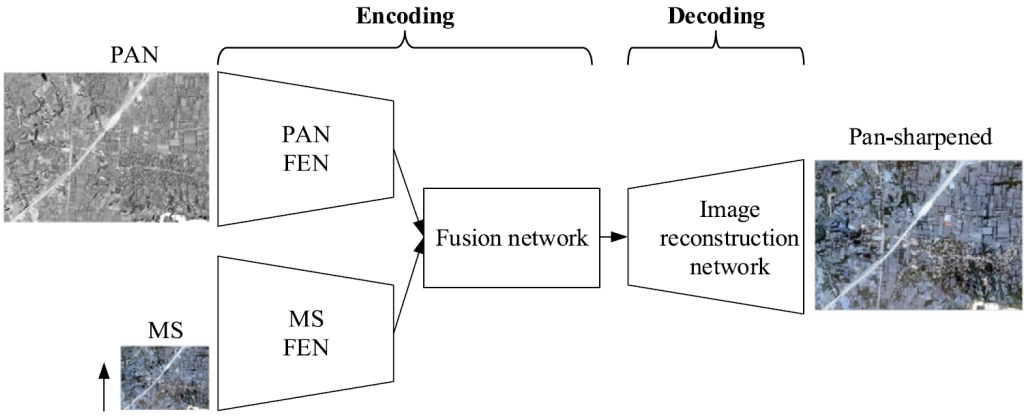
(b)



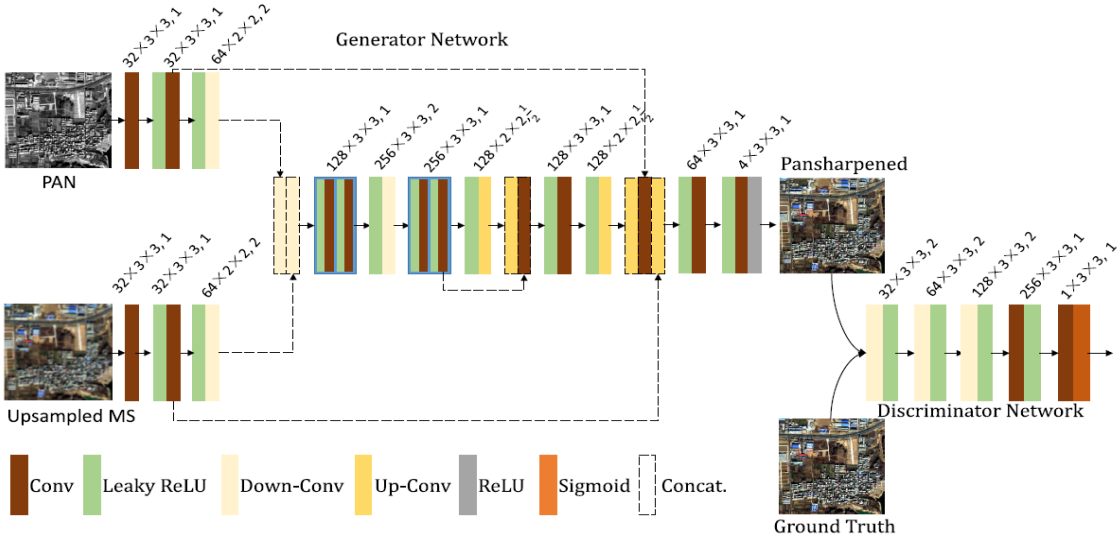
(c)


↑4
↑2
ResModules (with upsample)
Concatenate
ResModules
ResModules (with upsample)
Concatenate
ResModules
LMS Image
Predicted Image
Residual Strategy I
Residual Strategy II
High-pass
↓2
High-pass
PAN Image

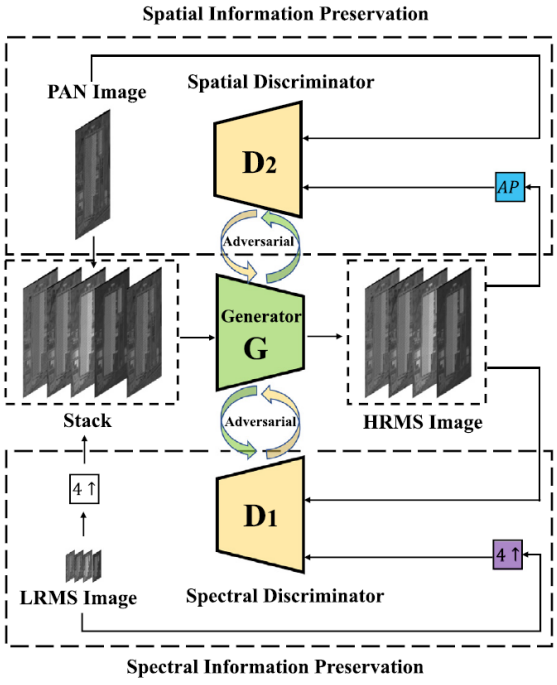
(d)


Encoding
Decoding
PAN
PAN FEN
Fusion network
Image reconstruction network
Pan-sharpened
MS
MS FEN

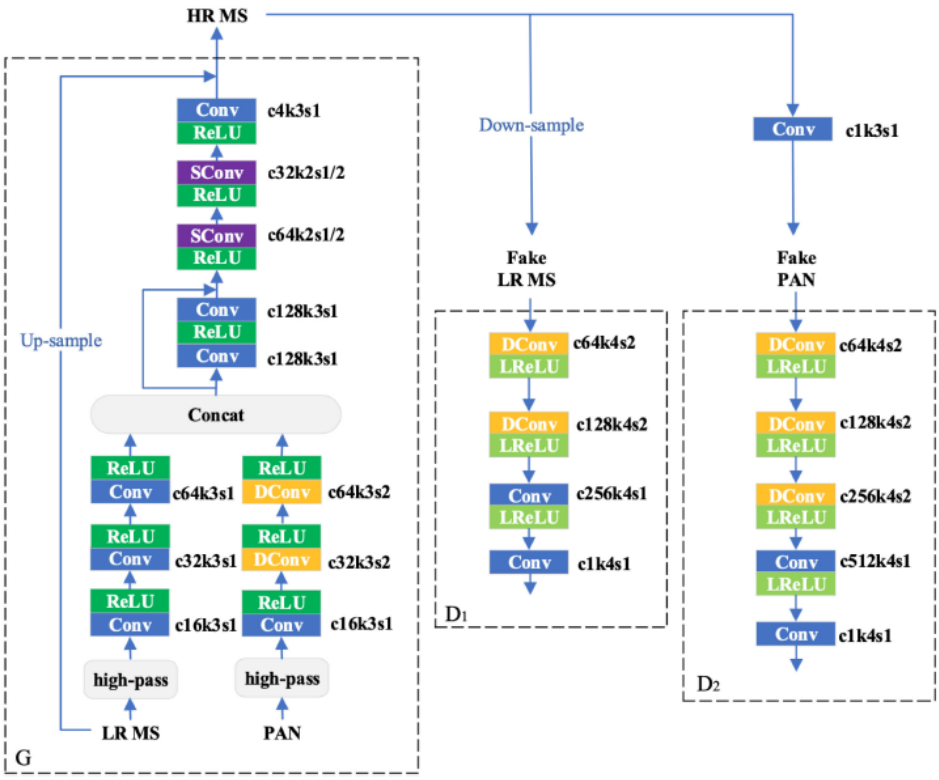
(e)


Generator Network
PAN
Upsampled MS
Pansharpened
Discriminator Network
Ground Truth
Conv
Leaky ReLU
Down-Conv
Up-Conv
ReLU
Sigmoid
Concat.

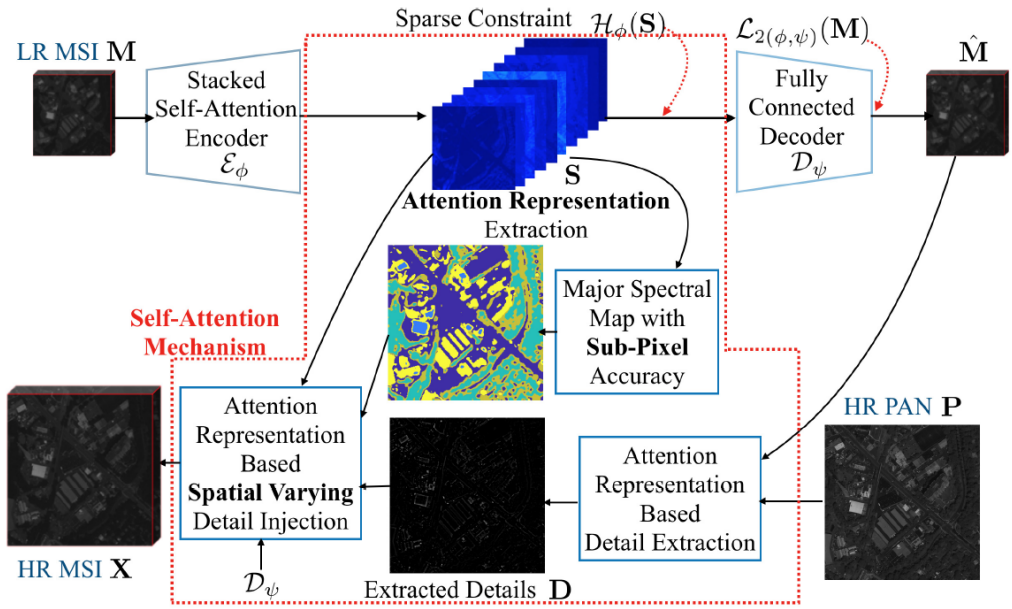
(f)


Spatial Information Preservation
PAN Image
Spatial Discriminator
D2
AP
Adversarial
Generator
G
Stack
HRMS Image
Adversarial
4 ↑
D1
4 ↑
LRMS Image
Spectral Discriminator
Spectral Information Preservation

(g)


HR MS
Conv
ReLU
c4k3s1
SConv
ReLU
c32k2s1/2
SConv
ReLU
c64k2s1/2
Conv
ReLU
Conv
c128k3s1
c128k3s1
Up-sample
Concat
ReLU
Conv
c64k3s1
ReLU
DConv
c64k3s2
ReLU
Conv
c32k3s1
ReLU
DConv
c32k3s2
ReLU
Conv
c16k3s1
ReLU
Conv
c16k3s1
high-pass
high-pass
LR MS
PAN
G
Down-sample
Conv
c1k3s1
Fake
LR MS
Fake
PAN
DConv
LReLU
c64k4s2
DConv
LReLU
c128k4s2
Conv
LReLU
c256k4s1
Conv
c1k4s1
D1
DConv
LReLU
c64k4s2
DConv
LReLU
c128k4s2
DConv
LReLU
c256k4s2
Conv
LReLU
c512k4s1
Conv
c1k4s1
D2

(h)



(i)

Figure 2.4 Deep-learning models for pan-sharpening: (a) PNN (Masi et al., 2016); (b) PanNet (Yang et al., 2017); (c) Multiscale and Multi-depth Convolutional Neural Network (Yuan et al., 2018a); (d) Super-resolution guided progressive pansharpening neural network (Cai and Huang, 2020); (e) Two-stream Fusion Network (Liu et al., 2020b); (f) PSGAN (Liu et al., 2020a); (g) Pan-GAN (Ma et al., 2020); (h) PGMAN (Zhou et al., 2021); (i) UP-SAM (Qu et al., 2020).

2.3.3 DNN Architectures

An overview of the DNN architectures used in this study including CNN, GAN, and the self-attention mechanism is presented in this section.

As the most commonly used architecture in the DL community, CNN is a multi-layer neural network proposed based on the animal visual cortex (Dargan et al., 2020). The first practical CNN architecture was developed by LeCun et al (LeCun et al., 2015) and has been widely applied in the field of image processing in the past few years. CNN is normally composed of three types of layers, convolutional layers to extract features, nonlinear layers applying activation function to the feature map, and pooling layers to reduce the spatial resolution based on certain statistical information derived from the neighbor pixels. The neuronal units in each layer are locally connected in receptive fields of shared wrights, which results in a significantly smaller number of parameters than fully connected neural networks (Zhu et al., 2017).

GAN was first proposed by Goodfellow (Goodfellow et al., 2014) designed based on game-theoretic principles. This architecture is composed of two adversarial networks including a generator and a discriminator. Taking the image generation as an example, the generator starts with Gaussian noise to produce images similar to the real samples, and the discriminator attempts to distinguish the generated fake samples from the real ones. This process continues until outputs of the generator are close enough to actual input samples, which can be characterized as a min-max game between the generator and the discriminator (Liu et al., 2020a).

As an integral component of Transformer, the self-attention mechanism is proposed in (Vaswani et al., 2017) to estimate the relevance of one item to the others in a given sequence. This is achieved by transforming the input sequence using three learnable weight matrices including query, key, and value. For a given entity in the sequence, the self-attention module computes the dot-product of the query with all keys, which is further divided by a scale factor and normalized by a softmax operator to obtain the weights on the values. Then the output of the self-attention module can be represented by the weighted sum of all entities in the sequence, which allows capturing long-term information and dependencies among each sequence element as compared to the RNN models (Khan et al., 2021).

Chapter 3 Forest Modeling for Canopy Height Estimation Using Airborne L-band Repeat-pass Dual-baseline PolInSAR Data

This chapter is dedicated to presenting the advances of PolInSAR-based forest height inversion and exploring the potential of model-based algorithms. All the experiments are carried out on airborne L-band repeat-pass acquisitions with temporal baselines of less than hours and a dual-baseline model-based method is specifically designed for the forest height estimation. This achievement is largely attributed to prior studies on the principle of microwave remote sensing (Ulaby et al., 1982, Tsang et al., 1985), the characterization of the electromagnetic scattering and radar processing (Treuhaft et al., 1996), and physical models expressing radar observations in terms of vegetation and ground parameters (Papathanassiou et al., 1998, Treuhaft and Siqueira, 2000). To characterize the interactions of radar waves with homogenous and heterogeneous scattering medium in the condition of airborne L-band repeat-pass configuration, the RMoG model (Lavalle and Hensley, 2015) is extended with different combinations among the Linear Volume Attenuation (LVA), Quadratic Volume Attenuation (QVA), Linear Volume Motion (LVM), and Quadratic Volume Motion (QVM) in the volume layer. The impacts of homogenous and heterogeneous attenuation and dynamic motion on the performance of forest height inversion are further explored thourgh PolInSAR coherence functions of the proposed model. To address the underdetermined problem between coherence observations and forest parameters, dual-baseline PolInSAR data is acquired to increase the degree of freedom (DOF) in the coherence observations and

thereby offers extra constraints on the forest parameters. A leveraging strategy is also employed during the forest height inversion to effectively utilize the advantages of LAV, QVA, LVM, and QVM depictions in the proposed model. Finally, experiments are performed on a boreal and a tropical forest site to test the proposed model and the associated inversion algorithm.

3.1 Introduction

PolInSAR provides an effective approach for forest characterization as it is sensitive to both physical properties and vertical structures of forest scattering medium (Moreira et al., 2009). This technique has been widely applied to deforestation monitoring, forest biomass estimation, and climate change investigation (Mette et al., 2004a, Mette et al., 2004b, Pardini et al., 2011). Prior studies have estimated the canopy height through a series of model-based methods applied to PolInSAR acquisitions. These sophisticated physical models establish mathematical relations between coherence observations and forest parameters from which the forest height can be estimated through the solution of non-linear equations (Garestier et al., 2008a, Garestier and Le Toan, 2009, Neumann et al., 2009b, Pichierri and Hajnsek, 2016, Lei et al., 2017). One mostly used model, the Random Volume over Ground (RVoG) model, represents the forest as a homogeneous volume of randomly oriented scattering particles statically distributed over the ground surface (Cloude and Papathanassiou, 2003). In this model, the volumetric attenuation of the radar scattering amplitude follows an exponential distribution with a constant extinction, and the PolInSAR coherence function derived is fully determined by four parameters including the ground phase, forest height, mean extinction coefficient, and ground to volume amplitude ratio. Several complementary

models have also been devised based on the RVoG model through integrating vertical heterogeneity into the forest representation (Garestier and Le Toan, 2009, Fu et al., 2017). These models have been applied to various forest types and sensor wavelengths using single-pass PolInSAR data (Praks et al., 2007, Garestier et al., 2008a, Garestier et al., 2008b, Hajnsek et al., 2009, Hensley et al., 2012, Kugler et al., 2015).

In the repeat-pass interferometry, temporal decorrelations derived from the dynamic changes of the vegetation and ground properties occurring between acquisition times affect the performance of forest parameter inversion. Hence, subsequent studies focus on overcoming the impact of temporal decorrelation by quantifying and compensating temporal components in repeat-pass PolInSAR coherences. The volumetric temporal decorrelation (RVoG+VTD) model was first proposed in (Papathanassiou and Cloude, 2003), describing the temporal decorrelation of the volume and ground layers as two complex coefficients embedded in the coherence function of the RVoG model. More physically-based models like Random Motion over Ground (RMoG) model (Lavalle and Hensley, 2012, Lavalle and Hensley, 2015) were designed to compensate for the temporal decorrelation components derived from the scatterer motion in the volume, e.g., due to the wind. A physical model accounting for both the position and dielectric property changes (Lei et al., 2017) was proposed based on the RMoG model, where the dielectric property changes of volume and ground layers were described by two complex coefficients integrated into the RMoG coherence function.

Each physical model mentioned above has its unique advantages when dealing with different types of PolInSAR acquisitions. Data involved in this study are collected by UAVSAR (Hensley et al., 2008), which is a fully polarimetric L-band radar system collecting repeat-pass interferometric data within a short temporal baseline (half an hour to hours) under various weather conditions. When data is collected with suitable weather conditions, dielectric changes derived from the moisture content variations in both the canopy and ground are negligible and it is reasonable to assume a stationary ground in the model. In this condition, the primary temporal decorrelation is assumed to be derived from wind-induced movements of the scatterers in the volume layer like in the RMoG model (Lee et al., 2012). Although the RMoG model has been applied to the forest height estimation using UAVSAR data in several studies (Lavalle et al., 2012, Lavalle and Hensley, 2015), the effect of different scattering attenuation and random motion characteristics on the inversion accuracy hasn't been investigated yet. This is especially relevant when dealing with SAR data acquired at lower frequencies such as L-band and P-band, where longer radar waves are more likely to interact with the scattering elements within the canopy and therefore more sensitive to the forest structure (Garestier and Le Toan, 2009). To this end, we considered both homogenous and heterogeneous scattering attenuation and random motion properties (LAV, QVA, LVM, QVM) in the proposed model and investigated their performance on the forest height inversion over different sites. Besides that, a pixel-wise optimization strategy is employed to leverage the advantages of different attenuation and motion profiles. The large footprint full-waveform Lidar data (Blair et al., 1999, Fatoyinbo et al., 2021) obtained by airborne LVIS system over boreal and tropical forest sites are used as the ground truth for the result analysis. Although LiDAR observations are much sparser

compared to the spatial resolution and coverage of radar images, they can characterize a variety of vegetation and terrain types with higher vertical accuracy.

This chapter is organized as follows. Test sites and datasets used in the experiments are shown in Section 3.2. Section 3.3 presents the proposed physical model, PolInSAR coherences, and forest height inversion method. Forest height inversion results over the boreal and tropical forest sites are shown in Section 3.4 following by a discussion in Section 3.5. Finally, conclusions are drawn in Section 3.6.

3.2 Test sites and datasets

3.2.1 Test sites

As shown in Figure **3.1**, test sites covered by boreal and tropical forests are involved in this experiment to investigate the applicability of the physical model with different volumetric profiles. This boreal forest site is part of the Boreal Ecosystem Research and Monitoring Sites (BERMS) in the Saskatchewan Province of Canada covering around 16 km along the east-west (104°46'38'' W to 105°1'15'' W) and 9 km along the north-south direction (53°52'29'' N to 53°56'10'' N). The southwest corner of the forest site (Candle Lake) is located ~70 km northeast of Prince Albert. Dominant tree species at the BERMS site are jack pine, black spruce, and aspen (Preston et al., 2014). The topography is relatively smooth with local elevation ranging from 504 m to 592 m above mean sea level (AMSL), and the mean annual precipitation is about 466 mm (Metsaranta and Lieffers, 2008b, Metsaranta and Lieffers, 2008a).

As presented in Figure **3.2**, the tropical test site is located in the Pongara National Park (PNP) on the south bank of Gabon Estuary. The image dataset covers a total area of around 12,008 ha within 0°4'15'' N to 0°7'50'' N and 9°45'20'' E to 9°55'1'' E. The dominant tree species in the PNP forest site are red mangroves covering about 55% of the total area (Giri et al., 2011). The PNP site exhibits a relatively smooth topography with an elevation range of less than 104 m. Annual precipitation over this area ranges from 2400 mm on the eastern portion to 2830 mm on the western side (Trettin et al., 2021), and the annual temperature ranges in a small interval averaging approximately 26.3 ± 0.9 °C (Hijmans et al., 2005).

3.2.2 Datasets

Three types of data including airborne SAR, airborne full-waveform LiDAR, and the simulated SAR data are involved in developing our forest height estimation method. The simulated SAR data are used to evaluate the sensitivity of the proposed theoretical models and airborne acquisitions are used as the real data for the algorithm validation.

The airborne data over the boreal forest site are collected as part of the Arctic Boreal Vulnerability Experiment (ABoVE) Airborne Campaign deployed by NASA's Earth Science Division (Miller et al., 2019). Starting in 2015, the ABoVE program tends to explore the vulnerability and resilience of the arctic and boreal forest to environmental change in North America (Greaves et al., 2016, Montesano et al., 2017, Bjorkman et al., 2018, Potter, 2018). As part of the ABoVE airborne data collections, airborne SAR

data over the BERMS forest site were acquired in late August 2018 and the associated airborne LiDAR data are collected by the LVIS in 2019.

The airborne data over the tropical site are collected within the AfriSAR campaign in 2016 (Fatoyinbo et al., 2021), which was deployed as a collaboration among National Park and international space agencies in support of ESA's BIOMASS, NASA-ISRO Synthetic Aperture Radar (NISAR), and the NASA's GEDI missions. Related airborne SAR and LiDAR data used in this study were collected over PNP in early February 2016.

3.2.2.1 UAVSAR

UAVSAR is a fully polarimetric L-band (1.26 GHz, 80 MHz bandwidth) SAR system typically flying at 12.5 km altitude and was built and operated by JPL (Hensley et al., 2008). This system is equipped with a precision autopilot system allowing for accurate repeat track acquisitions with short temporal baselines (Lavalle et al., 2017). Spatial resolutions of the UAVSAR single-look complex (SLC) image is 0.6 m in the azimuth direction and 1.6 m in slant range, which is further averaged by a rectangular window with sizes of 8 pixels in azimuth and 2 pixels in slant range to generate multi-looked polarimetric and interferometric image products with pixel spacings of 4.8 m in the azimuth direction and 3.2 m in slant range.

UAVSAR data over BERMS are collected in eight tracks with uniformly distributed vertical separations of 40-280 m in 40 m increments. These tracks were flown from

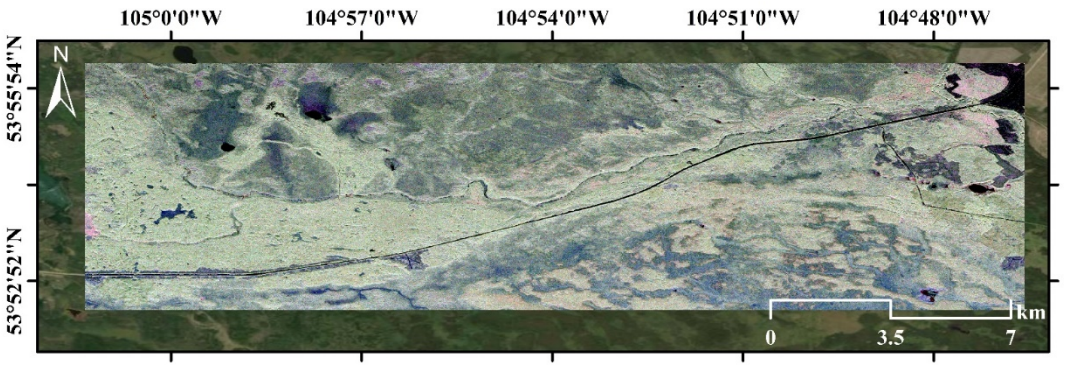
6:38:39 to 19:35:37 UTC on August 19, 2018, with roughly 25-minute time spacings providing temporal baselines of 25-175 minutes between pairs. The geocoded multi-looked SAR image in the Pauli color combination over BERMS is presented in Figure **3**.**1**(a).

Five repeat-pass tracks were acquired over PNP with nonuniform vertical baselines (0, 20, 45, 105 m) and roughly uniform temporal baselines from 12:30:11 to 14:13:21 UTC in 25-minute spacings on February 27, 2016. The geocoded multi-looked SAR image in the Pauli color combination over PNP is presented in Figure **3**.**2**(a).

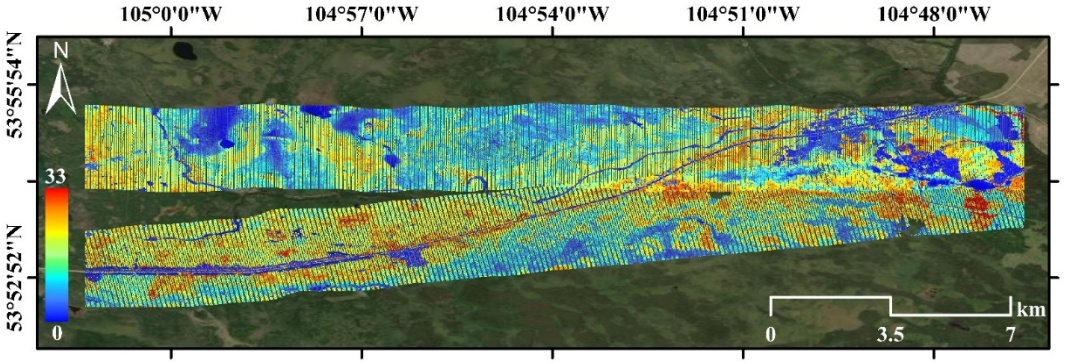
3.2.2.2 LVIS

LVIS, a medium-altitude imaging laser altimeter designed and developed by NASA's Goddard Space Flight Center, provides an accurate vertical structure of the canopy top and the underlying topography by digitally recording the returned signals (Blair et al., 1999). Relative Height 100 (RH100) metrics in the LVIS Level-2 collection represent heights above the detected ground surface at which 100% of the waveform energy has been returned and is typically associated with maximum tree height within a resolution beam of LiDAR (Simard and Denbina, 2017, Denbina et al., 2018). This study uses RH100 metrics as ground truth to evaluate the performance of forest height estimation. RH100 has been demonstrated to be a suitable forest height reference with RMSEs of 3.29 m in boreal forests and 2.87 m in tropical forests when compared with field data (Simard and Denbina, 2017, Denbina et al., 2018). The LVIS data over BERMS (10 m footprint diameter) were collected in July-August 2019 with a flight altitude of 12.5 km and a swath width of 2.5 km. Related data over the PNP (18 m footprint diameter)

were obtained in the same period as the UAVSAR data (February-March 2016) with a flight altitude of 7.3 km and a swath width of 1.5 km. RH100 heights in BERMS and PNP are illustrated in Figure **3.1**(b) and Figure **3.2**(b).

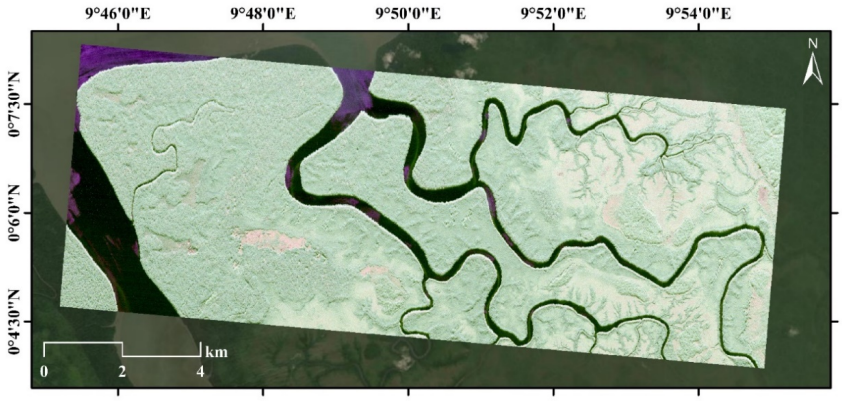


(a)

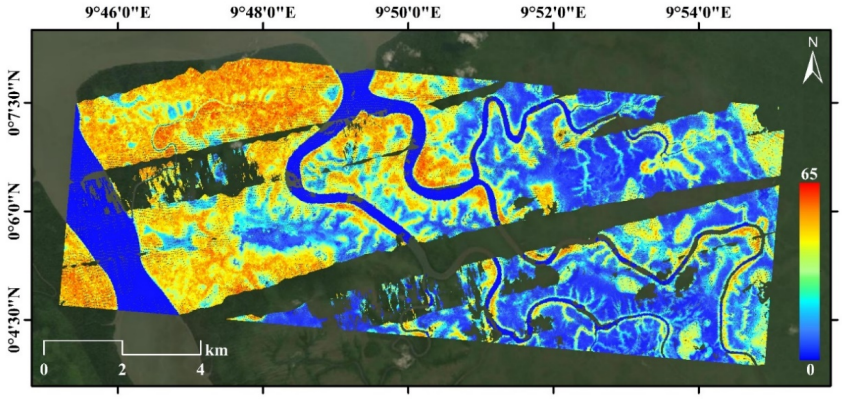


(b)

Figure 3.1 Study area in Boreal Ecosystem Research and Monitoring Sites (BERMS): (a) UAVSAR Pauli color composite image. Red: $|\boldsymbol{S_{HH}} - \boldsymbol{S_{VV}}|^2$, Green: $\mathbf{2}|\boldsymbol{S_{HV}}|^2$, Blue: $|\boldsymbol{S_{HH}} + \boldsymbol{S_{VV}}|^2$; (b) LVIS 2019 RH100 height.



(a)



(b)

Figure 3.2 Study area in the Pongara National Park (PNP): (a) UAVSAR Pauli color composite image; (b) LVIS 2016 RH100 height.

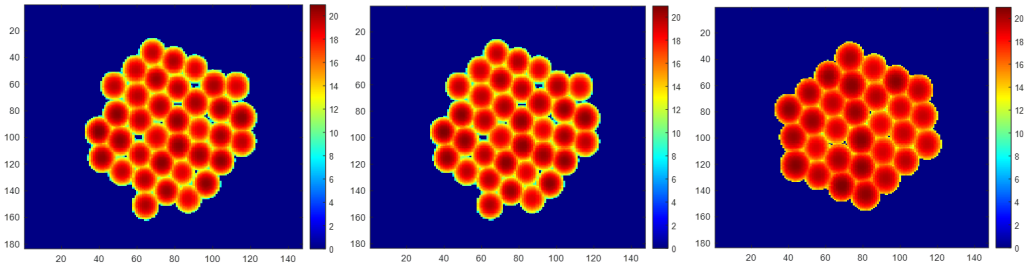
3.2.2.3 PolSARproSim+

Simulated SAR data generated by PolSARproSim+ are used to explore the scattering attenuation properties of different tree species. PolSARproSim+ is deployed on the distributed version of the ESA's software with improvements that allow more flexible simulations of different mission requirements, observational scenes, and instrument parameters (Hensley et al., 2014, Hensley et al., 2021). For example, this package allows the simulation of mixed forest stands with an external definition of allometric

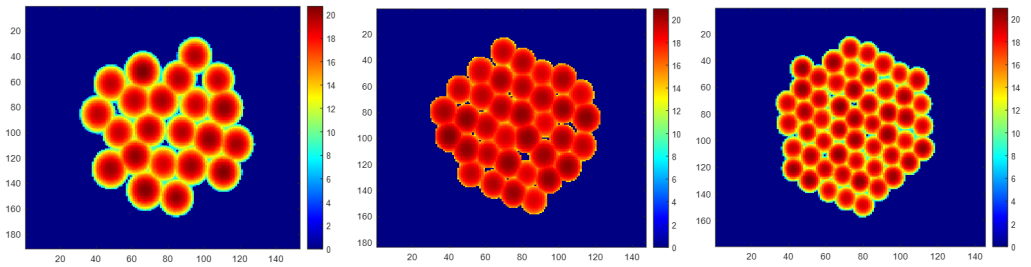
parameters while introducing temporal decorrelation effects originating from rain and wind. These extensions provide a powerful tool to explore the scattering attenuation properties in the forest medium and temporal decorrelation effects in the polarimetric interferometry and tomography. In this section, SAR simulations are performed using the system parameter of UAVSAR over 14 deciduous and 4 coniferous forest stands with average heights of 20 m as listed in Table 3.1. Related forest stand simulations of these 18 species are presented in Figure 3.3 and the average normalized scattering attenuation profiles are plotted versus height in Figure 3.4.

Table 3.1 Parameters of the simulated forest stands

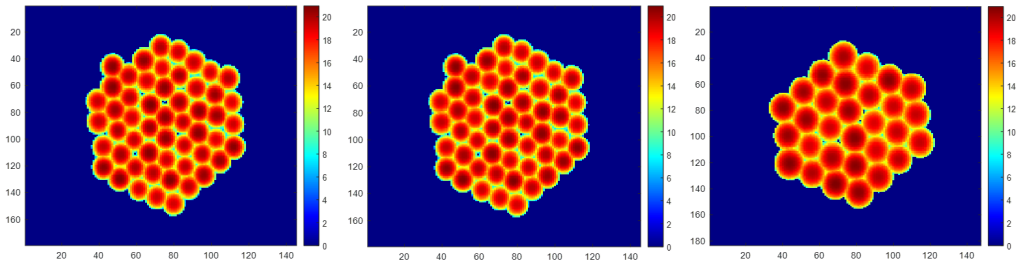
Tree species	**Type**	**Density (trees/ha)**
Striped maple (ACPE)	Deciduous	36
Red maple (ACRU)	Deciduous	36
Sugar maple (ACSA)	Deciduous	33
Yellow birch (BEAL)	Deciduous	19
Brazil nut (BEEX)	Deciduous	36
Sweet birch (BELE)	Deciduous	51
Paper birch (BEPA)	Deciduous	51
Grey birch (BEPO)	Deciduous	51
American beech (FAGR)	Deciduous	33
American hophornbeam (OSVI)	Deciduous	60
Red pine (PIRE)	Coniferous	44
Red spruce (PIRU)	Coniferous	25
White pine (PIST)	Coniferous	26
Black cherry (PRSE)	Deciduous	51
White oak (QUAL)	Deciduous	27
Red oak (QURU)	Deciduous	28
Red mangrove (RHMA)	Deciduous	69
Eastern hemlock (TSCA)	Coniferous	15



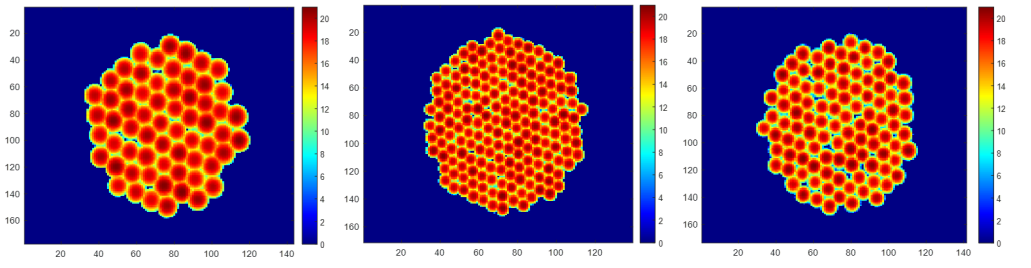
(a) (b) (c)



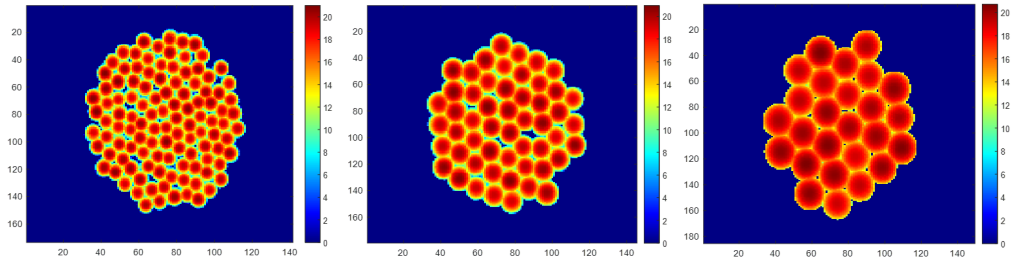
(d) (e) (f)



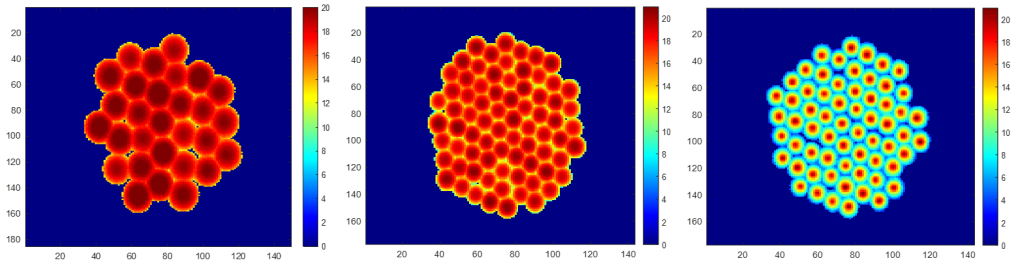
(g) (h) (i)



(j) (k) (l)

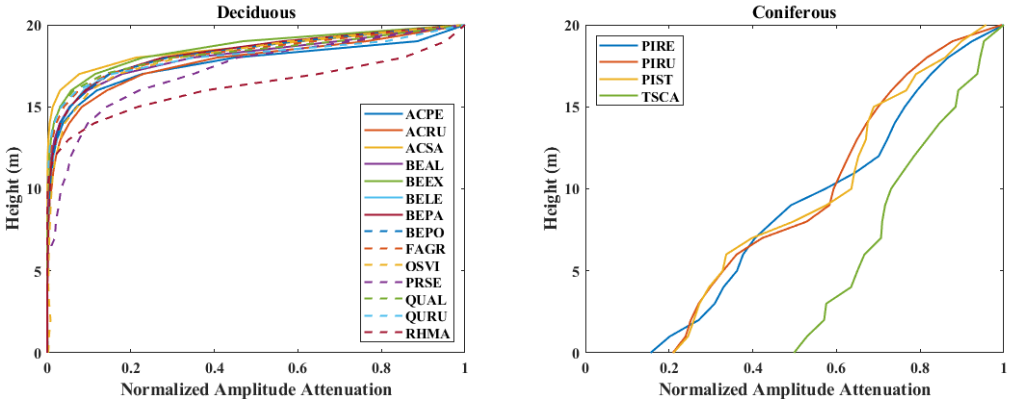


(m) (n) (o)



(p) (q) (r)

Figure 3.3 Simulated forest stands of 14 deciduous and 4 coniferous tree species. (a) ACPE; (b) ACRU; (c) ACSA; (d) BEAL; (e) BEEX; (f) BELE; (g) BEPA; (h) BEPO; (i) FAGR; (j) OSVI; (k) PIRE; (l) PIRU; (m) PIST; (n) PIST; (o) QUAL; (p) QURU; (q) RHMA; (r) TSCA.



(a) (b)

Figure 3.4 Normalized amplitude attenuation profile of different forest species. (a) Deciduous; (b) Coniferous.

3.3 Methodology

3.3.1 Physical model

As shown in Figure 3.5, forest stands are simulated as volumes of dynamic scatterers with either homogeneous or heterogeneous random motion and scattering attenuation

over a stable ground surface in repeat-pass UAVSAR PolInSAR configuration. The dynamic volume with a homogeneous scattering attenuation and random motion has been previously modeled in the RMoG model (Lavalle and Hensley, 2015). Dynamic volumes with heterogeneous attenuation and motion are given in Figure 3.5 depicted by the gradient color in the volume and arrows with different lengths, respectively.

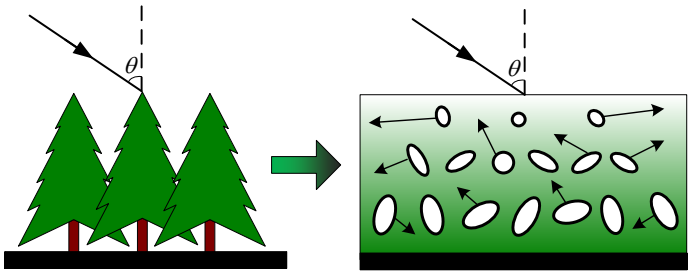


Figure 3.5 Schematic of the heterogeneous physical model

In the RMoG model, an exponential function with a constant mean extinction is used to describe the homogeneous scattering attenuation (LVA), while the homogeneous random motion of the dynamic scatterers is integrated as a Gaussian-statistic function with zero mean and vertically linear variance (LVM). In this condition, the scattering attenuation ρ_v and random motion η_v profiles of the volume layer are given by

$$\rho_v(z) = exp\left(-\frac{2\sigma_{LVA}}{cos(\theta)}(h_v - z)\right) \tag{3.1}$$

$$\eta_v(z) = exp(-\tau_{LVM} z) \tag{3.2}$$

where σ_{LVA} is the constant mean extinction, h_v is the forest height, θ is the incident angle, and τ_{LVM} is the gradient of the Gaussian motion variance.

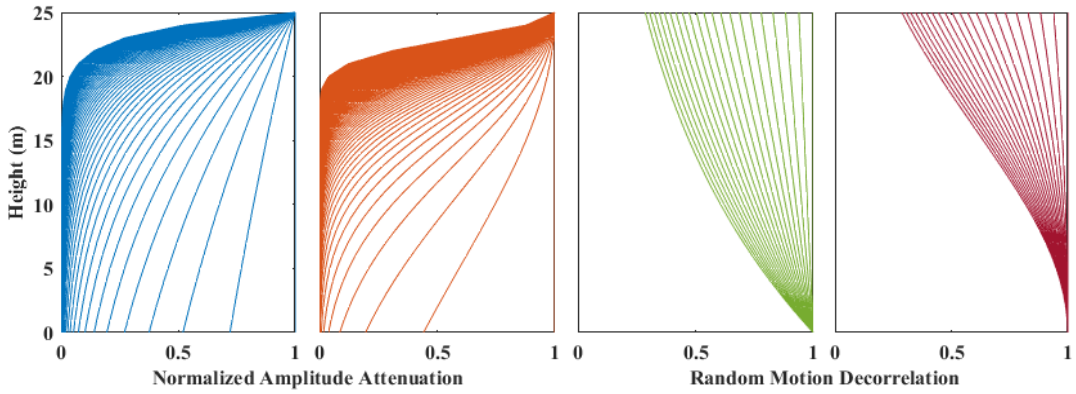
We further explore the suitability of a different functional form for the description of the scattering attenuation and random motion profiles.

$$\rho_v(z) = exp\left(-\frac{2\sigma_{QVA}(h_v - z)}{cos(\theta)}(h_v - z)\right) \tag{3.3}$$

$$\eta_v(z) = exp\left(-\tau_{QVM} z^2\right) \tag{3.4}$$

where σ_{QVA} is the gradient of the vertically linear extinction and τ_{QVM} is the second derivative of the Gaussian motion variance. Equation (3.3)-(3.4) give a heterogeneous scattering attenuation in the form of an exponential function with a vertically linear extinction (QVA) and a heterogeneous motion in the form of a Gaussian distribution with zero mean and vertically quadratic variance (QVM).

Scattering attenuation and random motion shape evolutions in Equations (3.1)-(3.4) are plotted versus volume heights ranging from 0 to 25 m for an incidence angle of 40° in Figure 3.6.



(a) (b)

Figure 3.6 Shape evolutions of the volumetric scattering attenuation and random motion. (a) LVA (left) and QVA (right); (b) LVM (left) and QVM (right).

3.3.2 PolInSAR coherence function

As an evaluation of the interferometric quality, the complex PolInSAR coherence is statistically calculated by

$$\hat{\gamma} = \frac{\langle \vec{\omega}_1^{\dagger} \Omega_{12} \vec{\omega}_2 \rangle}{\langle \vec{\omega}_1^{\dagger} T_{11} \vec{\omega}_1 \rangle \langle \vec{\omega}_2^{\dagger} T_{22} \vec{\omega}_2 \rangle} = |\hat{\gamma}| \exp(j\hat{\phi}) \tag{3.5}$$

where, $\vec{\omega}_1$ and $\vec{\omega}_2$ are the complex unitary vectors in a certain transmit and receive polarization, Ω_{12} is a 3×3 non-hermitian complex matrix containing both polarimetric and interferometric information, T_{11} and T_{22} are 3×3 hermitian coherency matrix describing polarimetric properties, $\hat{\phi}$ is the expected interferometric phase, and $|\hat{\gamma}|$ is the coherence magnitude related to the phase noise (Papathanassiou and Cloude, 2001).

After compensating spectral and SNR decorrelations (Kugler et al., 2015), the volume and temporal decorrelations are the two main components in the repeat-pass PolInSAR coherence. These two are directly related to the vertical structure, temporal stability, and physical characteristics of the forest scattering medium, and the compensated PolInSAR coherence can be further approximated as a coherence function based on the two-layer dynamic scattering model (Lavalle and Hensley, 2015).

$$\gamma = \exp(j\phi_g) \frac{\gamma_{vt} + \mu(\vec{\omega})}{1 + \mu(\vec{\omega})} \tag{3.6}$$

where, ϕ_g represents the ground phase, γ_{vt} is the coherence component with coupled effects from the volumetric scatterer and the wind-derived temporal change, and $\mu(\vec{\omega})$ is the ground to volume amplitude ratio varying with polarization $\vec{\omega}$.

Temporal changes in the ground layer are neglected in this coherence function because of the short temporal baselines of the UAVSAR acquisitions. The volumetric-temporal coherence can be further described by

$$\gamma_{vt} = \frac{\int_0^{h_v} \rho_v(z)\eta_v(z)\exp(jk_z z)dz}{\int_0^{h_v} \rho_v(z)dz} \tag{3.7}$$

where k_z is the vertical wavenumber given by

$$k_z = \frac{4\pi}{\lambda \sin\theta}\Delta\theta \approx \frac{4\pi B_\perp}{\lambda \sin\theta R} \tag{3.8}$$

where λ is the radar wavelength, $\Delta\theta$ is the difference between the repeat-pass incident angles, $B_\perp$ is the perpendicular baseline, and R is the slant range.

By substituting Equations (3.1)-(3.4) into (3.7), the volumetric-temporal coherences for physical models with different combinations of LVM, LVA, QVM, and QVA are after some algebraic manipulation given by

$$\gamma_{vt}(h_v, \sigma_{LVA}, \tau_{LVM}) = \frac{\frac{2\sigma_{LVA}}{\cos(\theta)}\left(\exp\left(\frac{2\sigma_{LVA}}{\cos(\theta)} + jk_z - \tau_{LVM}\right)h_v - 1\right)}{\left(\frac{2\sigma_{LVA}}{\cos(\theta)} + jk_z - \tau_{LVM}\right)\left(\exp\left(\frac{2\sigma_{LVA}}{\cos(\theta)}h_v\right) - 1\right)} \tag{3.9}$$

$$\gamma_{vt}(h_v, \sigma_{LVA}, \tau_{QVM}) = \sigma_{LVA}\sqrt{\frac{\pi}{\tau_{QVM}}}\exp\left(\frac{\left(\frac{2\sigma_{LVA}}{\cos(\theta)} + jk_z\right)^2}{4\tau_{QVM}}\right)$$

$$\cdot \frac{\left[\operatorname{erf}\left(\frac{2h_v\tau_{QVM} - \left(\frac{2\sigma_{LVA}}{\cos(\theta)} + jk_z\right)}{2\sqrt{\tau_{QVM}}}\right) + \operatorname{erf}\left(\frac{\frac{2\sigma_{LVA}}{\cos(\theta)} + jk_z}{2\sqrt{\tau_{QVM}}}\right)\right]}{\cos(\theta)\left(\exp\left(\frac{2\sigma_{LVA}}{\cos(\theta)}h_v\right) - 1\right)} \tag{3.10}$$

$$\gamma_{vt}(h_v, \sigma_{QVA}, \tau_{LVM}) = \exp\left(\frac{\left(\frac{4\sigma_{QVA}}{\cos(\theta)}h_v + jk_z - \tau_{LVM}\right)^2}{\frac{8\sigma_{QVA}}{\cos(\theta)}}\right)$$

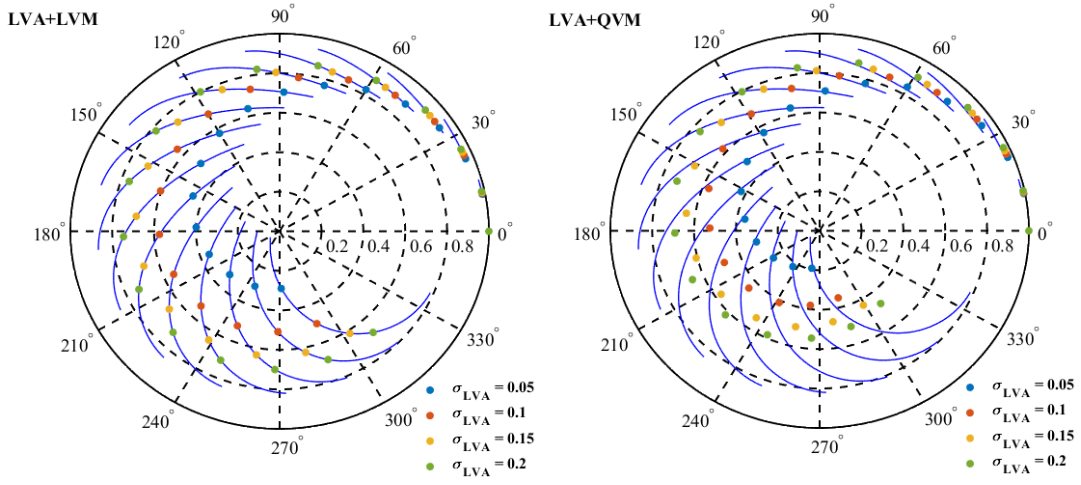
$$\cdot\frac{\left[\operatorname{erf}\left(\frac{\frac{4\sigma_{QVA}}{\cos(\theta)}h_v + jk_z - \tau_{LVM}}{\sqrt{\frac{8\sigma_{QVA}}{\cos(\theta)}}}\right) - \operatorname{erf}\left(\frac{jk_z - \tau_{LVM}}{\sqrt{\frac{8\sigma_{QVA}}{\cos(\theta)}}}\right)\right]}{\exp\left(\frac{2\sigma_{QVA}}{\cos(\theta)}h_v^2\right)\left[\operatorname{erf}\left(\sqrt{\frac{2\sigma_{QVA}}{\cos(\theta)}h_v^2}\right)\right]} \tag{3.11}$$

$$\gamma_{vt}(h_v, \sigma_{QVA}, \tau_{QVM}) = \sqrt{\frac{2\sigma_{QVA}}{\cos(\theta)}}\exp\left(\frac{\left(\frac{4\sigma_{QVA}}{\cos(\theta)}h_v + jk_z\right)^2}{\frac{8\sigma_{QVA}}{\cos(\theta)} + 4\tau_{QVM}}\right)$$

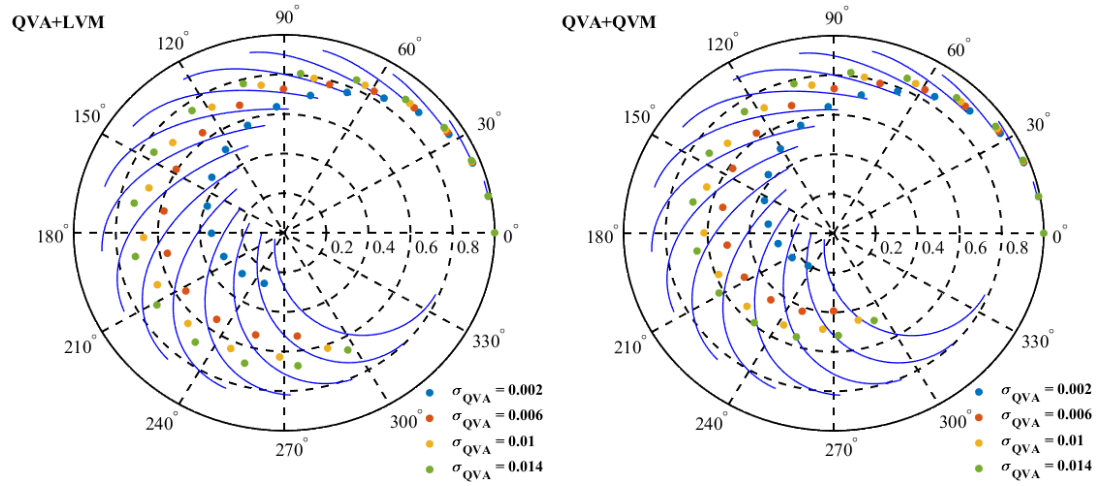
$$\cdot\frac{\left[\operatorname{erf}\left(\frac{2\tau_{QVM}h_v - jk_z}{\sqrt{\frac{8\sigma_{QVA}}{\cos(\theta)} + 4\tau_{QVM}}}\right) + \operatorname{erf}\left(\frac{\frac{4\sigma_{QVA}}{\cos(\theta)}h_v + jk_z}{\sqrt{\frac{8\sigma_{QVA}}{\cos(\theta)} + 4\tau_{QVM}}}\right)\right]}{\sqrt{\frac{2\sigma_{QVA}}{\cos(\theta)} + \tau_{QVM}}\exp\left(\frac{2\sigma_{QVA}}{\cos(\theta)}h_v^2\right)\left[\operatorname{erf}\left(\sqrt{\frac{2\sigma_{QVA}}{\cos(\theta)}h_v^2}\right)\right]} \tag{3.12}$$

The volumetric-temporal coherences described in Equations (3.9)-(3.12) are plotted in the complex plane with volume heights ranging from 0 to 25 m, a center frequency of 1.26 GHz, a vertical wavenumber of 0.2 for an incidence angle of 40° as in Figure 3.6. Each blue segment in Figure 3.7 corresponds to an RMoG volume with LVA and LVM of a given height and the mean extinction from 0 to 1 dB/m, where volumes with lower extinction have smaller radial distances and thereby lower coherences. Compared with

the blue segments of the RMoG model, the introduction of QVA and QVM in Figure 3.7(b)-(d) has little impact on volumetric-temporal coherences for lower heights and extinctions. However, larger differences are observed in volumes with higher forest heights and extinctions. Figure 3.7(b) illustrates that introducing QVM accelerates the loss of coherence magnitude with the increasing height, especially in volumes with higher extinction. As indicated in Figure 3.7(c), the introduction of QVA leads to a smaller interferometric phase at the same coherence magnitude level as RMoG volume, especially in the condition of higher extinction. Volumetric-temporal coherences that originated from the synergy of QVA and QVM are illustrated in Figure 3.7(d), where smaller interferometric phase and coherence magnitude can be observed at the same time compared with the RMoG volumes.



(a) (b)



(c) (d)

Figure 3.7 Volumetric-temporal coherence varying with different extinction coefficients: (a) LVA+LVM; (b) LVA+QVM; (c) QVA+LVM; (d) QVA+QVM

Each blue segment in Figure 3.8 corresponds to an RMoG volume with LVA and LVM of a given height and the random motion coefficient ranging from 0 to 0.1, where volumes with stronger random motions have smaller radial distances. Similar variation can be observed in Figure 3.8 as in Figure 3.7, such as the introduction of QVM has a stronger impact on the coherence magnitude than the interferometric phase. As shown in Figure 3.8(b), coherence magnitudes drop faster with the increased height. Likewise, the introduction of QVA causes a smaller interferometric phase for the same coherence magnitude. The volumetric-temporal coherences in Figure 3.8(d) are originated from the synergy of QVA and QVM, which also exhibit smaller interferometric phases and coherence magnitudes at the same time compared with the RMoG volumes.

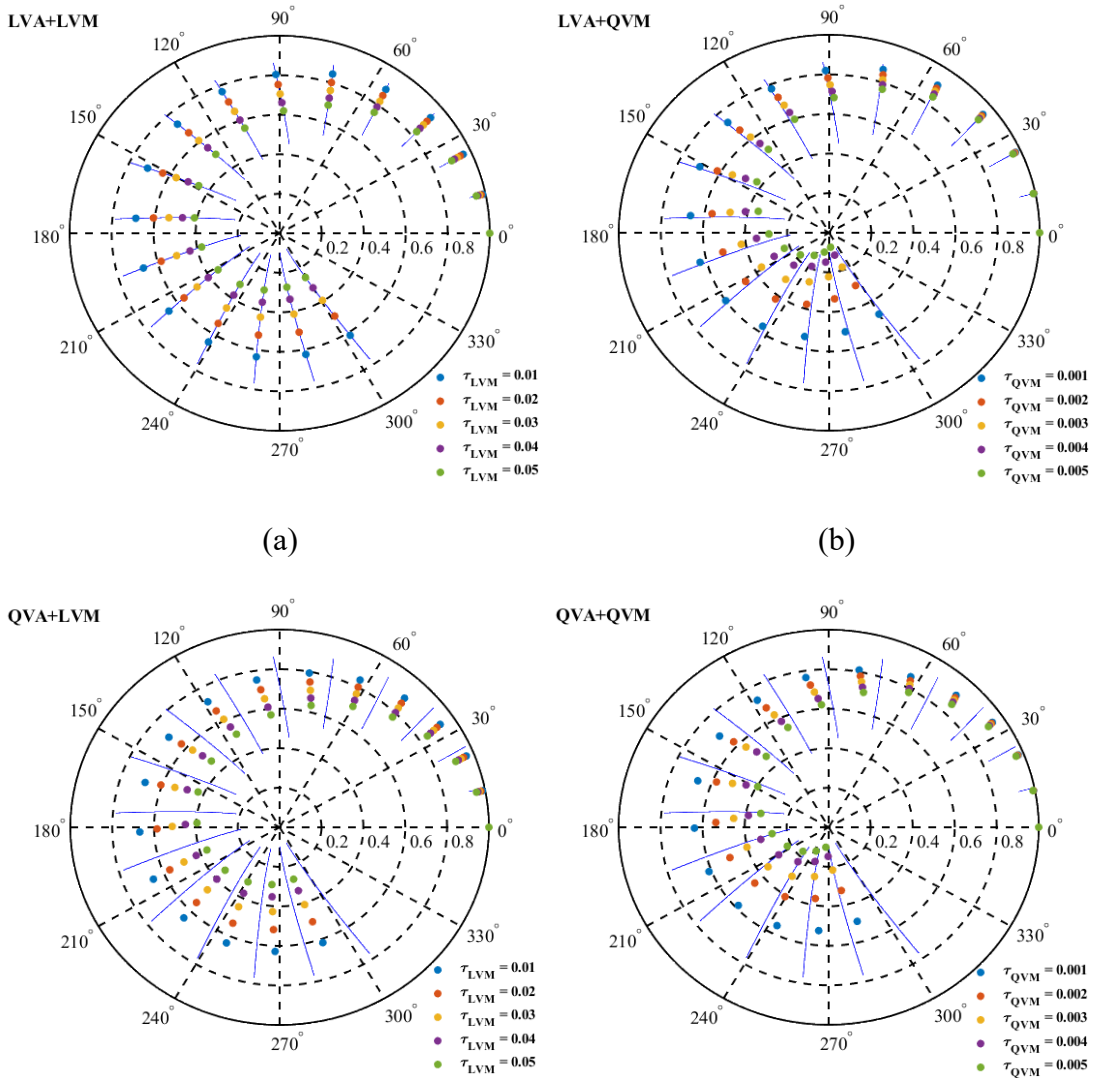


Figure 3.8 Volumetric coherence varying with different random motion coefficients: (a) LVA+LVM; (b) LVA+QVM; (c) QVA+LVM; (d) QVA+QVM

3.3.3 Forest height inversion

The PolInSAR coherence of a two-layer dynamic scattering model is fully determined by five parameters including the ground phase, forest height, random motion factor, ground to volume amplitude ratio, and the scattering attenuation coefficient. Three of them (attenuation coefficient, random motion factor, forest height) are accounted in

the polarization-independent volumetric-temporal coherence term equivalent to the PolInSAR coherence in the absence of ground backscatter contributions with ground phase removed. Equation (3.6) indicates a linear signature of coherences predicted by the proposed model on the unit circle, which can be further represented by a straight line varying with the ground-to-volume ratio in the complex plane. Theoretically, the value of the ground-to-volume ratio ranges from zero to infinity. When it is zero, all scattering components are derived from the volume layer and the complex PolInSAR coherence is simply given by $\hat{\gamma}_{vt}\exp(j\hat{\phi}_g)$. When the ground to volume ratio comes to infinite, all the scattering contributions are from the ground and related PolInSAR coherence is $\exp(j\hat{\phi}_g)$. The line fitting can be performed by a line regression of the complex coherence loci at a set of possible polarization channels including the typical linear and Pauli basis polarization channels (HH, HV, VV, HH+VV, HH-VV), and the phase diversity (PD) optimized polarizations with the highest and lowest phase center (Tabb et al., 2002). This fitted line intersects the unit circle at two points and one of them is the correct ground solution. We select the correct one by assuming the height of the highest observed phase center is less than π/k_z m above the ground as indicated in (Kugler et al., 2015, Denbina et al., 2018). After that, $\exp(j\hat{\phi}_g)\hat{\gamma}_{vt}$ is determined by the furthest coherence point from the estimated ground solution and the volumetric-temporal coherence $\hat{\gamma}_{vt}$ can be estimated by removing the ground phase.

However, one volumetric-temporal coherence estimation can only provide magnitude and phase information with two DOFs and thereby causes underdetermination for the inversion of three unknown parameters. Therefore, an extra pass of UAVSAR data is

acquired to form the dual-baseline repeat-pass interferometry. Each interferometry can provide one volumetric-temporal coherence estimation through the above-mentioned method. Since repeat-pass UAVSAR data are acquired at slightly different altitudes and incidence angles within hours, we assume a stable attenuation property (a constant attenuation coefficient) in the forest canopy within different passes. In this condition, one additional interferometry can provide volumetric-temporal coherence estimations with two more DOFs while only introducing one additional unknown motion factor. Then, the forest height can be estimated by solving nonlinear equations using iterative optimization to minimize the least square distance between the predicted and observed volumetric-temporal coherences.

$$\begin{cases} \hat{\gamma}_{vt_1} = f_i(\rho_v, \eta_{v_1}, h_v) \\ \hat{\gamma}_{vt_2} = f_i(\rho_v, \eta_{v_2}, h_v) \end{cases}, \text{with} \quad i = 1,2,3,4 \tag{3.13}$$

where, f_i is the volumetric-temporal coherence function as presented in Equations (3.9) -(3.12).

Equations (3.9)-(3.12) provide four descriptions of the volumetric-temporal coherence for the forest height inversion. Each one corresponds to a certain scattering attenuation and random motion profiles in the two-layer dynamic scattering model. To leverage various scattering attenuation and random motion profiles, we choose the forest height that derived from the coherence function f_i with the minimum distance between the predicted and observed volumetric-temporal coherences in each pixel during the forest height inversion, that is

$$\min_{f_i} \left\{ \left| f_i(\rho_v, \eta_{v_1}, h_v) - \hat{\gamma}_{vt_1} \right|^2 + \left| f_i(\rho_v, \eta_{v_2}, h_v) - \hat{\gamma}_{vt_2} \right|^2 \right\}, \text{with} \quad i = 1,2,3,4 \tag{3.14}$$

Comparative experiments are conducted between the proposed dual-baseline method and the prior Cloude's dual-baseline method based on the RVoG model (Cloude, 2002) to validate the superiority of the proposed method. Forest height estimation results are quantitatively evaluated by computing the bias and Root Mean Square Error (RMSE) with the LVIS RH100 trees height, given by

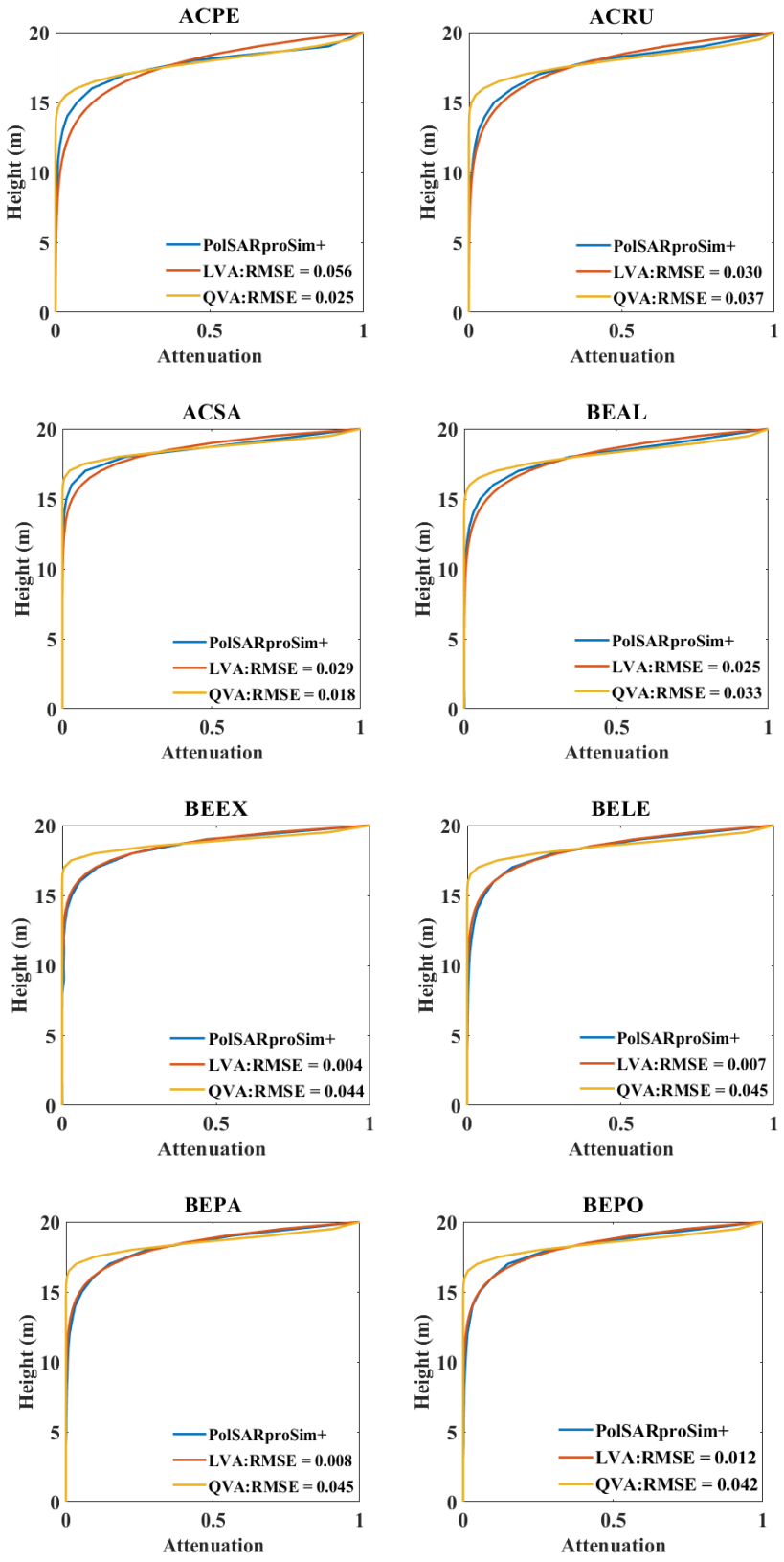
$$Bias = \frac{\sum_{i=1}^{n} h_v(i)}{n} - \frac{\sum_{i=1}^{n} ref(i)}{n} \tag{3.15}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(h_v(i) - ref(i)\right)^2}{n}} \tag{3.16}$$

3.4 Result

3.4.1 Scattering attenuation fitting

PolSARproSim+ allows building attenuation grids using the effective permittivities determined by integrating forward scattering amplitudes of the plant elements and interpreting them as effective polarizabilities (Hensley et al., 2014). These attenuation grids can provide approximate estimates of the attenuation by location during the scattering calculation, which maintains the heterogeneity of the tree crown distribution. To assess the applicability of LVA and QVA to a particular tree species, the forest stands and the scattering attenuation maps of fourteen deciduous and four coniferous species in Table 3.1 are generated using tree models embedded in the PolSARProSim+. Average attenuation profiles derived from these output attenuation maps are plotted as blue curves in Figure 3.9 and Figure 3.10. The red and yellow curves are the best-fit LVA and QVA to the generated average attenuation profiles. Related RMSEs of the fitting results are also shown in the plots in Figure 3.9 and Figure 3.10.


ACPE
Height (m)
Attenuation
PolSARproSim+
LVA:RMSE = 0.056
QVA:RMSE = 0.025
ACRU
PolSARproSim+
LVA:RMSE = 0.030
QVA:RMSE = 0.037
ACSA
PolSARproSim+
LVA:RMSE = 0.029
QVA:RMSE = 0.018
BEAL
PolSARproSim+
LVA:RMSE = 0.025
QVA:RMSE = 0.033
BEEX
PolSARproSim+
LVA:RMSE = 0.004
QVA:RMSE = 0.044
BELE
PolSARproSim+
LVA:RMSE = 0.007
QVA:RMSE = 0.045
BEPA
PolSARproSim+
LVA:RMSE = 0.008
QVA:RMSE = 0.045
BEPO
PolSARproSim+
LVA:RMSE = 0.012
QVA:RMSE = 0.042

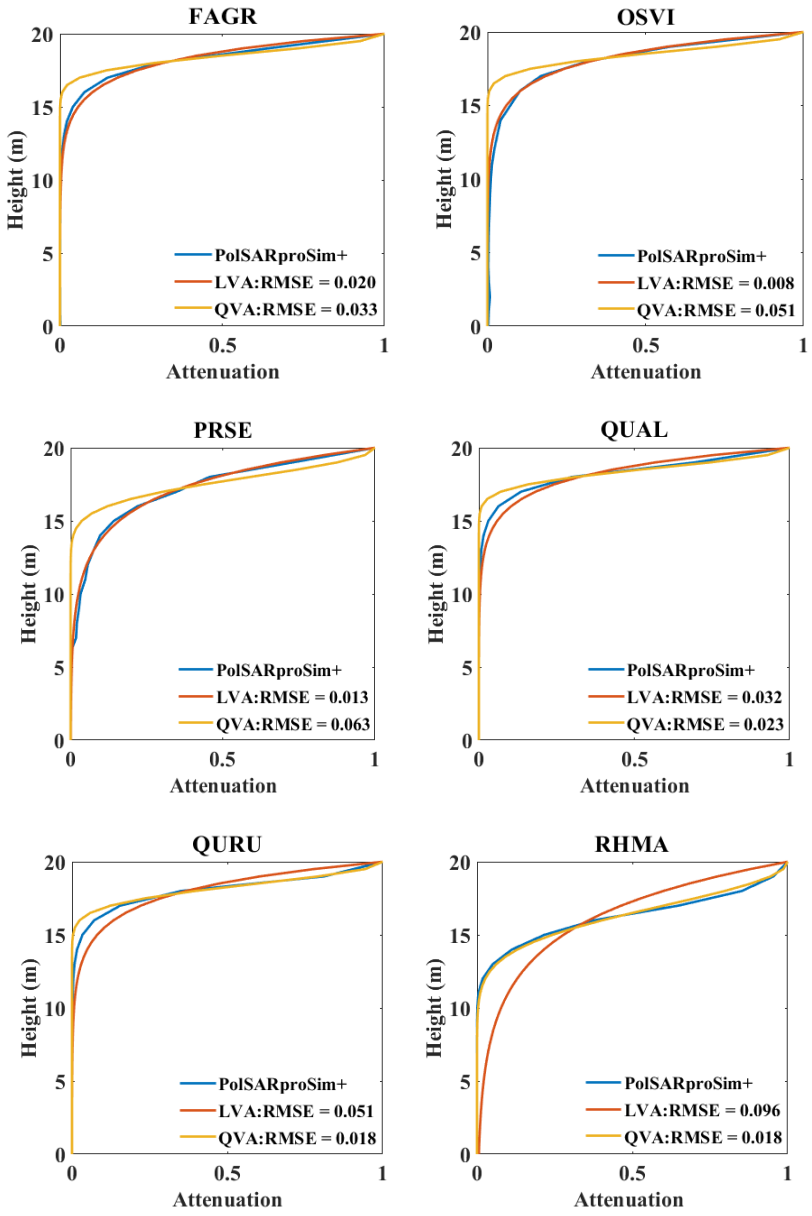


Figure 3.9 Vertical scattering attenuation fitting results over the simulated deciduous forest stands.

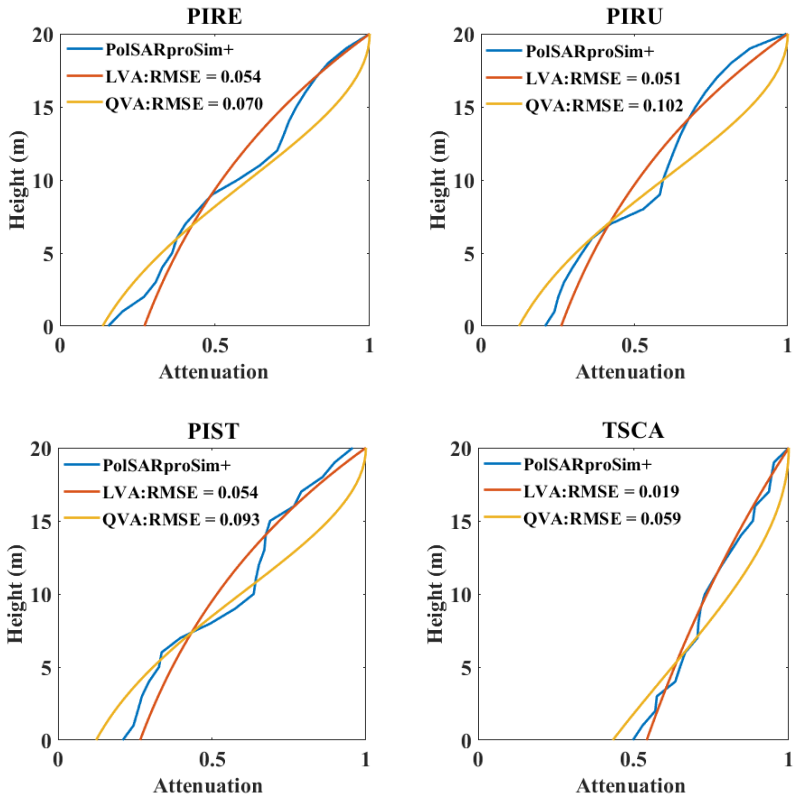


Figure 3.10 Vertical scattering attenuation fitting results over the simulated coniferous forest stands.

The attenuation description with a smaller RMSE is chosen as the best-fit attenuation model for the forest stands of a certain species as listed in Table 3.2, which indicates that LVA is more suitable to depict the coniferous species while a mixture of LVA and QVA is needed for different deciduous species.

Table 3.2 Best fit attenuation description for the forest stands of each species

Species	Best-fit model	Species	Best-fit model	Species	Best-fit model
ACPE[1]	QVA	ACRU[1]	LVA	ACSA[1]	QVA
BEAL[1]	LVA	BEEX[1]	LVA	BELE[1]	LVA
BEPA[1]	LVA	BEPO[1]	LVA	FAGR[1]	LVA
OSVI[1]	LVA	PRSE[1]	LVA	QUAL[1]	QVA
QURU[1]	QVA	RHMA[1]	QVA	PIRE[2]	LVA
PIRU[2]	LVA	PIST[2]	LVA	TSCA[2]	LVA

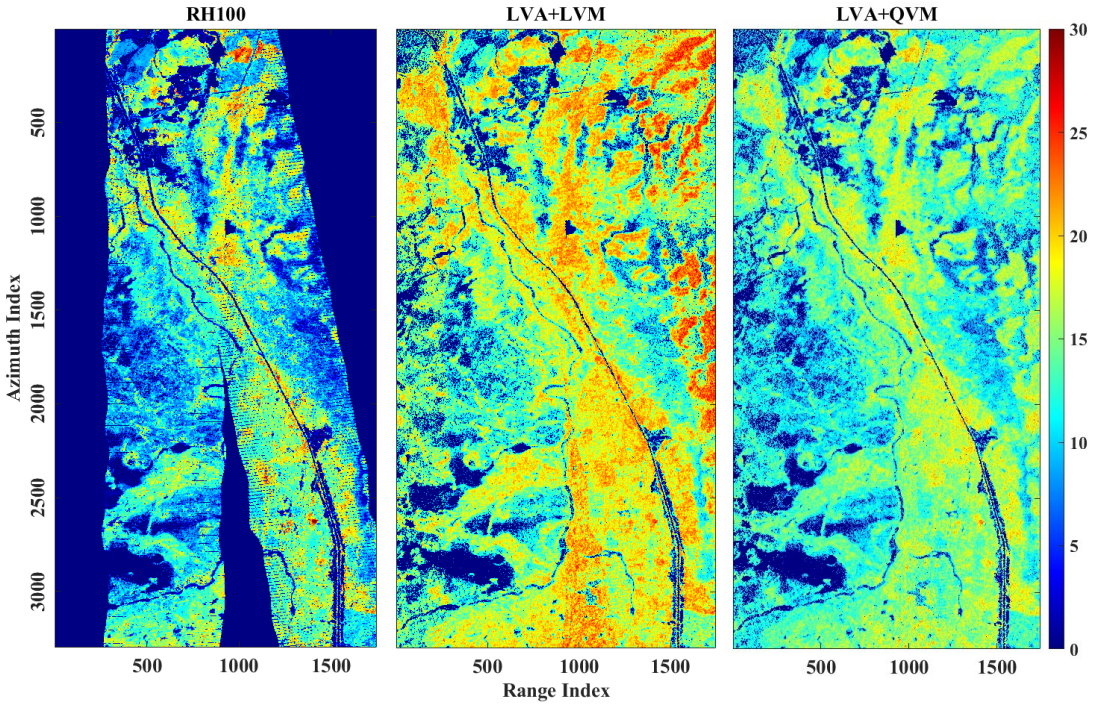
[1] Deciduous species; [2] Coniferous species.

3.4.2 Forest height inversion

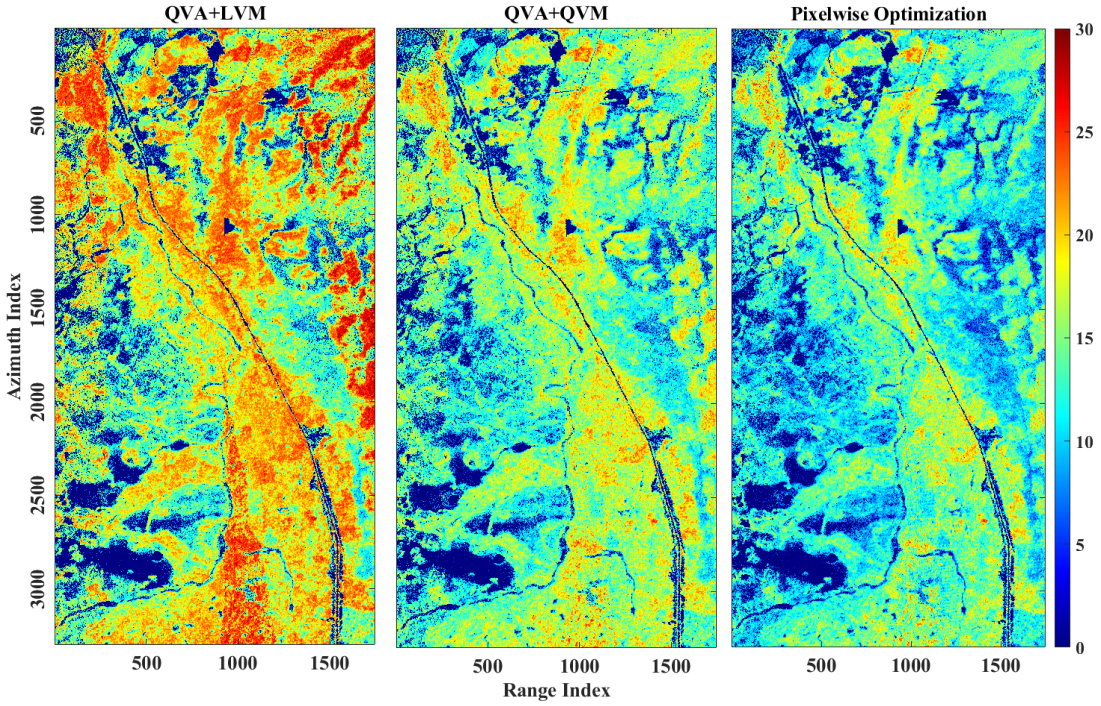
3.4.2.1 BERMS

Dual-baseline PolInSAR data are selected from multi-baseline UAVSAR collections over BERMS collected in uniformly distributed vertical and temporal baselines (see Section 3.2.2.1). Vertical baselines between the primary and two auxiliary images are 40 m and temporal baselines are 25 min with an average absolute vertical wavenumber of 0.09 and the π/k_z of 35 m. Since the maximum forest height in BERMS is about 33 m, the selection of longer vertical baselines will result in a smaller ambiguity height than the canopy height and hence an invalid assumption during the ground solution selection. Forest height results derived from the proposed dual-baseline method using physical models with various attenuation and motion properties and the Cloude's dual-baseline method using the RVoG model are illustrated in Figure 3.11 along with the RH100 reference height, where the areas with low backscatter (less than -15 dB for an incidence angle from 0-35°, less than -21 dB for an incidence angle from 35-45°, less than -24 dB for an incidence angle from 45-55°, and less than -28 dB for an incidence

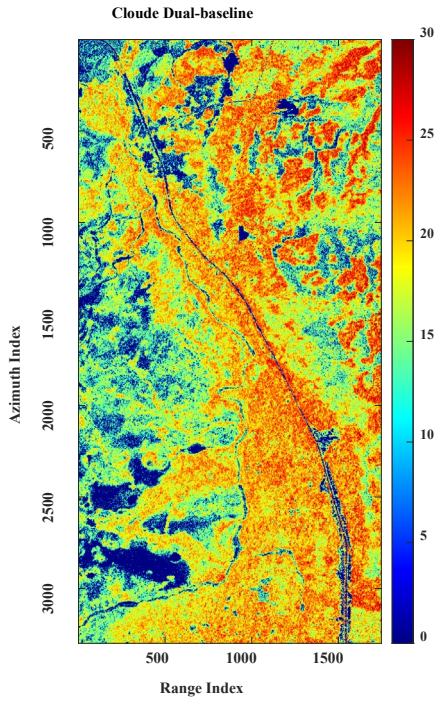
angle larger than 55°) are masked out (Denbina and Simard, 2017) and a 3×3 moving average is performed to smooth forest height maps. Related PolInSAR-derived heights are plotted versus LVIS RH100 height for a quantitative comparison as presented in Figure 3.12 and the associated RMSE, BIAS, and R^2 metrics are given in Table 3.3, where the calculated RMSE and BIAS are precise to the centimeter, and R^2 is precise to two decimal places as in the prior studies (Denbina et al., 2018, Simard and Denbina, 2018). During the comparison, height samples are uniformly selected from the whole forest height image with a separation of 10 pixels in both the slant range and azimuth directions, and pixels corresponding to null height values in the RH100 map are not included in the calculation of any accuracy metrics.



(a) (b) (c)



(d) (e) (f)



(g)

Figure 3.11 Forest height inversion results (in meters) over BERMS based on different models. (a) LVIS RH100 height; (b) LVA+LVM; (c) LVA+QVM; (d)

QVA+LVM; (e) QVA+QVM; (f) Pixel-wise optimization; (g) Cloude's dual-baseline

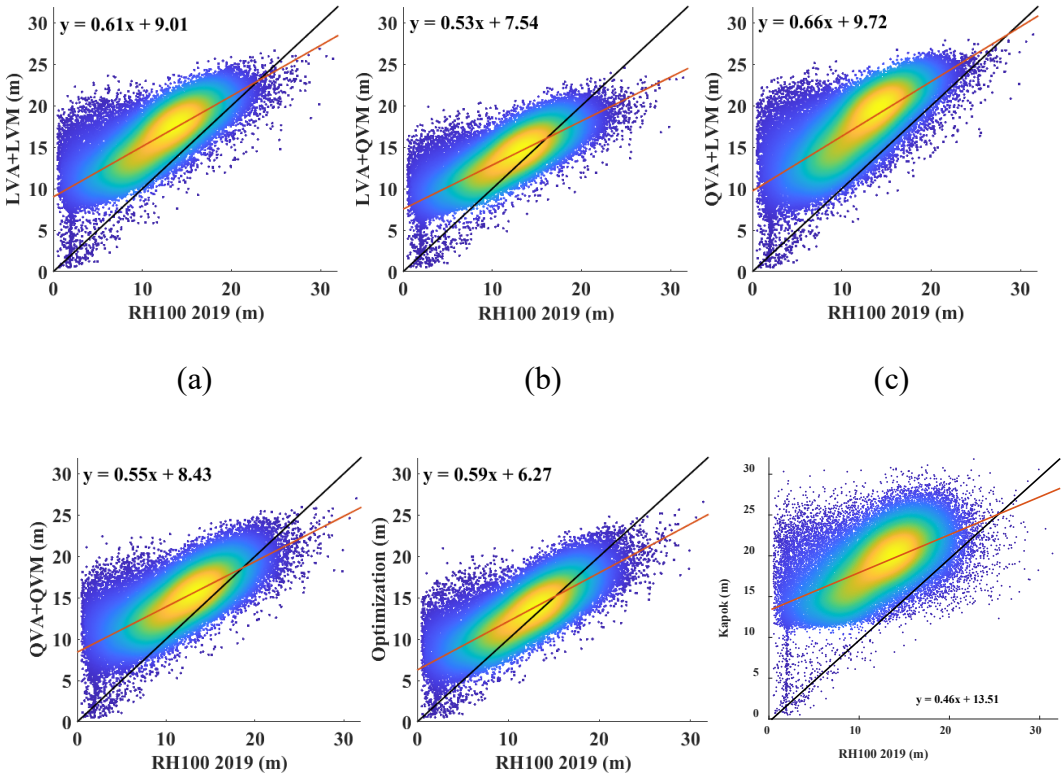


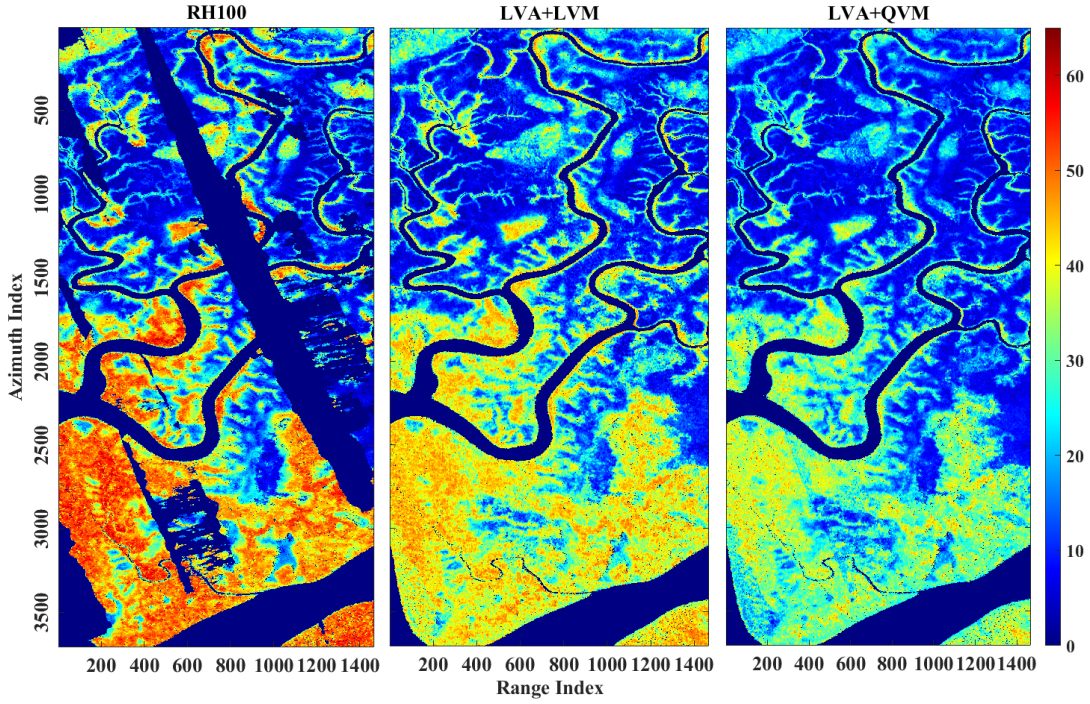
Figure 3.12 Density plots of forest height inversion results versus the LVIS RH100 metric over the BERMS site. Black lines are y=x and red lines are the linear fits. (a) LVA+LVM; (b) LVA+QVM; (c) QVA+LVM; (d) QVA+QVM; (e) Pixel-wise optimization; (f) Cloude's dual-baseline

Table 3.3 Accuracy metrics of forest height inversion results over BERMS.

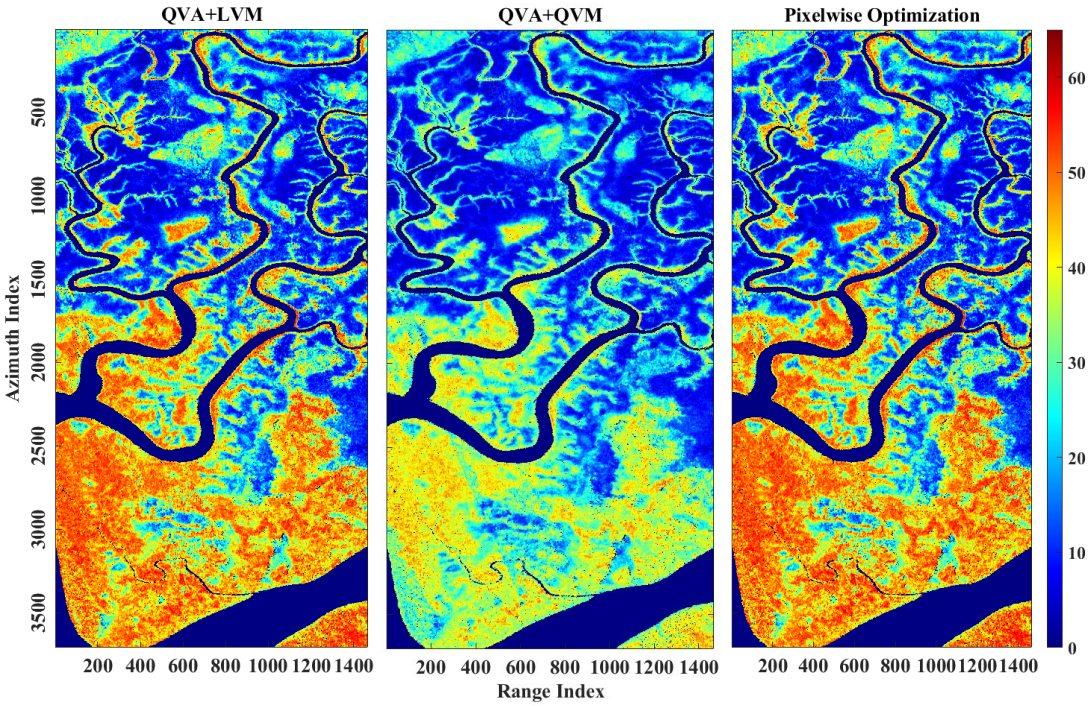
	LVA+LVM	**LVA+QVM**	**QVA+LVM**	**QVA+QVM**	**Opt**	**Cloude**
RMSE	5.31 m	3.56 m	6.56 m	4.43 m	3.21 m	8.32 m
Bias	4.43 m	2.05 m	5.74 m	3.13 m	1.45 m	7.24 m
R^2	0.63	0.65	0.58	0.57	0.65	0.33

3.4.2.2 PNP

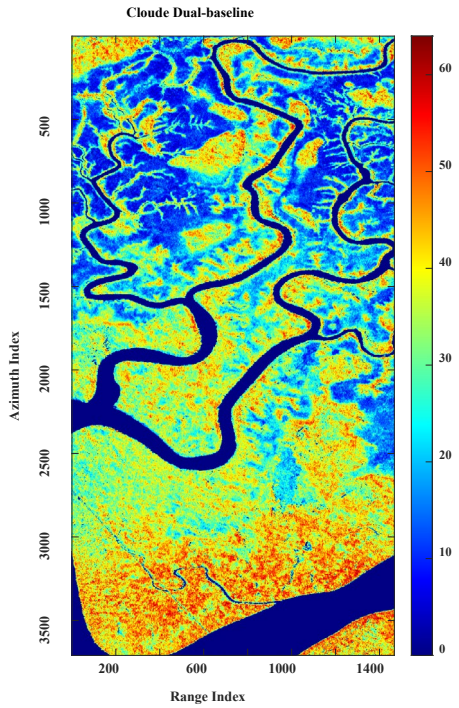
Dual-baseline PolInSAR data are also selected from the multi-baseline collection over the PNP site as depicted previously in Section 3.2.2.1. The maximum forest height in PNP is much higher than the BERMS (~65 m) and thus requests for a smaller vertical baseline with a larger ambiguity height. Vertical baselines between the primary and the two auxiliary images are both 20 m with temporal baselines of 25 min and 50 min, respectively. Related average absolute vertical wavenumber is about 0.05 with π/k_z of ~63 m. Forest height results in the PNP site derived from the proposed dual-baseline method using physical models with various attenuation and motion properties and the Cloude's dual-baseline method using the RVoG model are presented in Figure 3.13. Similar masking and smoothing operations are conducted on the estimated height for PNP as was done for BERM. Density plots between these estimated forest heights and the LVIS RH100 height are presented in Figure 3.14 along with the associated RMSE, BIAS, and R^2 metrics with the same decimal places given in Table 3.4.


RH100
LVA+LVM
LVA+QVM
Azimuth Index
Range Index
200 400 600 800 1000 1200 1400
500 1000 1500 2000 2500 3000 3500
0 10 20 30 40 50 60

(a) (b) (c)

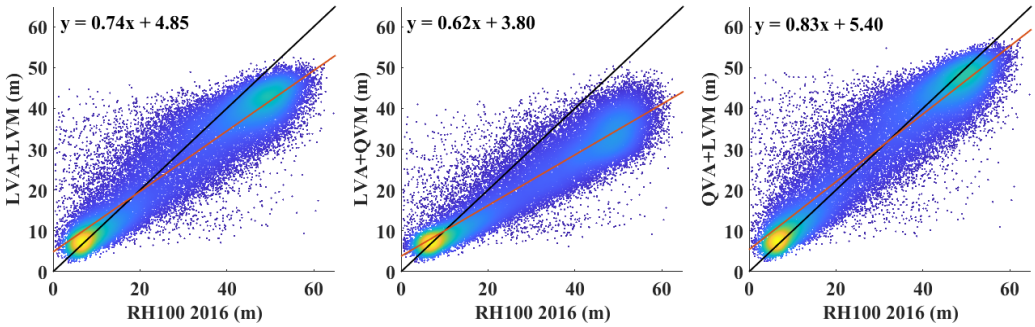

QVA+LVM
QVA+QVM
Pixelwise Optimization
Azimuth Index
Range Index
200 400 600 800 1000 1200 1400
500 1000 1500 2000 2500 3000 3500
0 10 20 30 40 50 60

(d) (e) (f)

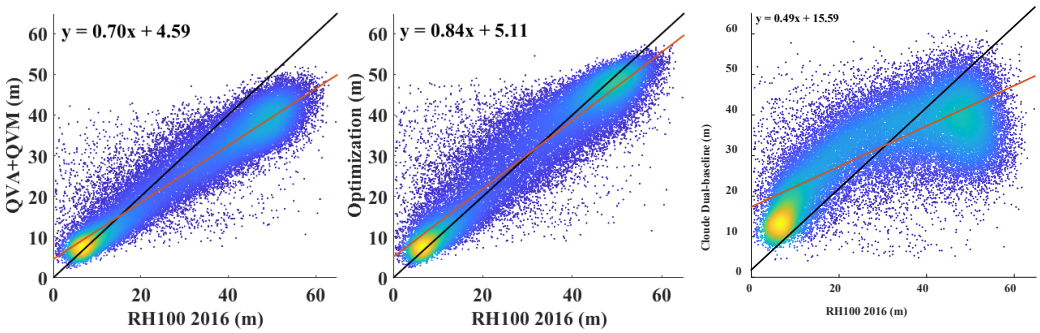


(g)

Figure 3.13 Forest height inversion results (in meters) over the PNP site based on different physical models. (a) LVIS RH100 height; (b) LVA+LVM; (c) LVA+QVM; (d) QVA+LVM; (e) QVA+QVM; (f) Pixel-wise optimization; (g) Cloude's dual-baseline



(a) (b) (c)



(d) (e) (f)

Figure 3.14 Density plots of forest height inversion results versus the LVIS RH100 metric over PNP. Black lines are y=x and red lines are the linear fits. (a) LVA+LVM; (b) LVA+QVM; (c) QVA+LVM; (d) QVA+QVM; (e) Pixel-wise optimization; (f) Cloude's dual-baseline

Table 3.4 Accuracy metrics of forest height inversion results over PNP

	LVA+LVM	LVA+QVM	QVA+LVM	QVA+QVM	Opt	Cloude
RMSE	7.71 m	10.83 m	6.83 m	8.09 m	6.48 m	11.70 m
Bias	-2.74 m	-7.29 m	0.43 m	-4.25 m	0.41 m	0.91 m
R^2	0.83	0.91	0.92	0.93	0.92	0.76

3.5 Discussion

3.5.1 BERMS

Compared with RH100 metrics, forest heights derived from the proposed and the prior Cloude's dual-baseline method over BERMS are generally overestimated as presented in Figure 3.11 and Figure 3.12. The overestimation is more prominent in smaller trees than the taller ones. The proposed physical model with QVM significantly reduces the overestimation and exhibits a better performance in either the LVA or QVA conditions

as illustrated in Figure 3.12(b) and Figure 3.12(d). A 1.75 m reduction in the RMSE and a 2.38 m reduction in the bias are observed in the LVA+QVM result and a 2.13 m reduction in the RMSE and a 2.61 m reduction in bias are achieved by the QVA+QVM result. This suggests that Gaussian distribution with zero mean and vertically quadratic variance better fits the dynamic motion properties of forests in BERMS. Moreover, the proposed physical model with LVA surpasses QVA for forest height inversion in either LVM or QVM conditions because the heterogeneous attenuation aggravates the overestimation of smaller trees as indicated in Figure 3.12(c) and Figure 3.12(d). A 1.25 m reduction in the RMSE and a 1.31 m reduction in the bias are achieved by the LVA+LVM result and a 0.87 m reduction in RMSE and a 1.08 m reduction in bias are observed in the LVA+QVM result. This is consistent with the scattering attenuation fitting results derived from the PolSARproSim+ simulated data (Section 3.4.1), where the physical model with LVA is more suitable for coniferous species such as the jack pine and black spruce in BERMS. Advantages of different scattering attenuation and random motion structures can be observed through using the pixel-wise optimization strategy, which leverages LVA, LVM, QVA, and QVM on a pixel basis and therefore achieves better performance than the others that don't vary spatially. As presented in

Figure 3.11(f) and Figure 3.12(e), a 0.35 m reduction in RMSE and a 0.60 m reduction in bias are observed compared with the best-performing single model (LVA+QVM). Besides that, as indicated in the high RMSE and bias and the low R^2 in the last column of Table **3.3**, the overestimation is much more prominent in the result originating from the prior Cloude's dual-baseline method based on the RVoG model compared with the

results derived from the proposed dual-baseline method. We relate this discrepancy to the neglect of temporal decorrelation in the RVoG model.

3.5.2 PNP

Compared with RH100 metrics, forest heights derived from the proposed model over the PNP site are generally underestimated as presented in Figure 3.13 and Figure 3.14, particularly in the areas with taller trees. The LVM description in the proposed model significantly reduces the underestimation and exhibits a better forest height estimation performance in either LVA or QVA conditions as shown in Figure 3.14(a) and Figure 3.14(c). A 3.12 m reduction in the RMSE and a 4.55 m increase in the bias are achieved by the LVA+LVM result, whereas a 1.26 m reduction in RMSE and a 4.68 m increase in bias are observed in the QVA+LVM result indicating that a Gaussian distribution with zero mean and vertically linear variance better fits the dynamic motion properties of tropical forests in PNP. Moreover, the physical model with QVA better performs on the forest height inversion than the LVA ones in either LVM or QVM conditions by reducing the underestimation of taller trees as shown in Figure 3.14(c) and Figure 3.14(d). A 0.88 m reduction in RMSE and a 3.17 m increase in the bias are obtained by the QVA+LVM result and a 2.74 m reduction in the RMSE and a 3.04 m increase in the bias are achieved by the QVA+QVM result. This is consistent with the scattering attenuation fitting results over the simulated red mangrove (RHMA) forest stands (see section 3.4.1), where physical models with QVA appear to be more suitable to describe the scattering attenuation of RHMA species in PNP. Figure 3.13(f) and Figure 3.14(e) also provide evidence to support the superiority of the pixel-wise optimization, where a 0.35 m reduction in the RMSE and a 0.02 m reduction in bias are achieved compared

with the best-performing single model (QVA+LVM). Moreover, as indicated in Figure 3.14(f), though neglecting the temporal decorrelation in the RVoG model reduces the underestimation in the associated forest height result, the RMSE is still very high due to the strong discrepancy between the scattering attenuation properties of the RVoG model and the real scenario.

3.6 Conclusion

In this chapter, we demonstrated the potential of dual-baseline repeat-pass airborne PolInSAR for estimating forest height at the landscape scale. A physical model with LVA, QVA, LVM, and QVM profiles in the volume layer is proposed based on the RMoG model to investigate the impacts of homogeneous or heterogeneous attenuation and random motion properties on the performance of forest height inversion. Experiments are carried out over a boreal and tropical forest site to show the efficacy of different volumetric profiles over various forest types. Related forest height results are compared with LVIS RH100 references for the quantitative evaluation. Volumes with LVA and QVM achieve the best performance (RMSE of 3.56 m and bias of 2.05 m) over the boreal forest dominated by coniferous tree species, while volumes with QVA and LVM exhibit the best performance (RMSE of 6.83 m and bias of 0.43 m) in the tropical forests dominated by deciduous red mangroves. This is consistent with the scattering attenuation fitting results derived from the PolSARproSim+ simulated data. Forest height results generated by the pixel-wise optimization strategy surpass the best-performing single models in both the boreal (RMSE of 3.21 m and bias of 1.45 m) and the tropical (RMSE of 6.48 m, bias of 0.41 m) forests, which indicates its superiority in leveraging the advantages of different volumetric profiles.

Chapter 4 Pol-GAN: A deep-learning-based unsupervised forest height estimation method based on the synergy of PolInSAR and LiDAR data

The previous chapter explores the advance and potential of model-based forest height inversion using PolInSAR data, which retrieved the forest height by solving nonlinear equations established on the relationship between the forest parameters and coherence observations. These methods estimate forest height at pixel level and exhibit a strong dependency on the theoretical models. Hence, the discrepancy between the theoretical models and actual backscattering behaviors and the neglect of useful information like shape, texture, and spatial characteristics among pixels inevitably lead to errors in the model-based forest height estimates. Hence, LiDAR acquisitions with higher vertical accuracy are introduced into the forest height estimation through the synergy with the PolInSAR data. A deep-learning-based unsupervised forest height estimation method is proposed for data fusion. This method is based on a designed GAN architecture, which reformulates the forest height inversion as a pan-sharpening process between the sparse LiDAR-derived height and the high-spatial-resolution PolInSAR features to address the challenges in traditional PolInSAR-based forest height inversion. The forest height estimates with high spatial resolution and vertical accuracy are generated through a continuous generative and adversarial process, where a progressive network architecture underpins the generator to overcome the significant difference between spatial resolutions of spaceborne LiDAR and SAR acquisitions. Finally, experiments

are performed on a boreal and a tropical forest site to test the proposed forest height estimation algorithm.

This work is accepted and in press in ISPRS Journal of Photogrammetry and Remote Sensing.

Zhang, Q., Ge, L., Hensley, S., et al. PolGAN: A deep-learning-based unsupervised forest height estimation based on the synergy of PolInSAR and LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing* 2022; 186, 123-139.

4.1 Introduction

One aspect of forest characterization encompasses describing the spatial and vertical distributions of trees, which plays a significant role in carbon sequestration estimation, sustainable forest management, and climate change monitoring (Hardiman et al., 2011, Wulder et al., 2012, Berndes et al., 2016, Kiat et al., 2020). As a critical part of the forest characterization, forest height estimation aims to provide a key indicator of the forest productivity and biomass level, statically or/and dynamically (Englhart et al., 2011). Prior studies have been dedicated to measuring the forest height through in-situ or/and remote sensing methods (Schneider et al., 2014, Pause et al., 2016). In-situ methods employ plenty of expensive and time-consuming field works (Liang et al., 2018), whereas spaceborne and airborne remote sensors can provide consistent and cost-efficient observations at regional or/and global scales. In particular, active remote sensors such as SAR and LiDAR are superior to passive ones thanks to their capability of providing data for inferences on vertical distribution and internal structure of the vegetation (Aghababaei et al., 2020, Guo et al., 2020a).

Advantages of SAR remote sensing such as cloud penetration, weather independence, wide coverage, and high spatial resolution are widely documented (Moreira et al., 2013). However, in SAR images, 3D radar reflectivities are projected into an azimuth-range domain where height information needs to be further retrieved through multi-channel observations. For this purpose, Polarimetric SAR Interferometry (PolInSAR) (Papathanassiou and Cloude, 2001) emerged as an effective approach for estimating forest height due to its sensitivity to both the physical properties and vertical structure of the forest scattering medium. Several forest height inversion algorithms have been developed based on PolInSAR data, most of them are based on the theoretical models. The most commonly used Random Volume over Ground (RVoG) model (Cloude and Papathanassiou, 2003) simplifies the forest as a homogeneous volume of randomly oriented scattering particles statically distributed over the ground, where volumetric attenuation of the radar scattering amplitude follows an exponential distribution with a constant extinction coefficient. Other complementary models have also been devised based on RVoG model by integrating vertical heterogeneity into forest representation (Garestier and Le Toan, 2007) or/and compensating the temporal decorrelation caused by the dynamic change of vegetation and ground surface properties occurring between acquisitions (Papathanassiou and Cloude, 2003, Lavalle and Hensley, 2015, Lei et al., 2017), which have been further applied to various forest types and sensor wavelengths for single-pass or repeat-pass interferometry (Praks et al., 2007, Garestier et al., 2008a, Garestier et al., 2008b, Hajnsek et al., 2009, Hensley et al., 2012, Kugler et al., 2015). Qualities of these forest height results normally exhibit a strong dependency on the theoretical model employed because they are estimated through solving non-linear

equations approximately established between the coherence observations and forest parameters. Besides that, other factors related to PolInSAR system (vertical baseline, signal to noise ratio, etc.), the data collection method (single or repeat-pass, single or multiple-baseline, etc.), and imaging scenario (topography, vegetation condition, etc.) may also influence the vertical accuracy of forest height inversion (Ahmed et al., 2011, Kugler et al., 2015, Zhang et al., 2016b).

LiDAR measurements are based on the time elapsed from the round-trip travel of each laser pulse between the sensor and target (Sexton et al., 2009), which are less affected by PolInSAR-related errors such as the speckle noise and temporal decorrelation and normally exhibit a higher vertical accuracy. However, LiDAR data are economically and computationally constrained at a local scale because of the small swath. The most commonly used LiDAR systems in forest height estimation are the discrete return and full-waveform LiDAR (Silva et al., 2018). Discrete systems record single or multiple returned peaks from a given pulse whereas full-waveform systems digitize the entire reflected energy to provide complete vertical profiles (Sumnall et al., 2016). Studies confirmed that full-waveform systems are less affected by signal occlusion and are better at describing the details in the forest canopy and understory vegetation (Crespo-Peremarch et al., 2020), which are equipped onboard spaceborne platforms to produce consistent but coarse-resolution forest structure maps at the continental or global scale (Qi and Dubayah, 2016, Martino et al., 2019). Full-waveform LiDARs operating from orbit measure the amount of reflected laser light from structures distributed vertically throughout the canopy volume and determine the forest canopy height and structure from metrics related to the vertical distribution of the reflected laser light, such as the

hundredth percentile of the cumulative waveform energy relative to the ground surface (RH100) (Anderson et al., 2006, Silva et al., 2018).

To produce wide-coverage forest height maps with high spatial resolution and vertical accuracy, related studies have attempted to generate forest height estimates through the synergy of LiDAR and PolInSAR. Support vector machine (SVM) is employed (Pourshamsi et al., 2018) to extrapolate the LiDAR-based canopy height using multi-baseline PolInSAR inverted parameters. SVM is also employed for the LiDAR-aided baseline selection (Denbina et al., 2018) to determine the best pair from multi-baseline PolInSAR data for the forest height inversion over each pixel. Besides, some studies improved the performance of PolInSAR-based forest height inversion by extracting prior information (ground phase, mean extinction, etc.) from the LiDAR metrics (Fu et al., 2017, Qi et al., 2019). However, these methods still have a strong dependency on the simplified model used, which may lead to errors in the estimated vegetation parameters when the model largely departures from the real scene. Therefore, instead of inversion, this study reformulates the multi-source (LiDAR and PolInSAR) forest height estimation as an image pan-sharpening task aiming at the generation of high spatial resolution and vertical accuracy forest height through the synergy of LiDAR and PolInSAR data. The recent advances in deep learning and its success in computer vision and information fusion inspired us to explore its potential on this specific forest height estimation task.

Pan-sharpening can be regarded as a super-resolution process guided by high-spatial-resolution panchromatic images. Prior studies introduced the Convolutional Neural

Networks (CNN) (LeCun and Bengio, 1995) and Generative Adversarial Networks (GAN) (Goodfellow et al., 2014) into pan-sharpening tasks in terms of supervised and unsupervised learning. Following Wald's protocol, supervised methods such as PNN (Masi et al., 2016), PanNet (Yang et al., 2017), PSGAN (Liu et al., 2020a), and RED-cGAN (Shao et al., 2019a) use the spatial-degraded images as low-spatial-resolution inputs and the original images as ground truth for training. These methods either suffer from spectral or/and spatial distortions or require additional supervision. To address these issues, unsupervised methods such as Pan-GAN (Ma et al., 2020) and PGMAN (Zhou et al., 2021) directly use the original images as inputs and preserve the spectrum and spatial information by two separate discriminators throughout the generative and adversarial game.

Since GAN has achieved remarkable performance on image pan-sharpening or super-resolution, the reformulated forest height pan-sharpening task in this paper is realized by a tailored GAN method called PolGAN. Instead of parameter inversion, PolGAN is inversely a forest height generation process, where the generator produces a high spatial resolution and vertical accuracy height map based on the PolInSAR and LiDAR inputs and the feedback provided by the discriminators. The mathematical relations established between forest parameters and PolInSAR coherences in traditional model-based methods are employed in one of the discriminators as the criteria to assess the quality of forest height generation. The designed structure of PolGAN not only reduces its dependence on the theoretical model but also mitigates the limitation of the spatial baseline during PolInSAR data acquisition because interferometric phase unwrapping reverses into a wrapping process, where the vertical wavenumber on longer needs to

fit with the various vegetation height such that the highest observed phase center can match the requirement of less than π (Kugler et al., 2015, Simard and Denbina, 2018). Moreover, as an unsupervised method, PolGAN fits an individual set of weights for each new image and the training process does not rely on the ground truth by designing specific loss functions.

The rest of this chapter is organized as follows. Test sites and datasets used in the experiments are presented in Section 4.2. Section 4.3 describes details of the proposed PolGAN method. Forest height estimation results over the boreal and tropical forest sites are presented in Section 4.4 following by some discussions. Finally, conclusions are drawn in Section 4.5.

4.2 Test sites and datasets

4.2.1 Test sites

A boreal and a tropical forest site are used to investigate the application of PolGAN to forest height estimation. As presented in Figure 4.1, the boreal forest site is part of the Howland research forest northeast of Bangor in the U.S. state of Maine. The whole imaging area covers about 24,018 ha from 44°48'41'' N to 45°10'18'' N and from 68°34'22'' W to 68°45'48'' W. As a typical northern hardwood transition forest, the dominant tree species at this site are spruce-hemlock-fir, aspen-birch, and hemlock-hardwood mixtures with an average height of around 20 m (Lei and Siqueira, 2014). The topography varies from relatively flat to gently rolling with a maximum elevation change of fewer than 70 m within 10 km. The climate is primarily cold and humid in

the Howland site with an annual temperature of 5.4±5.0 °C and an annual rainfall of 1050±170 mm. The MODIS Land Cover Type Product (MCD12Q1) in 2009 provides a rough statistic of the landcover types and components in the selected boreal forest (evergreen needle-leaf forests of 0.9%, deciduous broadleaf forests of 6.4%, mixed forests of 67.5%, savannas of 18.3%, grasslands of 0.6%, built-up lands of 4.0%, and water bodies of 2.3%).

As shown in Figure 4.2, the tropical forest site is part of the Lope National Park located in central Gabon (0°2'12'' S to 0°19'34'' S, 11°27'31'' E to 11°43'29'' E) covering a total area of around 33,062 ha. The north of the Lope site is mainly characterized by savannas while the south part is dominated by tropical forests providing a gradually varying canopy height and vertical structure through the whole imaging coverage (El Moussawi et al., 2019). Significant topographic variations are observed around this site with altitudes ranging from 0 m to 683 m and the largest slope greater than 20%. The annual rainfall is about 1500 mm in the northern part and 2500 mm in the southern (Mitchard et al., 2012). Likewise, the 2016 MCD12Q1 product gives a rough statistic of the landcover types and components in the tropical forest site (evergreen broadleaf forests of 52.1%, savannas of 42.2%, grasslands of 5.1%, wetlands of 0.3%, and water bodies of 0.1%).

4.2.2 Datasets

4.2.2.1 UAVSAR

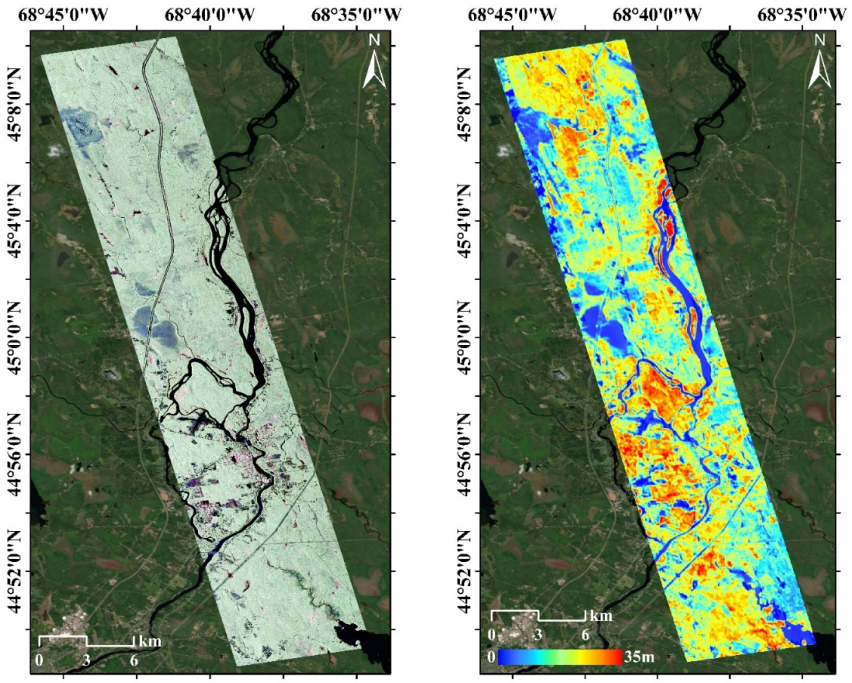
The SAR data involved in this study were collected by UAVSAR, which can be further used as a proxy for the NISAR's L-band product (Lavalle et al., 2017, Chapman et al.,

2019). UAVSAR data over the Howland forest site were collected in a zero-spatial-baseline configuration with a repeat-pass time interval of 42 min on August 5, 2009. The UAVSAR data over the Lope site are collected on February 25, 2016, during the AfriSAR campaign (Fatoyinbo et al., 2021), with a vertical baseline of 20 m and a temporal baseline of 23 min. A typical UAVSAR SLC product has a ground range and an azimuth resolution of 2.5 m and 1.0 m (Fore et al., 2015), respectively, and can be further averaged by a rectangular window to reduce speckle noise and generate the multi-looked polarimetric and interferometric images. The geocoded and multi-looked SAR images in the Pauli color combination over the Howland and Lope forest sites are presented in Figure 4.1(a) and Figure 4.2(a).

4.2.2.2 Land Vegetation and Ice Sensor (LVIS)

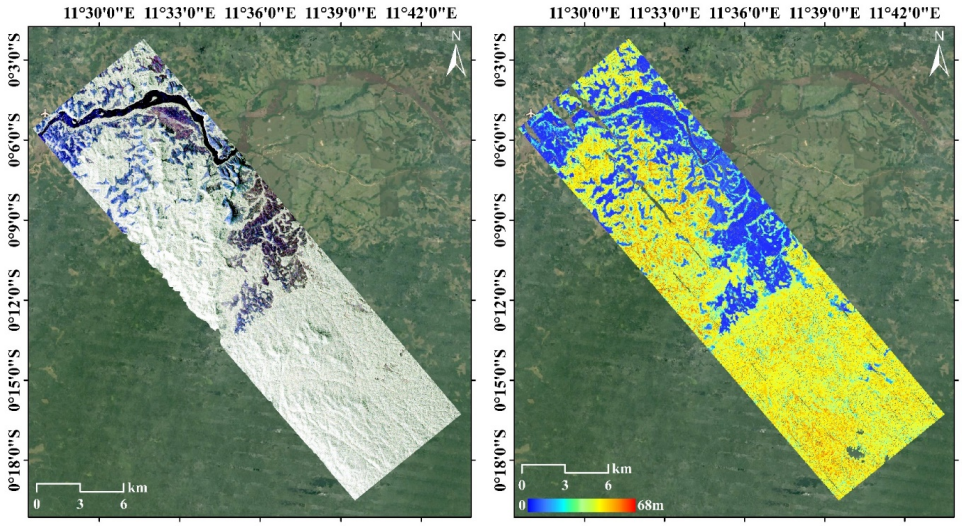
The large-footprint full-waveform LiDAR data used in this chapter are RH100 metrics originating from the LVIS level-2 collection. LVIS products are can be used as a proxy for the sparse orbital LiDAR measurements from systems such as GEDI (Dubayah et al., 2020). The nominal footprint diameter of LVIS is 20 m to 25 m separated by 9 m (overlapping footprints) in the along-track direction while GEDI samples the Earth's surface with 25 m nominal footprints distributed at 60 m spacing along-track and 600 m spacing cross-track (Dubayah et al., 2020). Therefore, RH100 metrics derived from LVIS data are downsampled to simulate GEDI-like data. LVIS data over the Howland site were captured on 6 August 2009 with a 20 m footprint diameter, an 8.5 km flight altitude, and a 1.7 km swath width. LVIS data over the Lope site were collected on 02 March 2016 with an 18 m footprint diameter, a 7.3 km flight altitude, and a 1.5 km

swath width. LVIS RH100 metrics in the geographic projection over the Howland and Lope sites are illustrated in Figure 4.1(b) and Figure 4.2(b).



(a) (b)

Figure 4.1 Howland site: a) UAVSAR Pauli color composite image. Red: $|\boldsymbol{S_{HH}} - \boldsymbol{S_{VV}}|^2$, Green: $\mathbf{2}|\boldsymbol{S_{HV}}|^2$, Blue: $|\boldsymbol{S_{HH}} + \boldsymbol{S_{VV}}|^2$. b) LVIS RH100 height.

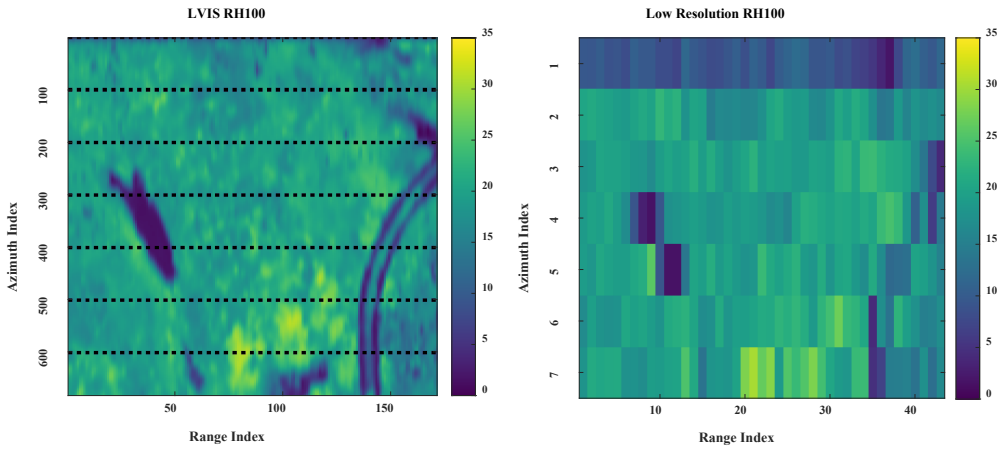


(a) (b)

Figure 4.2 Lope site: a) UAVSAR Pauli color composite image. b) LVIS RH100 height.

4.2.2.3 Data preparation

Downloaded UAVSAR SLC images are first multi-looked by a rectangular window with 6 pixels in both azimuth and range directions to reduce speckle noise and generate multi-looked polarimetric images with azimuth and ground range sampling spacings of 6 and 15 m. These multi-looked images are subsequently input into the polarimetric decomposition and the polarimetric interferometry modules in the preprocessing part of PolGAN (Section 4.3.2) to produce high-spatial-resolution features. LVIS RH100 metrics are projected into the same slant range and azimuth frame as the multi-looked polarimetric images which are used as the ground truth during performance evaluation. This high-resolution RH100 height is downsampled by slicing with an azimuth factor of 100 and a range scaling factor of 4 (black dots in Figure 4.3(a)) to generate the low-resolution LiDAR height inputs with azimuth and ground range sampling spacings of 600 m and 60 m (same as GEDI).



(a) (b)

Figure 4.3 (a) Example of the projected LVIS RH100 height. (b) Example of the downsampled low-spatial-resolution height map.

4.3 Methodology

4.3.1 Overview of Pol-GAN

Goodfellow proposed the generative adversarial networks (GAN) to generate realistic images through estimating the underlying distribution of the given unlabeled samples (Goodfellow et al., 2014). The network is designed based on game-theoretic principles with two adversarial parts (a generator and a discriminator). The generator learns the distribution of the given data by mapping a random sample z from any distribution to a sample x from the data space and is trained to output images to fool the discriminator. The discriminator learns to distinguish the difference between the generated and real samples and outputs a scalar indicating the probability of a sample being real or fake (Creswell et al., 2018). This adversary continues until the generator-derived images cannot be further distinguished by the discriminator and the two-player min-max game can be formulated as an optimization problem as

$$\min_G \max_D E_{x\sim p_{\text{data}}(x)}[log\, D(x)] + E_{z\sim p_z(z)}\left[log\left(1 - D\big(G(z)\big)\right)\right] \quad (4.1)$$

where G is the generator, D is the discriminator, x is a sample from the distribution of training data $p_{\text{data}}(x)$, and z is a sample from an arbitrary random distribution $p_z(z)$.

In this chapter, the traditional forest height inversion process is reformulated as an unsupervised pan-sharpening task aimed at generating forest heights with high spatial resolution and vertical accuracy based on the synergy of LiDAR and PolInSAR data.

A preprocessing step is employed to extract the high-spatial-resolution PolInSAR coherences and PolSAR decomposition features (see Section 4.3.2) before feeding into the pan-sharpening network. The preprocessing step is necessary because the repeat-pass acquisitions have both necessary as well as superfluous information not needed for the forest height estimation. The pan-sharpening deep neural network is designed based on the generative adversarial strategy in GAN with one generator and two discriminators. The generator absorbs the spatial distributions of different scattering mechanisms in polarimetric decomposition features $Pol_{HR}(r_h h \times r_w w \times d_s)$, vertical distributions of the PolInSAR coherences COR_{HR} $(r_h h \times r_w w \times d_c)$, and low-spatial-resolution heights derived from LiDAR sparse samples H_{LR} $(h \times w \times 1)$ to produce a pan-sharpened forest height map $\widehat{H}_{HR}$ $(r_h h \times r_w w \times 1)$ with greater spatial details and higher vertical accuracy. Here, h and w are the height and width of the low-resolution LiDAR height map, r_h and r_w are ratios between LiDAR and radar pixel spacings in azimuth and range directions, d_c and d_s are the dimensions of PolInSAR coherences and PolSAR decomposition features.

Training of the generator is dominated by the sum of pixel-wise spatial and vertical losses (see Section 4.3.3.1) along with two discriminators playing a refinement role to force the spatial and vertical distributions in the generated images consistent with that in high-resolution inputs. The structure of PolGAN is illustrated in Figure 4.4.

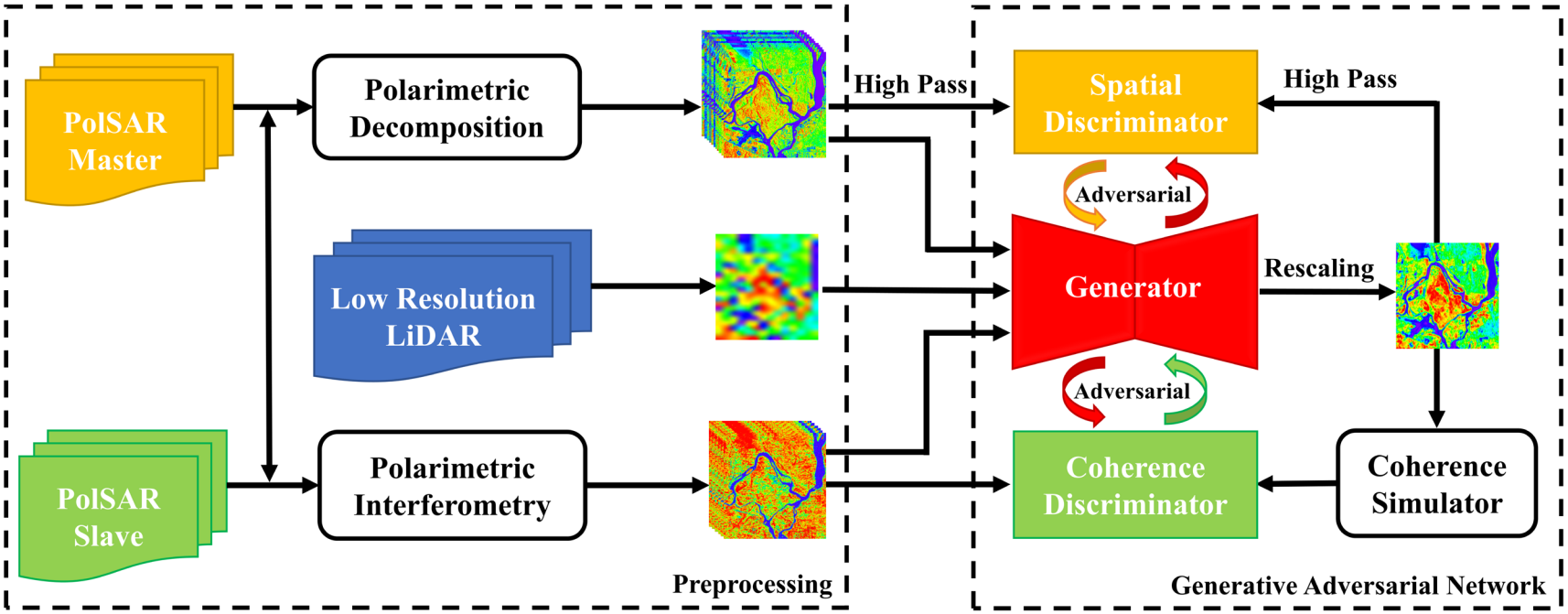


Figure 4.4 Structure of PolGAN with a multi-source data preprocessing module in the left box and the generative adversarial network with one generator and two discriminators in the right box.

4.3.2 Data preprocessing

The repeat-pass PolInSAR acquisition includes a master and a slave fully polarimetric SAR (PolSAR) image collected with a time separation of less than an hour. As listed in Table 4.1, two PolSAR images are fed into a polarimetric decomposition module to extract spatial features of different scattering mechanisms using different polarimetric decomposition algorithms (Cloude and Pottier, 1996, Freeman and Durden, 1998, Lee et al., 1999, Pottier and Lee, 2000, Yamaguchi et al., 2005, Neumann et al., 2009a). Taking the three-component Freeman decomposition as an example, each pixel in a multi-looked PolSAR image can be represented by a coherency matrix observation $\boldsymbol{T}$

$$\boldsymbol{T} = \langle \vec{k}_p \vec{k}_p^{\dagger} \rangle = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{12}^* & T_{22} & T_{23} \\ T_{13}^* & T_{23}^* & T_{33} \end{bmatrix}, with\ \vec{k}_p = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ 2S_{HV} \end{bmatrix}. \quad (4.2)$$

Where S_{HH}, S_{HV} and S_{VV} are the scattering elements in scattering matrix $\boldsymbol{S}$.

In Freeman-Durden decomposition, the coherency matrix $\boldsymbol{T}$ is modeled as a sum of contributions from the surface, double-bounce, and volume scattering mechanisms, given by

$$\boldsymbol{T} = P_s \boldsymbol{T}_s + P_d \boldsymbol{T}_d + P_v \boldsymbol{T}_v, \quad (4.3)$$

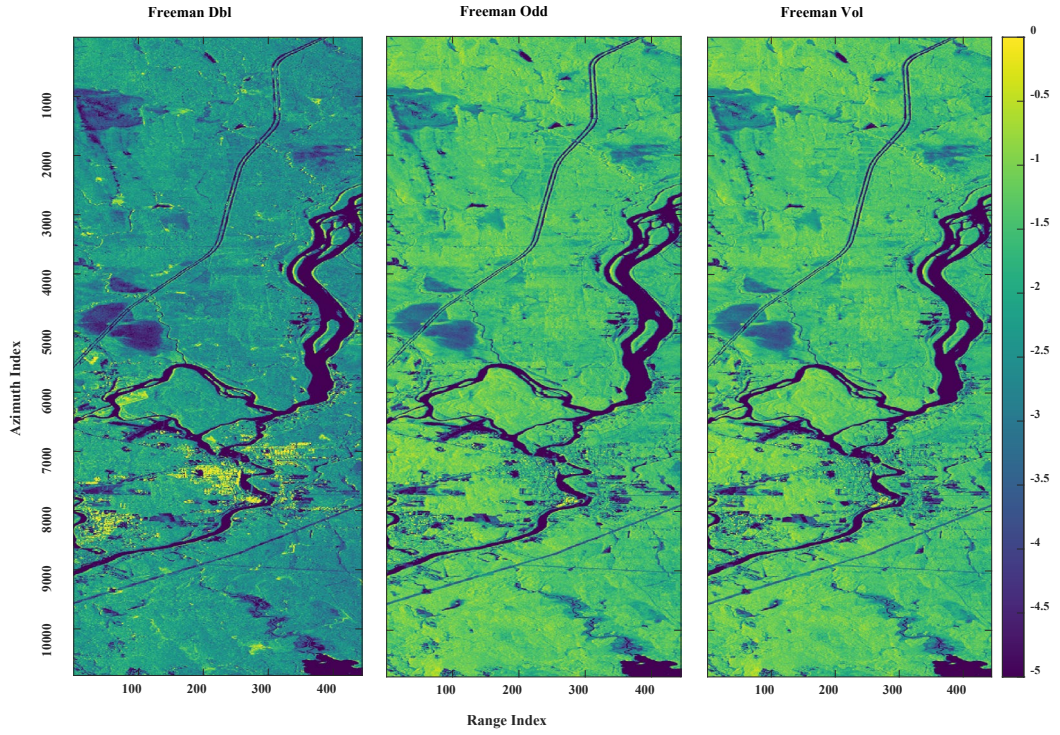
where P_s, P_d, and P_v are powers of the surface, double-bounce, and volume scattering components, $\boldsymbol{T}_s$ is the surface coherency matrix derived from a first-order Bragg surface scatterer, $\boldsymbol{T}_d$ is the double-bounce coherency matrix from a dihedral corner reflector, and $\boldsymbol{T}_v$ is a volume coherency matrix.

$$\boldsymbol{T}_s = \frac{1}{1+|\beta|^2} \begin{bmatrix} 1 & \beta & 0 \\ \beta^* & |\beta|^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \boldsymbol{T}_d = \frac{1}{1+|\alpha|^2} \begin{bmatrix} 1 & \alpha & 0 \\ \alpha^* & |\alpha|^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \boldsymbol{T}_v = \frac{1}{4} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.4)$$

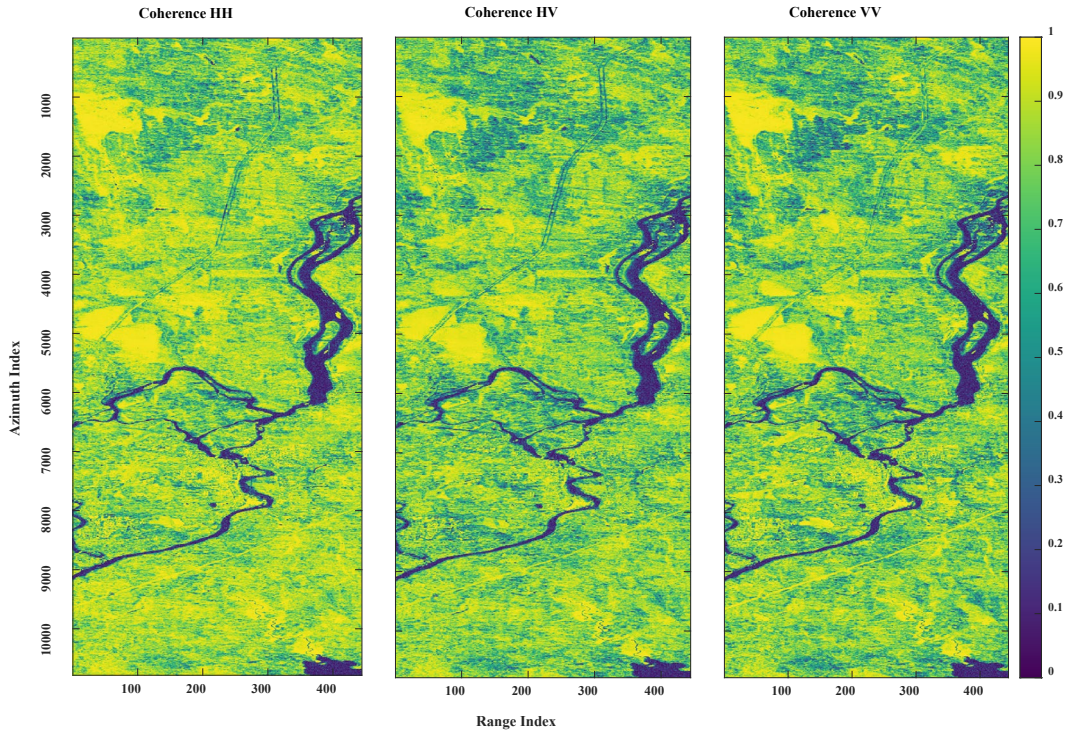
where α and β are scattering parameters with constraints that $|\alpha| < 1$ and $|\beta| < 1$. The surface, double-bounce, and volume powers can be extracted by solving Equation (4.3), which are then fed into the generator and spatial discriminator along with other features (surface, double-bounce, volume, and helix scattering features derived from the Yamaguchi decomposition (Yamaguchi et al., 2005); anisotropy and orientation randomness features derived from Neumann decomposition (Neumann et al., 2009a); anisotropy, entropy, alpha angle, and orientation angle features from the H/A/Alpha decomposition (Cloude and Pottier, 1997); eigendecomposition of coherency matrix (Cloude and Pottier, 1996)). Examples of high-resolution polarimetric decomposition features are presented in Figure 4.5(a).

Table 4.1 Polarimetric decomposition features

Decomposition	**Feature**	**Dimension**
Yamaguchi	Odd, Dbl, Vol, Hlx	4
Neumann	Delta, Psi, Tau	3
H/A/Alpha	Alpha, Anisotropy, Beta, Delta, Entropy, Gamma, Lambda, HA, (1-H)A, H(1-A), (1-H)(1-A)	11
Freeman	Odd, Dbl, Vol	3
Cloude	T11, T22, T33	3



(a)



(b)

Figure 4.5 Examples of the high-resolution features after preprocessing: (a) Polarimetric decomposition features; (b) Polarimetric interferometry features

Additionally, the PolSAR image pair is processed by a polarimetric interferometry module with a 3×3 moving window to generate PolInSAR coherences at any desired polarization state (e.g., HH, HV, VV, HH+VV, HH-VV, etc.) (Lavalle and Hensley, 2015, Ghasemi et al., 2018, Simard and Denbina, 2018), where spectral and signal-to-noise ratio (SNR) decorrelations are calibrated by methods in (Kugler et al., 2015). As illustrated in Figure 4.5(b), PolInSAR coherences carry the vertical information of the scattering volume and can be utilized by the generator and coherence discriminator for the forest height estimation.

4.3.3 Generative and adversarial network

4.3.3.1 Generator

Unlike the traditional pan-sharpening tasks with upsampling scales of 4 or 8, scaling factors in the forest height generation process are much larger due to the significant difference between the spatial resolutions of spaceborne LiDAR and SAR acquisitions. As indicated in Section 4.2.2.3, high-resolution PolSAR and PolInSAR features are with azimuth (height) and ground range (width) sampling spacings of 6 m and 15 m, and the simulated low-resolution GEDI LiDAR height is with azimuth and ground range sampling spacings of 600 m and 60 m. Therefore, the scaling factors are set as 100 and 4 in the azimuth and range directions. For a better performance over pan-sharpening tasks with large upsampling scales, a progressive pan-sharpening strategy (Lai et al., 2018, Cai and Huang, 2020) is adopted, where the generator is decomposed into two sub-networks and each one of them is responsible for a 10×2 pan-sharpening as presented in Figure 4.6. Referring to the architecture of Pan-GAN (Ma et al., 2020),

the sub-network of the generator is built by a dense connection (Iandola et al., 2014) among four convolution layers with filter sizes of 9×9, 5×5, 5×5, and 3×3, strides of 1 with padding, and channels of 256, 128, 64, and 1. This dense connection takes high-resolution PolInSAR and PolSAR features and the upsampled low-resolution LiDAR height (nearest neighbor) as inputs and the activation function in each layer is set as ReLU with batch normalization (BN) except for the last one. All inputs are normalized to the same scale before feeding into the generator.

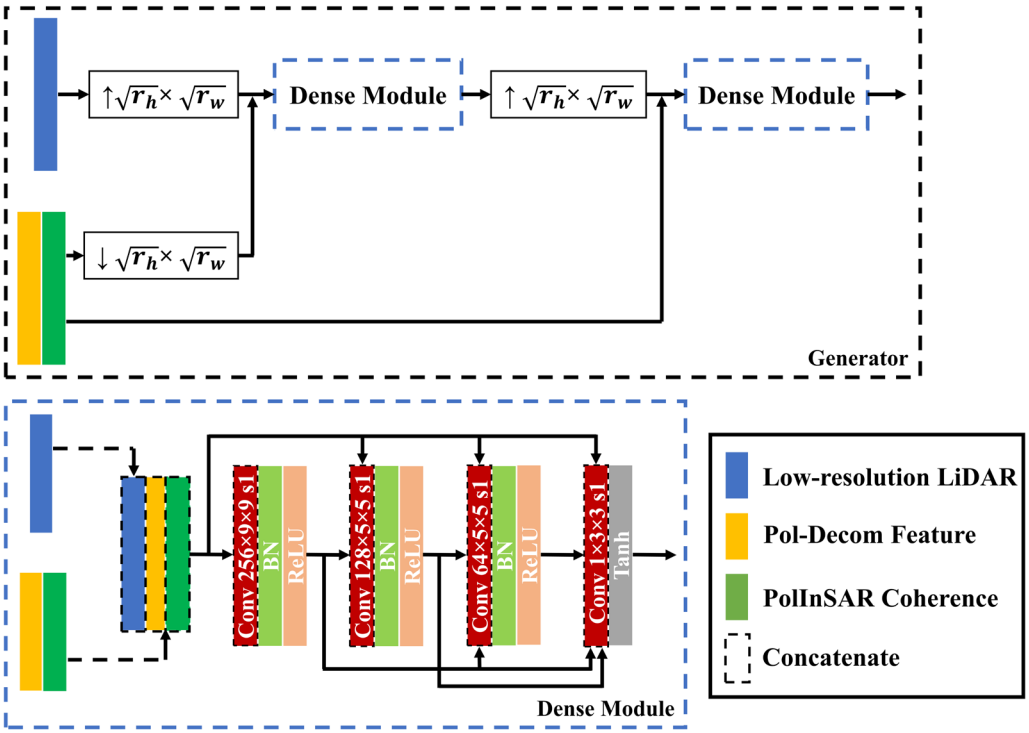


Figure 4.6 The architecture of the progressive generator with the basic dense module in the blue box.

The loss function of the generator is defined by the sum of vertical and spatial losses.

$$\mathcal{L}_G = \mathcal{L}_{vertical} + \mathcal{L}_{spatial} \tag{4.5}$$

where $\mathcal{L}_G$ is the total loss of the generator, $\mathcal{L}_{vertical}$ is the sum of all vertical-related losses, and $\mathcal{L}_{spatial}$ includes all spatial-related losses.

$$\mathcal{L}_{vertical} = \frac{1}{2N}\sum_{n=1}^{N}\left(\left\|\downarrow \widehat{H}_{HR_1} - H_{LR}\right\|_F^2 + \left\|\downarrow \widehat{H}_{HR_2} - H_{LR}\right\|_F^2\right) + \alpha\mathcal{L}_{adv1} \quad (4.6)$$

where N is the number of training data, $\downarrow$ represents the slicing downsampling, $\widehat{H}_{HR_1}$ and $\widehat{H}_{HR_2}$ are the generated forest heights from two sub-networks, $\|\cdot\|_F^2$ is the matrix Frobenius norm and $\mathcal{L}_{adv1}$ represents the adversarial loss derived from the coherence discriminator D_c balanced by a regularization parameter α.

$$\mathcal{L}_{adv1} = \frac{1}{N}\sum_{n=1}^{N}\left(D_c\left(CS\left(\widehat{H}_{HR_2}\right)\right) - 1\right)^2 \quad (4.7)$$

where CS is the coherence simulation operation. The spatial loss in Equation (4.5) can be written as

$$\mathcal{L}_{spatial} = \frac{\beta}{2N}\sum_{n=1}^{N}\left(\left\|\nabla\widehat{H}_{HR_1} - AP\left(\nabla(\downarrow Pol_{HR})\right)\right\|_F^2 + \left\|\nabla\widehat{H}_{HR_2} - AP(\nabla Pol_{HR})\right\|_F^2\right) + \gamma\mathcal{L}_{adv2} \quad (4.8)$$

where ∇ denotes the high pass filter extracting the high-frequency spatial information, $AP(\cdot)$ denotes the average pooling operation in the channel dimension, β and γ are the regularization parameters, and $\mathcal{L}_{adv2}$ represents the adversarial loss derived from the spatial discriminator D_s.

$$\mathcal{L}_{adv2} = \frac{1}{N}\sum_{n=1}^{N}\left(D_s\left(\widehat{H}_{HR_2}\right) - 1\right)^2 \quad (4.9)$$

4.3.3.2 Coherence simulator

The output of the generator is then processed by a coherence simulator before feeding into the coherence discriminator. This coherence simulator is constructed based on a simplified version of the RMoG model to generate a simulated volumetric-temporal

coherence, $\tilde{\gamma}_{vt}$, for a non-zero baseline geometry or a simulated temporal coherence, $\tilde{\gamma}_{t}$, for a zero-baseline geometry using forest height, random motion, and extinction coefficient inputs. In an RMoG model, coherences of the repeat-pass acquisition with a short temporal baseline can be approximated as a coherence function in Equation (4.10) after the compensation of the spectral and SNR decorrelations.

$$\gamma = \exp(j\phi_g)\frac{\gamma_{vt} + \mu(\vec{\omega})}{1 + \mu(\vec{\omega})} = \exp(j\phi_g)\left(\gamma_{vt} + \frac{\mu(\vec{\omega})}{1 + \mu(\vec{\omega})}(1 - \gamma_{vt})\right) \tag{4.10}$$

where ϕ_g is the ground phase, $\mu(\vec{\omega})$ is the ground to volume amplitude ratio varying with polarization $\vec{\omega}$, and γ_{vt} represents the coherence component with coupled effects from volume scattering and wind-derived temporal change.

Under the assumption of exponential scattering decay and a Gaussian random motion in the volume layer, γ_{vt} can be described by

$$\gamma_{vt} = \frac{\int_0^{h_v} \exp\left(\frac{2\sigma}{\cos\theta} z\right) \exp\left(-\frac{1}{2}\left(\frac{4\pi}{\lambda}\right)^2 \sigma_v^2(z)\right) \exp(jk_z z) dz}{\int_0^{h_v} \exp\left(\frac{2\sigma}{\cos\theta} z\right) dz} \tag{4.11}$$

where h_v denotes the forest height, σ is the mean extinction coefficient, θ is the local incidence angle, λ is the radar wavelength, k_z is the vertical wavenumber, and $\sigma_v(z)$ is the standard deviation of the Gaussian random motion which linearly varies with the height (Lei and Siqueira, 2014). γ_{vt} can be computed in a closed form as function of forest height, mean extinction coefficient, and random motion factor τ_1, where θ and k_z are data layers supplied on the UAVSAR website given by

$$\gamma_{vt} = \sigma\sqrt{\frac{\pi}{\tau_1}}\exp\left(\frac{\left(\frac{2\sigma}{\cos\theta}+jk_z\right)^2}{4\tau_1}\right)$$

$$\cdot\frac{\left[\operatorname{erf}\left(\frac{2h_v\tau_1-\left(\frac{2\sigma}{\cos\theta}+jk_z\right)}{2\sqrt{\tau_1}}\right)+\operatorname{erf}\left(\frac{\frac{2\sigma}{\cos\theta}+jk_z}{2\sqrt{\tau_1}}\right)\right]}{\cos\theta\left(\exp\left(\frac{2\sigma}{\cos\theta}h_v\right)-1\right)} \tag{4.12}$$

The PolInSAR coherence function in Equation (4.10) is a straight line in the complex plane, which can be fitted by coherence observations at different polarizations using the least square method (Cloude and Papathanassiou, 2003, Hajnsek et al., 2009). This line intersects the unit circle at two points one of which is the ground phase $\exp(j\hat{\phi}_g)$. Methods for selecting the correct ground phase have been discussed in several studies (Kugler et al., 2015, Denbina et al., 2018, Pourshamsi et al., 2018) and thus won't be described in detail here. Then, $\exp(j\hat{\phi}_g)\hat{\gamma}_{vt}$ is determined by the furthest coherence point from the estimated ground solution and the observation of volumetric-temporal coherence $\hat{\gamma}_{vt}$ can be extracted by removing the ground phase.

In the zero-baseline scenario with $k_z \approx 0$, γ_{vt} degrades into a real temporal coherence γ_t based on the mean value theorem for integration, which is fully determined by the forest height and the random motion factor τ_2,

$$\gamma_t = \exp(-\tau_2 h_v^2). \tag{4.13}$$

In this situation, the effect of volumetric scattering is negligible and the PolInSAR coherence in Equation (4.10) can be reformulated as a real function as,

$$\gamma = \frac{\gamma_t + \mu(\vec{\omega})}{1+\mu(\vec{\omega})}. \tag{4.14}$$

Since $\gamma_t < 1$ and $\mu(\vec{\omega}) > 0$, it can be concluded that the reformulated PolInSAR coherence in Equation (4.14) monotonically increases with the ground-to-volume ratio and the observation of temporal coherence $\hat{\gamma}_t$ can be estimated through the minimum of coherence observations at different polarizations.

The coherence simulator is constructed based on Equation (4.12) in non-zero baseline geometry (Lope) or Equation (4.13) in zero-baseline geometry (Howland), where the forest height input is produced by the generator. The random motion and extinction coefficient inputs are generated by upsampling the relatively low-spatial-resolution estimates derived from the low-spatial-resolution PolInSAR coherences and the sparse LiDAR heights. This is reasonable because the scattering attenuation and random motion effects are relatively consistent in a local area and thus can be considered as a constant value among locally neighboring pixels (Simard and Denbina, 2017).

4.3.3.3 Discriminator

A spatial and a coherence discriminator are used in PolGAN to refine the spatial and vertical distributions of the generated forest height, respectively. Instead of a vertical discriminator, we developed a coherence discriminator for the vertical refinement because PolInSAR coherences implicitly provide high-spatial-resolution descriptions of the scattering vertical structure. This coherence discriminator learns to distinguish the difference between the observed and simulated temporal or volumetric-temporal coherences and gives feedback to the generator through coherence adversarial loss. Associated observations of volumetric-temporal $\hat{\gamma}_{vt}$ or temporal $\hat{\gamma}_t$ coherences and the simulated volumetric-temporal $\tilde{\gamma}_{vt}$ or temporal $\tilde{\gamma}_t$ coherences are obtained based

on the methods in Section 4.3.3.2. Like Pan-GAN, a spatial discriminator is deployed to refine the spatial distribution in the generated height image. Inputs of the spatial discriminator are the high-pass filtering results of the generated forest height and the polarimetric decomposition features. An averaging operation is conducted along the feature dimension before feeding into convolution layers. As shown in Figure 4.7, both discriminators are fully constructed by convolutional neural networks (CNN) but with different inputs. Each of them includes six layers with filter sizes of 3×3, strides of 2 without padding except for the last one (filter: 6×6, stride: 1), and channels of 16, 32, 64, 128, 256, and 1. Batch normalization is used except for the first layer and activation functions are set as Leaky ReLU for each layer.

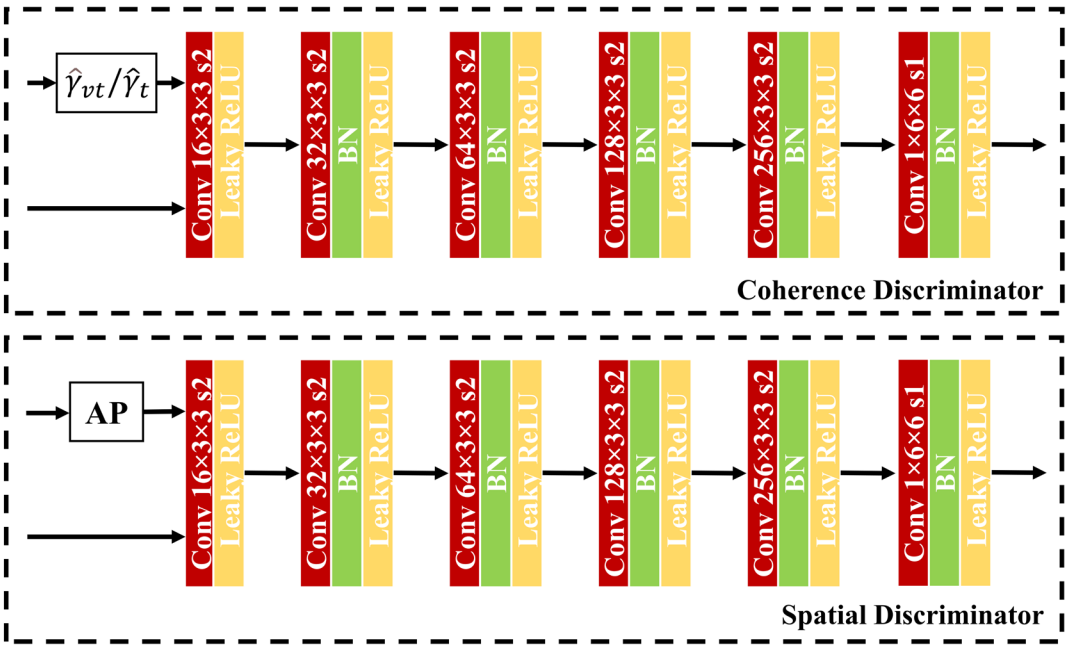


Figure 4.7 The architecture of the coherence and spatial discriminators

4.3.4 Implementation details

High-spatial-resolution features and the associated low-spatial-resolution height over each site are cropped into small patches (50% overlaps) in sizes of 200×200 and 2×50, respectively, and then fed into the generators and discriminators. The batch size is set as 16 in the training process, the RMSProp optimizer is initialized with a base learning rate of 0.001 and a decay rate of 0.99, the regularization parameters α, β, and γ in Equations (4.6) and (4.8) are set as 0.05, 1, and 0.01, and the training epochs are set as 80 for both Howland and Lope sites. Ablation experiments are conducted over three more structures (a progressive generator with two dense connections and a spatial discriminator, a progressive generator with two dense connections and a coherence discriminator, and a standard generator with one dense connection and dual discriminators) to explore the contributions of each discriminator and different generator architectures to the pan-sharpening result. Related training details are the same as mentioned above, except that $\alpha = 0$ when there is no coherence discriminator and $\gamma = 0$ when there is no spatial discriminator. All experiments are implemented in the Pytorch framework based on two NVIDIA GeForce RTX 2080Ti GPUs. Besides that, comparative experiments are also conducted between PolGAN and a prior model-based Kapok method (Denbina and Simard, 2017, Simard and Denbina, 2018) to validate the superiority of the proposed method.

4.4 Result and Discussion

4.4.1 PolGAN results

Forest heights $\widehat{H}_{HR}$ produced by PolGAN at different training stages are shown in this section. Quantitative evaluations are conducted based on accuracy metrics including the Root Mean Square Error (RMSE), Mean Signed Error (BIAS), and Coefficient of Determination (R^2) (Denbina et al., 2018, Simard and Denbina, 2018). The RMSE, BIAS, and R^2 measure the absolute difference, the signed difference, and dependency between the predicted and the reference values, respectively. Figure 4.8 represents the evolution of the generated forest height over the Howland site. Related quantitative evaluations are conducted over a region of interest (ROI) in the red frame and the PolGAN derived heights are plotted versus high-resolution RH100 heights as shown in Figure 4.9. The calculated RMSE and BIAS are precise to the centimeter, and R^2 is precise to two decimal places as given in Table 4.2, where the negative BIAS indicates that the estimated height is overall lower than the reference RH100 height and vice versa. Epoch 0 corresponds to the upsampling height produced by the nearest neighbor.

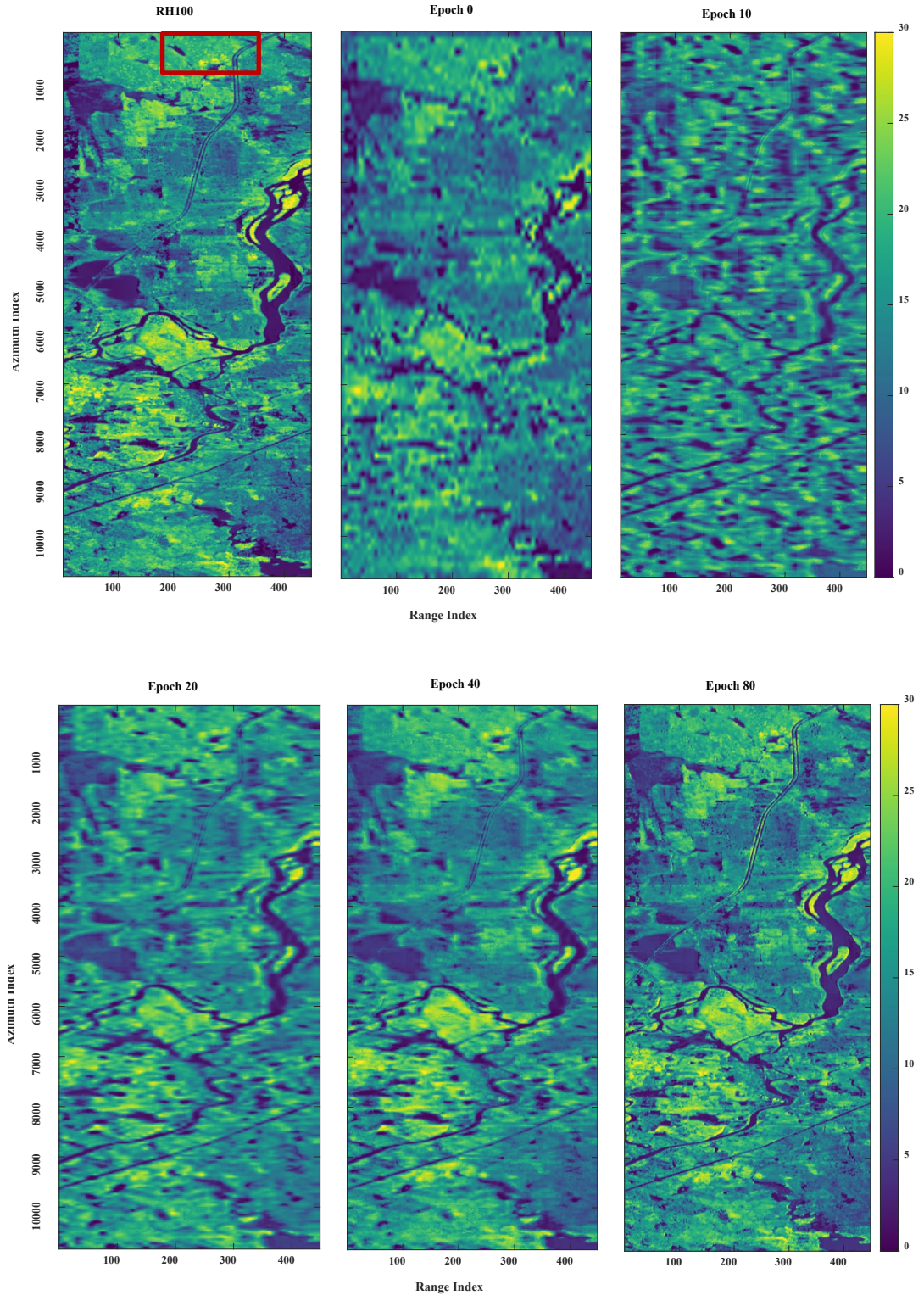
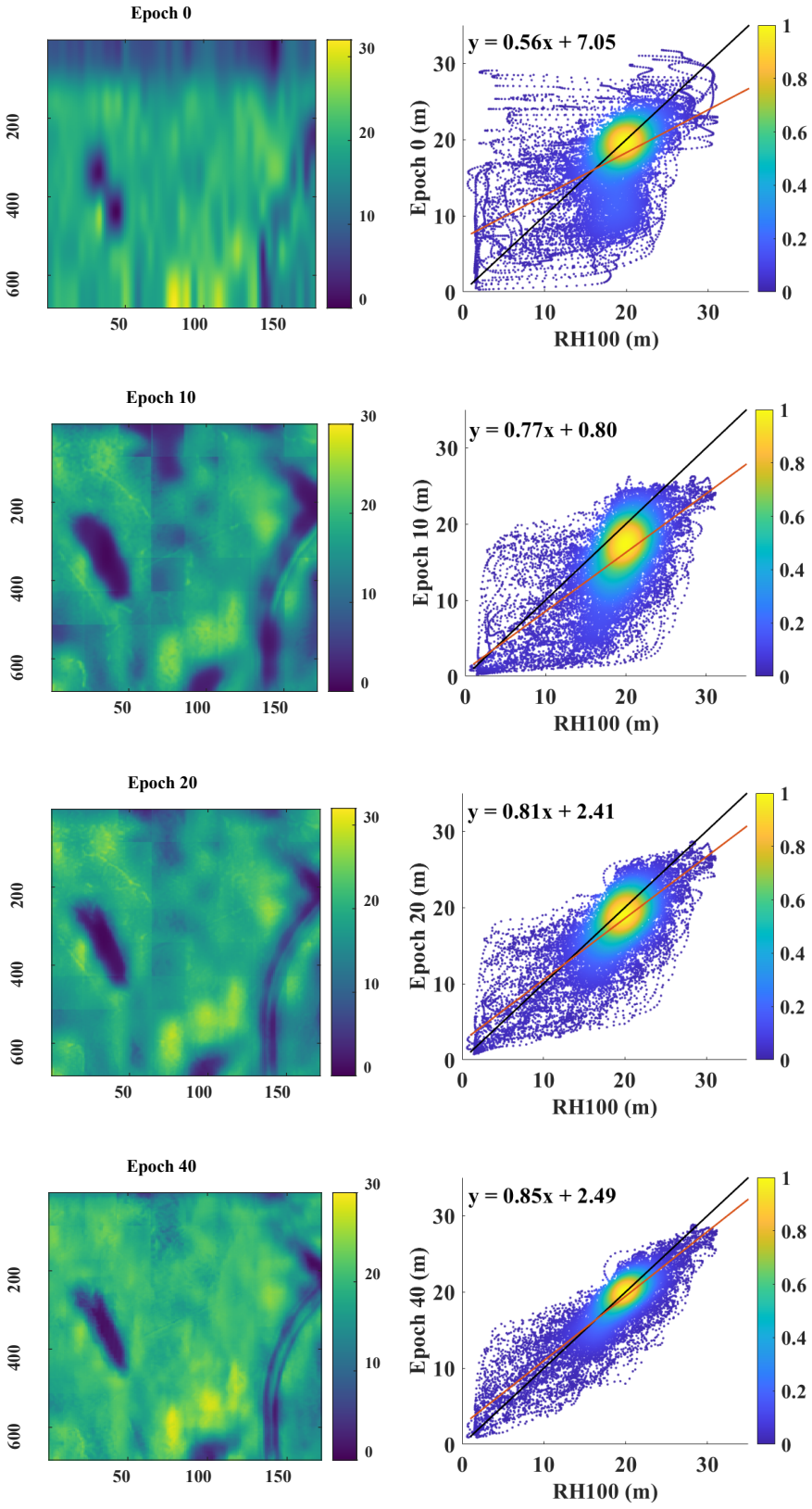


Figure 4.8 PolGAN derived forest height maps over the Howland site at different training stages.


Epoch 0
y = 0.56x + 7.05
Epoch 0 (m)
RH100 (m)
Epoch 10
y = 0.77x + 0.80
Epoch 10 (m)
RH100 (m)
Epoch 20
y = 0.81x + 2.41
Epoch 20 (m)
RH100 (m)
Epoch 40
y = 0.85x + 2.49
Epoch 40 (m)
RH100 (m)

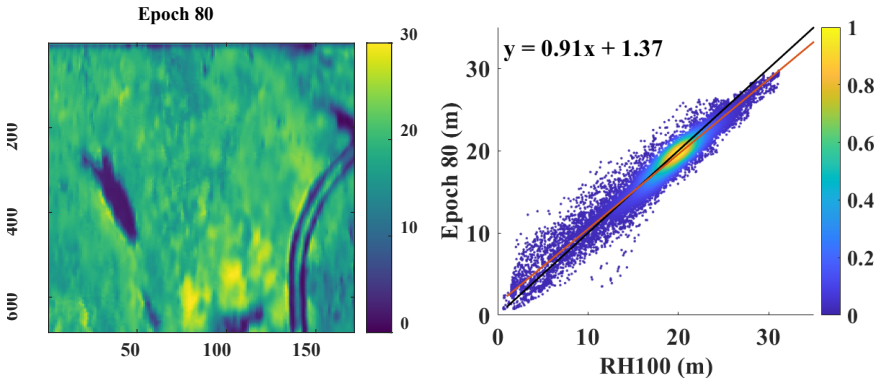


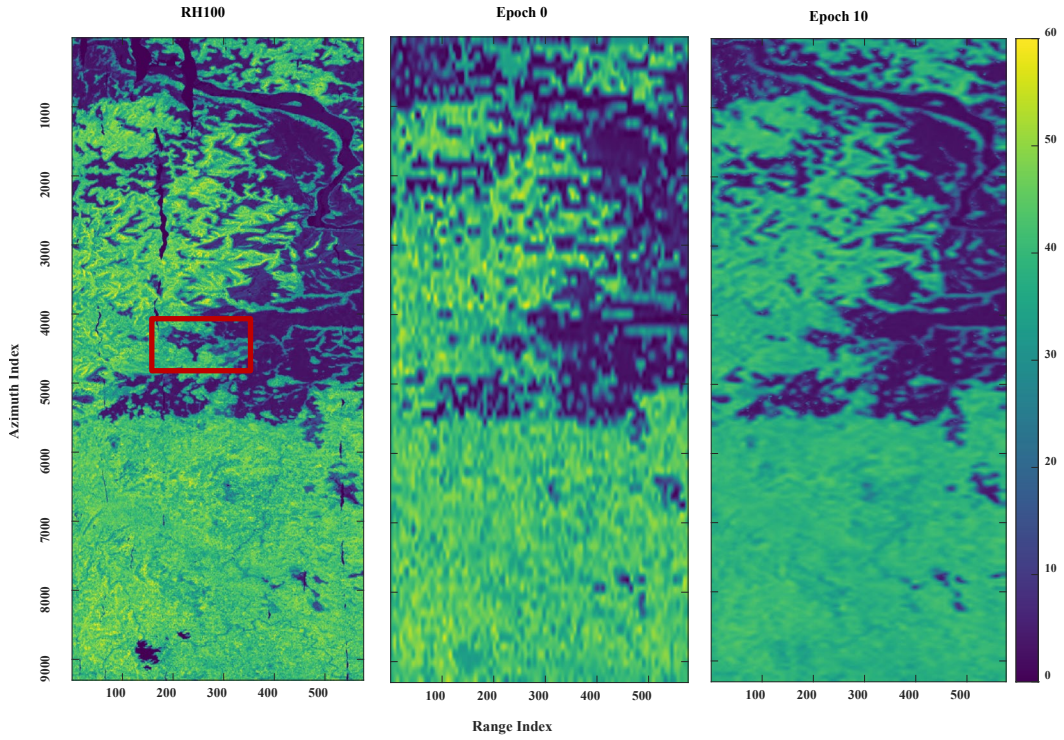
Figure 4.9 Evolution of PolGAN derived forest height over ROI (left column) and associated density plots versus RH100 heights (right column), where black lines are y=x, red lines are the linear fits

Table 4.2. Evolution of the accuracy metrics over ROI

Epoch	0	10	20	40	80
RMSE	6.72 m	5.26 m	3.39 m	2.19 m	1.21 m
BIAS	-1.10 m	-3.42 m	-1.21 m	-0.33 m	-0.27 m
R^2	0.26	0.45	0.59	0.78	0.93

As illustrated in Figure 4.8, at the earlier stages of the training, the generated forest heights are with coarse spatial details and lower vertical accuracies. After several iterations, spatial and vertical distributions are gradually refined by the spatial and coherence discriminator, and the generated forest height is converged to the high-resolution RH100 height. Results over ROI in Figure 4.9 present the evolution of the generator more visually, where boundaries of water bodies and bare roads are getting clearer and more details in the vegetation height are gradually revealed by training. This is consistent with the decreasing RMSE (6.72 m to 1.21 m) and the increasing R^2 (0.26 to 0.93) metrics in Table 4.2.

Figure 4.10 shows the evolution of the generated forest height map over the Lope site. Some areas are not covered by the LVIS LiDAR acquisitions and associated pixels are set as NoData in the high-resolution RH100 height map. The number of NoData pixels is largely reduced in the low-resolution LiDAR height after downsampling and the NoData values are further set as zero in the normalized low-resolution input. Similarly, quantitative evaluation is conducted over the ROI in the red frame, where the NoData pixels in the high-resolution RH100 height map are not included in the calculation of any accuracy metrics. PolGAN-derived results are plotted versus the reference RH100 height as shown in Figure 4.11 along with related RMSE, BIAS, and R^2 metrics listed in Table 4.3.



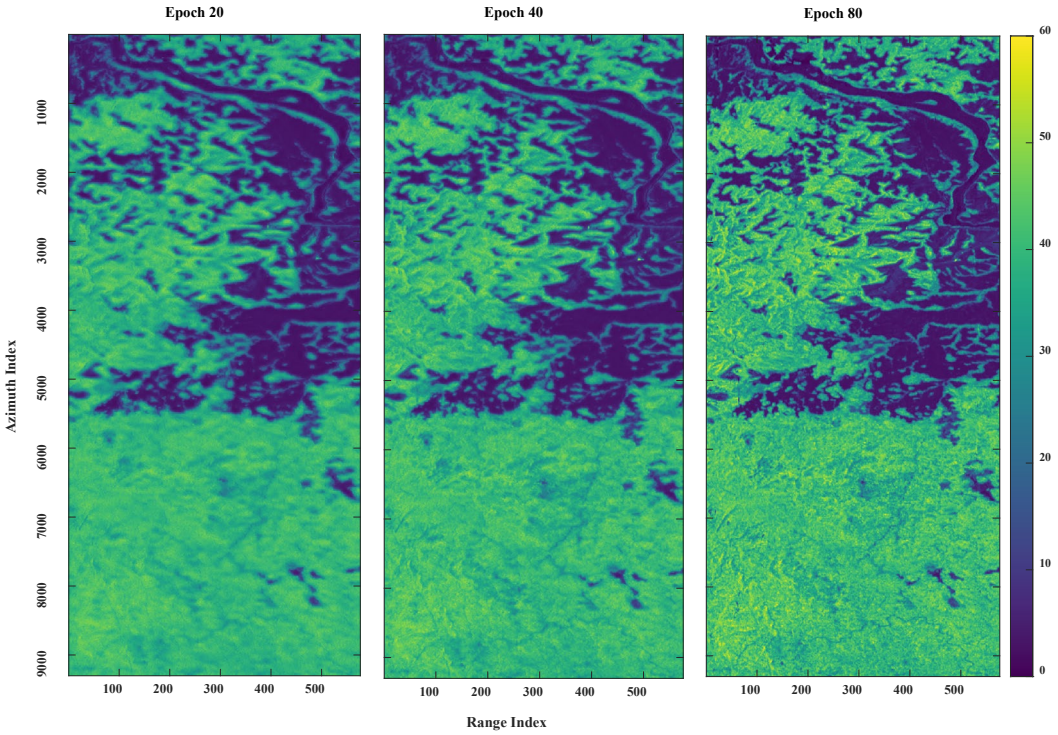
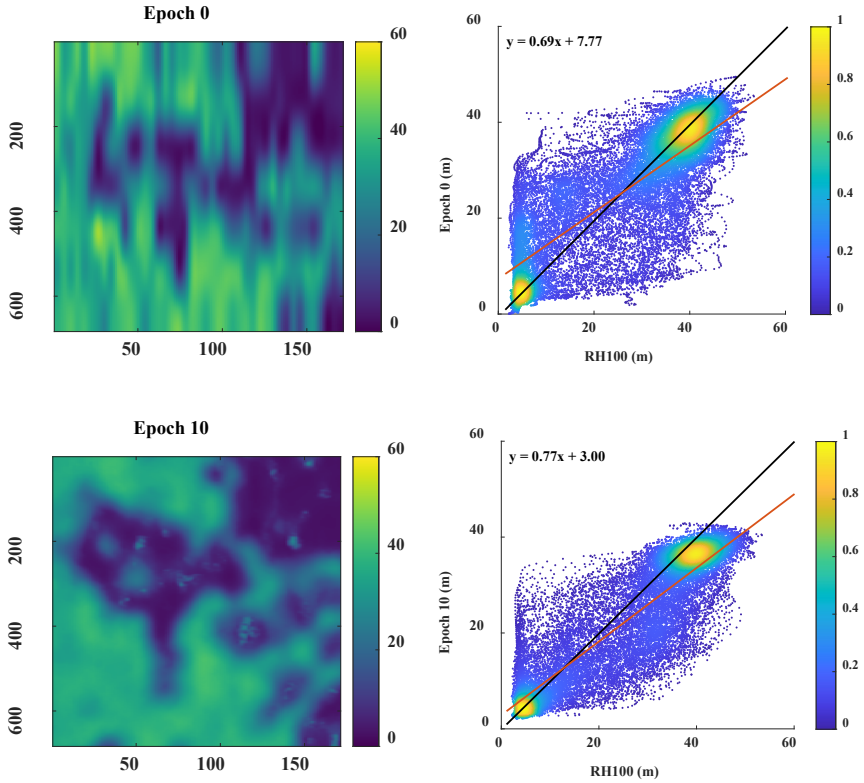


Figure 4.10 PolGAN derived forest height maps over the Lope site at different training stages.



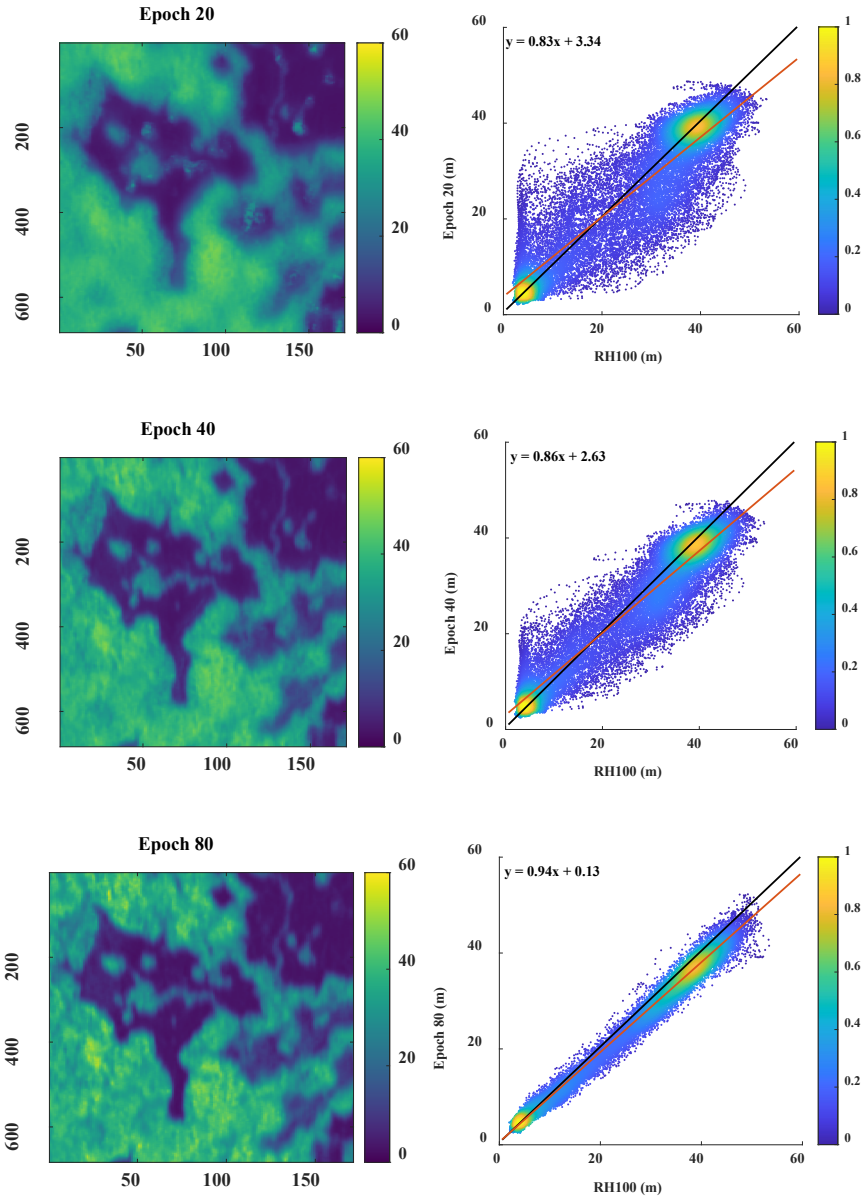


Figure 4.11 Evolution of PolGAN derived forest height over ROI (left column) and associated density plots versus RH100 heights (right column).

Table 4.3. Evolution of the accuracy metrics over ROI

Epoch	0	10	20	40	80
RMSE	9.14 m	7.62 m	6.64 m	5.08 m	2.37 m
BIAS	-0.06 m	-2.94 m	-0.94 m	-1.01 m	-1.25 m
R^2	0.62	0.76	0.80	0.87	0.98

Similar evolution progress is observed in the generated forest height maps over the Lope site in Figure 4.10, where boundaries of short savannas become prominent and the vertical accuracies of tall evergreen broadleaf forests increase with training. This is also indicated in the decreasing RMSE (9.14 m to 2.37 m) and the increasing R^2 (0.62 to 0.98) in Figure 4.11. Note that RMSEs of the Lope site are overall higher than those of the Howland site, which is related to the height ranges at the two forest sites (Howland: 0-35 m, Lope: 0-68 m). Low-resolution LiDAR inputs are normalized to the range of (-1, 1) before feeding into the generator and the generator outputs are rescaled to the initial range as the estimated forest height. Errors in the outputs are amplified due to this scaling and hence larger scaling ratios lead to lower accuracies in the forest height results.

4.4.2 Ablation studies

As mentioned in section 4.3.4, ablation studies are conducted on models with different generator and discriminator structures. Related experiments are first performed on the progressive (P) generator with different discriminators (spatial only, coherence only, and dual discriminators) to explore the contributions from each discriminator. This is followed by an experiment on the standard (S) generator with dual discriminators to validate the superiority of the progressive architecture. Related forest height maps are presented in Figure 4.12 and quantitative evaluation is conducted over the same ROI in Howland. The generated heights are plotted versus high-resolution RH100 heights as presented in Figure 4.13 and related accuracy metrics with the same decimal places are summarized in Table 4.4.

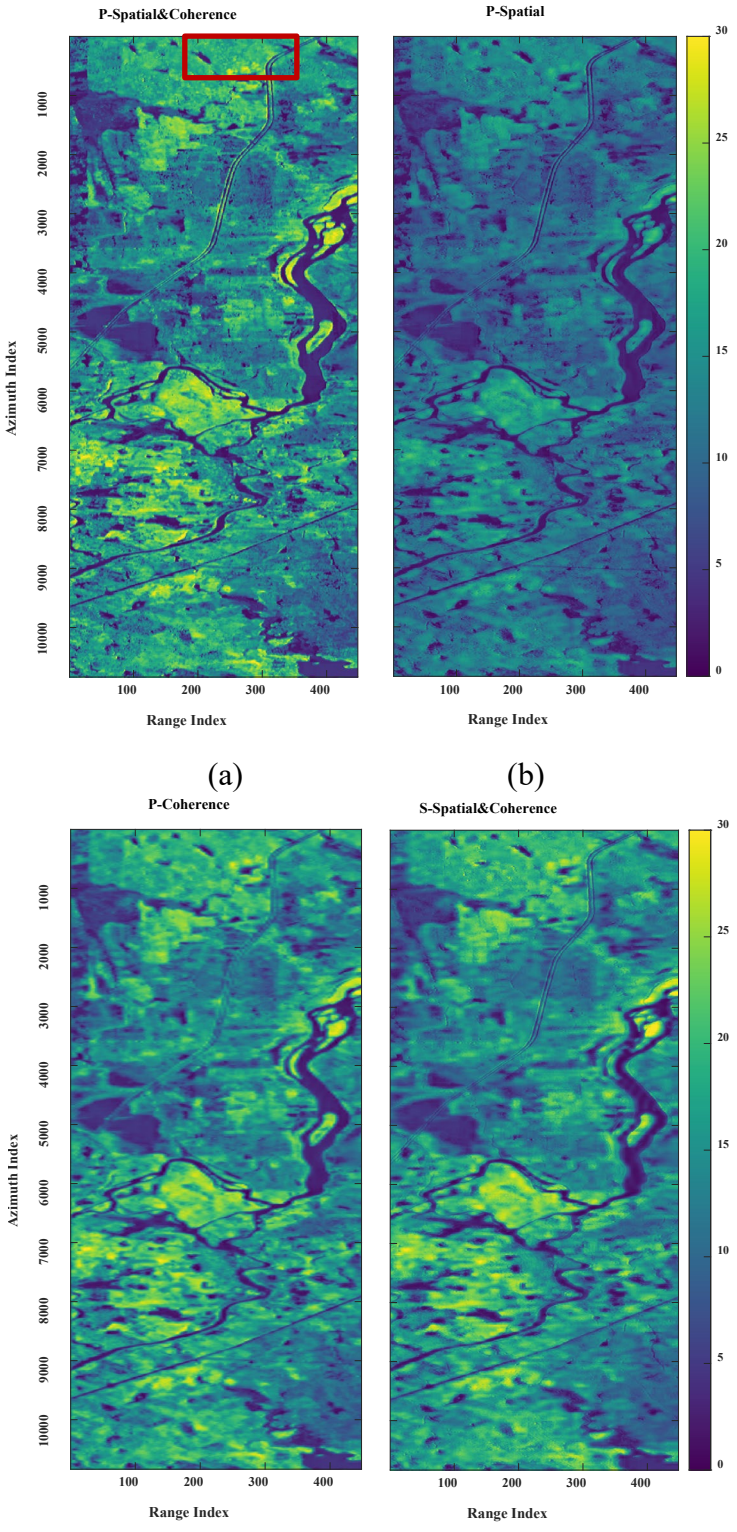
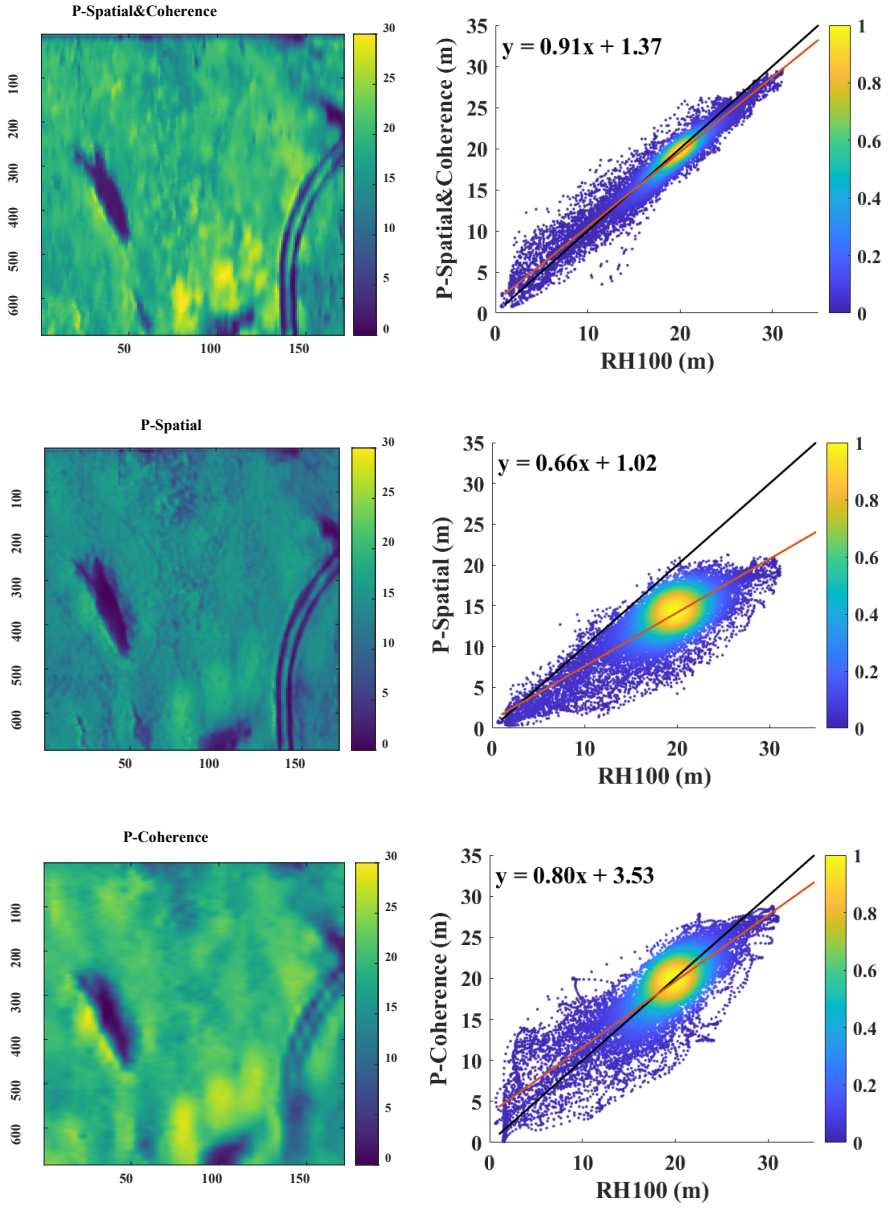


Figure 4.12 Forest height maps generated by models with different generator and discriminator structures: (a) Progressive generator with dual discriminator (P-

Spatial&Coherence); (b) Progressive generator with spatial discriminator (P-Spatial); (c) Progressive generator with coherence discriminator (P-Coherence); (d) Standard generator with dual discriminators (S-Spatial&Coherence).



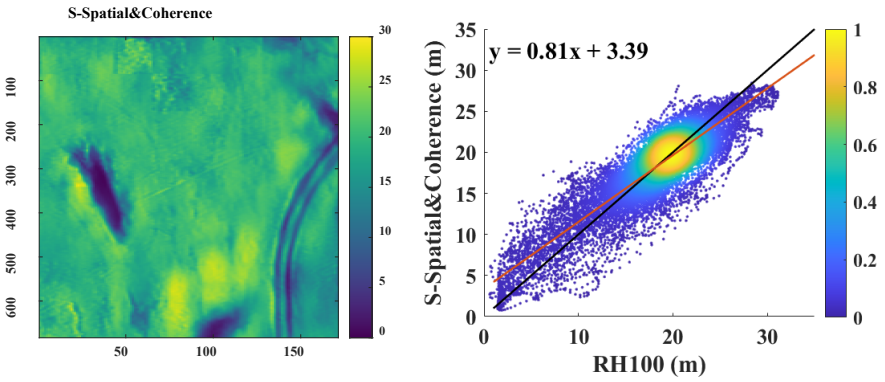


Figure 4.13 Forest height over ROI generated by models with different generator and discriminator structures (left column) and associated density plots versus RH100 heights (right column).

Table 4.4 Accuracy metrics of the ablation experiments

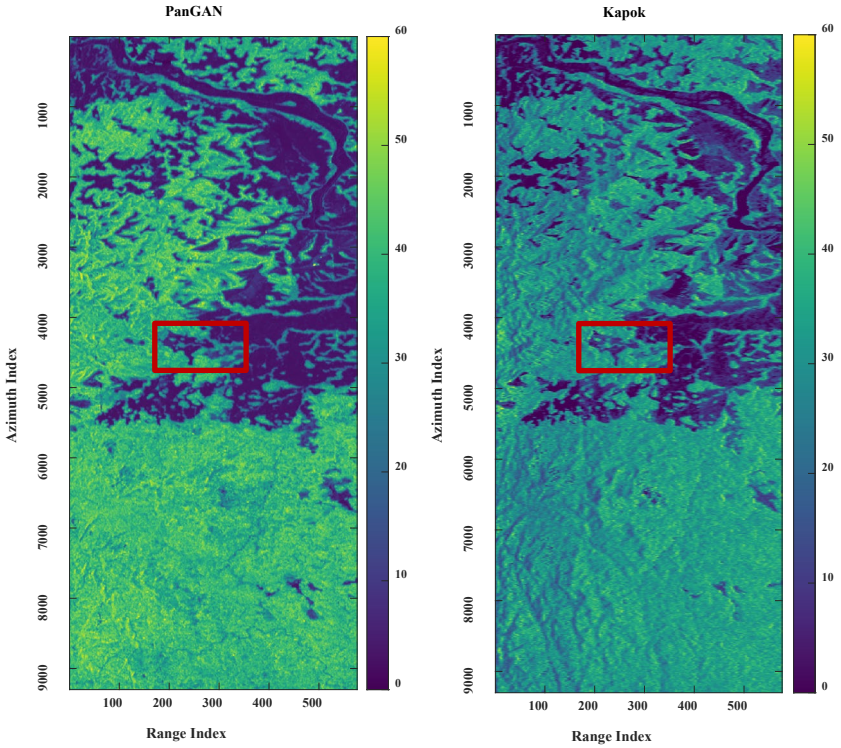
Metric	P-Spatial&Coherence	P-Spatial	P-Coherence	S-Spatial&Coherence
RMSE	1.21 m	5.87 m	2.74 m	2.43 m
BIAS	-0.27 m	-5.34 m	-0.10 m	-0.08 m
R^2	0.93	0.72	0.67	0.80

As presented in Figure 4.12, Figure 4.13, and Table 4.4, the forest height generated by the progressive generator with spatial discriminator has much lower vertical accuracy than the PolGAN-derived result (P-Spatial&Coherence) although the spatial details are quite clear. On the contrary, coarse spatial details but higher vertical accuracy can be observed in the forest height result derived from the progressive generator with a coherence discriminator. This indicates the responsibility of each discriminator and the necessity of dual discriminators during the forest height pan-sharpening process, where the spatial discriminator mainly focuses on the refinement of the spatial details while the coherence discriminator is dedicated to refining the vertical details.

It can also be concluded from Figure 4.12, Figure 4.13, and Table 4.4, the progressive generator with two dense connections used in the PolGAN (RMSE: 1.21 m; R2: 0.93) exhibits a better performance on the pan-sharpening task with large upsampling scales compared with the standard generator with one dense connection (RMSE: 2.43 m; R^2: 0.80). This is because the progressive generator has a much deeper network structure than the standard one and hence offers more parameters to simulate the pan-sharpening process. Additionally, training of the progressive generator is guided by two ground-truth-derived losses at different spatial resolutions, whereas the training of standard architecture is only guided by one.

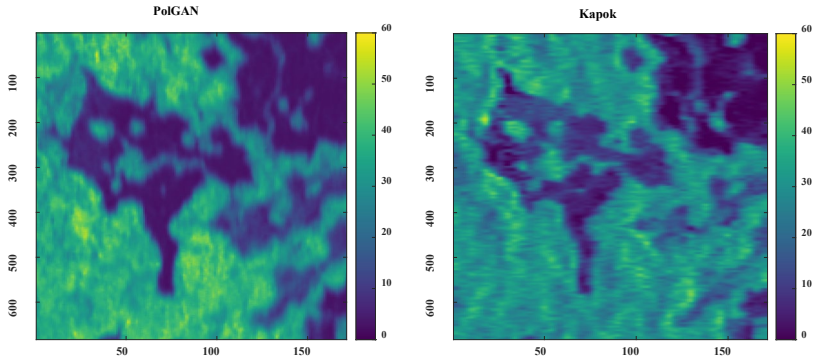
4.4.3 Comparative experiments

The comparison experiment is conducted between the PolGAN and Kapok methods. Kapok is a Python library developed by JPL to estimate forest height using the repeat-pass PolInSAR data. The library can import UAVSAR acquisitions and generate forest height estimates based on the RVoG model. Forest height maps derived from the two methods are illustrated in Figure 4.14 and quantitative evaluation is conducted over the same ROI in the Lope site. The generated heights are plotted versus high-resolution RH100 heights as presented in Figure 4.15, where the RMSE, BIAS, and R^2 of the PanGAN result are 2.37 m, -1.25 m, and 0.98 as listed in Table 4.3, and the RMSE, BIAS, and R^2 of the Kapok result are 8.02 m, -1.91 m, and 0.73.

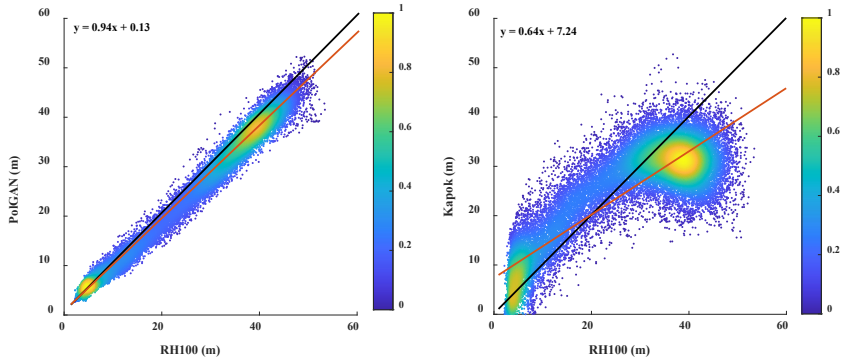


(a) (b)

Figure 4.14 (a) PolGAN-derived forest height map; (b) Kapok-derived forest height map.



(a)



(b)

Figure 4.15 (a) Forest height over ROI generated by PolGAN and Kapok methods; (b) Density plots versus RH100 heights.

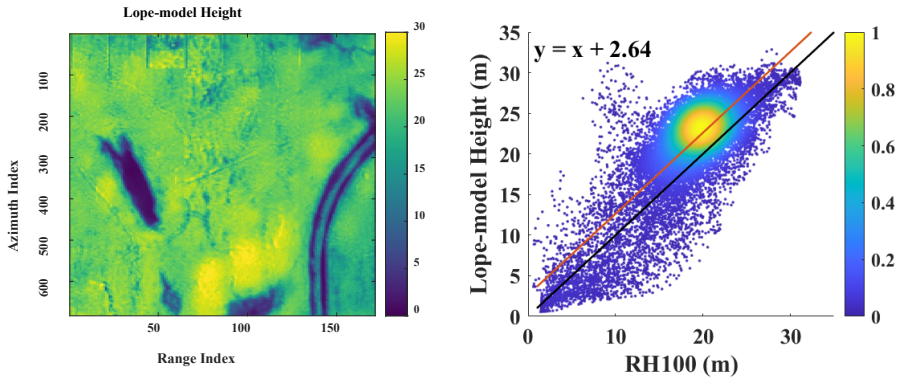
As a typical model-based forest height inversion method, Kapok estimates the forest height by solving nonlinear equations of forest parameters, imaging geometry, sensor parameters, and coherence observations established based on the RVoG model. It has been mentioned in Section 4.1, RVoG model simplifies the forest as a homogeneous volume of randomly oriented scattering particles statically distributed over the ground, where the temporal decorrelation and the vertical heterogeneity are not considered in this model. Therefore, discrepancies between RVoG model and actual backscattering behaviors inevitably exist especially when using repeat-pass L-band PolInSAR data, where dynamic changes of forest properties occur within different acquisition times and longer radar waves in L-band are more likely to interact with scattering elements in the forest canopy (more sensitive to the forest structure). Besides, the Kapok method estimates pixel-wise forest heights using the coherence observations in each radar pixel meaning that the spatial characteristics among pixels like shape and texture are not considered during the inversion.

The proposed PolGAN method is intended to redress the above-mentioned challenges. Firstly, the forest height inversion is reformulated as an unsupervised pan-sharpening process on basis of the high-resolution polarimetric decomposition and polarimetric interferometry features and the low-resolution LiDAR height. The relation established based on the theoretical model is only employed before the coherence discriminator as a part of the vertical refinement of the generator. This significantly reduces the conformity requirement between the theoretical model and the actual scene. Secondly, both the generator and the two discriminators are constructed based on CNN, which has been widely used to extract shape, texture, and spatial characteristics in computer vision tasks and therefore provide PolGAN with the capability to use information among pixels that were not considered in the prior methods.

4.4.4 Discussion

At the current stage, PolGAN experiments are mainly conducted at the regional scale using the airborne PolInSAR data and the simulated GEDI data. For further application at the continental or/and global scale, several challenges need to be addressed. The first one is the availability of large-scale Lidar and PolInSAR collections. Spaceborne LiDAR missions like GEDI and Ice, Cloud, and land Elevation Satellite-1&2 (ICESat-1&2) have been dedicated to providing global LiDAR observations since 2003 and the upcoming spaceborne SAR missions such as the NISAR, the European Space Agency (ESA)'s BIOMASS (Carreiras et al., 2017), and German Aerospace Center (DLR)'s Tandem-L (Moreira et al., 2015) are designed to get L-band PolInSAR measurements in single- or repeat-pass configurations within days. The second challenge would be the efficiency of the data processing method. For the two test sites involved in this

study, it averagely takes about 6 hours to get the final forest height estimation result over each forest site when implemented by two NVIDIA GeForce RTX 2080Ti GPUs. Attemptations have been made to explore whether PolGAN needs to be retrained for every new region, where the model trained over the Lope site is applied to Howland acquisitions to get the forest height pan-sharpening result. As presented in Figure 4.16, the RMSE, BIAS, and R^2 of the Howland forest height generated by the model trained on the Lope site are 4.11 m, 2.61 m, and 0.67. RMSE is 2.90 m higher and R^2 is 0.26 lower than the result in Figure 4.9 indicating that retraining is necessary to get forest height estimates with higher accuracy. A possible way to improve the efficiency of PolGAN is to employ pre-trained weights from previous training during the retraining process. The third challenge is about the difference between airborne and spaceborne remote sensing. For example, the spatial resolution of spaceborne data is normally lower compared with airborne, the temporal baseline of the spaceborne repeat-pass SAR acquisition is much longer, and the vertical accuracy of the spaceborne LiDAR is lower. Therefore, further adjustments on the PolGAN model (architectures and hyperparameters) are inevitable to obtain better forest height estimations on different spaceborne missions. In summary, there are still ways to go for the global application of the proposed PolGAN, but its potential on the forest height estimation with high spatial resolution and vertical accuracy will greatly benefit further applications like global land management and biomass level monitoring (Mette et al., 2002, Mette et al., 2004a, Mette et al., 2004b, Duncanson et al., 2020).



(a) (b)

Figure 4.16 (a) Forest height over Howland ROI generated by the model trained on the Lope site; (b) Density plots versus RH100 heights.

4.5 Conclusion

This chapter proposed an unsupervised deep-learning-based forest height estimation method based on the synergy of the high-resolution repeat-pass PolInSAR and low-resolution LiDAR acquisitions. The traditional model-based forest height inversion is reformulated as a pan-sharpening process between low-resolution LiDAR heights and high-resolution PolInSAR features and is further realized by a novel PolGAN network with one progressive generator and two discriminators. The generator absorbs the input features and produces forest height estimates with high spatial resolution and vertical accuracy and two discriminators distinguish spatial and coherence differences between the generated result and remote sensing observations and provide feedback to refine the generator. Compared with traditional PolInSAR forest height inversion methods, the proposed method exhibits less dependency on theoretical models and effectiveness in using shape, texture, and spatial information in high-spatial-resolution features. Results indicate that this method can achieve RMSEs of around 1.21 m over

the ROI in Howland and 2.37 m over the ROI in Lope and hence provide an effective approach for the forest height estimation with the synergy of PolInSAR and LiDAR data.

Chapter 5 Towards A Deep-Learning-Based Automated Active Fire Detection Framework Using Sentinel-2 Imagery

Most prior studies on active fire detection take advantage of the emitted radiance in the thermal infrared and mid-infrared bands with coarse spatial resolutions. However, the background radiance in a coarse pixel tends to overwhelm signals from subtle fires making it hard to reveal more spatial details on small fires. Though further efforts have been placed on active fire detection using the reflectance in the short wave infrared bands with medium resolutions, these methods are still based on pixel or region-level comparisons using sensor-specific thresholds and neighborhood statistics. Since prior researches have indicated that active fires normally exhibit higher reflectances in the SWIR bands while negligible contributions in the red band, the active fire detection process is formulated as a semantic segmentation on the SWIR-Red false-color images in this chapter. Studies mainly focus on exploring the potential of deep-learning-based semantic segmentation models on active fire detection, which is a relatively new field with open demands for datasets and architectures. To address the challenges, a dataset is constructed using the open-access Sentinel-2 imagery and the performance of three deep-learning-based segmentation models trained on this dataset are further evaluated. Modifications are performed on the architecture of the best-performing model based on the unique behaviors of active fires in SWIR-Red false-color images to boost the detection accuracy. An automated active fire detection framework is developed based on this modified model for its further application on global fire monitoring, which allows an automated Sentinel-2 data download, processing, and the generation of the

active fire detection results through the time and location information provided by the user. Related performance is evaluated in terms of detection accuracy and processing efficiency.

I have published this work in Remote Sensing (Zhang et al., 2021b).

Zhang Q, Ge L, Zhang R, et al. Towards a Deep-Learning-Based Framework of Sentinel-2 Imagery for Automated Active Fire Detection. *Remote Sensing* 2021; 13(23), 4790.

5.1 Introduction

Australia has experienced one of its most devastating fire seasons in 2019-2020, colloquially known as "Black Summer" (Kemter et al., 2021). Commenced in June 2019, the fire was out of control across Australia from September 2019 to March 2020 causing at least 33 deaths, a total burned area of almost 19 million hectares, and economic loss of over 103 billion AUD (Killian, 2020). Likewise, the United States has also experienced a series of wildfires since May 2020. By the end of that year, at least 46 deaths, a total burned area of over 4 million hectares, and an economic cost of almost 20 billion USD have been reported (Higuera and Abatzoglou, 2021). This dramatic increase in frequency and severity of wildfire events around the world is directly related to the changing climate that enhanced conditions for wildfires (low precipitation, high air temperature, and strong wind), urging more than ever before governments and fire management agencies to develop a reliable, timely, and cost-efficient fire monitoring system (Hua and Shao, 2017, Jones et al., 2020). Associated earth observation satellites have been deployed to monitor fire activity in two aspects,

burned area mapping and active fire detection (Giglio et al., 2008, Schroeder et al., 2016, Kumar and Roy, 2018, Chuvieco et al., 2019). This study will focus on the second one and detect the location of actively burning spots during the wildfire event.

Most prior studies on the active fire detection use the emitted radiance in the thermal infrared and mid-infrared bands because the radiance from the coolest fires is more than two orders of magnitude greater than the radiance from a nonburning land surface in the middle-infrared and about an order of magnitude greater in the thermal infrared (Kumar and Roy, 2017). Acquisitions in the thermal infrared and mid-infrared bands acquired by sensors like AVHRR, Visible Infrared Imaging Radiometer Suite (VIIRS), MODIS, Sea and Land Surface Temperature Radiometer (SLSTR), Visible and Infra-Red Radiometer (VIRR), etc., onboard the near-polar orbit satellites or/and Advanced Baseline Imager (ABI), Spinning Enhanced Visible and Infra-Red Imager (SEVIRI), and Space Environment Data Acquisition monitor (SEDI) onboard the geostationary satellites (Stroppiana et al., 2000, Wooster et al., 2012, Roberts and Wooster, 2014, Schroeder et al., 2014, Bessho et al., 2016, Giglio et al., 2016, Lin et al., 2018, Hall et al., 2019) are normally with coarse spatial resolutions. These data provide consistent measurements in near-real-time and have been deployed in several global monitoring systems such as the Digital Earth Australia Hotspots from Geoscience Australia (GA), European Forest Fire Information System (EFFIS) from ESA, Fire Information for Resource Management System (FIRMS) from NASA, etc. However, the background radiance in coarse resolution pixels tends to overwhelm signals from subtle hotspots due to the large instantaneous field of view (IFOV), and hence is hard to reveal active fires with more spatial details (Murphy et al., 2016).

Further efforts have been placed on the active fire detection using medium resolution acquisitions from Landsat and/or Sentinel-2 and several algorithms were developed based on the reflective wavelength bands (Murphy et al., 2016, Schroeder et al., 2016, Kumar and Roy, 2018, Hu et al., 2021). The Landsat-8 OLI data are mostly used to detect active fires at a spatial resolution of 30 m. The latest GOLI algorithm (Kumar and Roy, 2018) is proposed based on a statistical examination on the top of atmosphere (TOA) reflectance in the red and SWIR bands and exhibited comparable commission error (CE) and slightly lower omission error (OE) than other algorithms on the active fire detection results. This statistics-based method has been further applied to Sentinel-2 acquisitions and an AFD-S2 method (Hu et al., 2021) is proposed for active fire detection at a spatial resolution of 20 m.

Fast-growing deep-learning techniques (Zhu et al., 2017) provide great tools for large-volume remote sensing data processing inspiring us to explore their potentials on active fire detection. This is a relatively new field and still lacks large-scale datasets and architectures for training and testing. The first contribution to this field is based on the Landsat-8 imagery (de Almeida Pereira et al., 2021), which introduces a large-scale dataset to investigate how U-Net architecture can be applied to approximate the handcraft active fire maps (Ronneberger et al., 2015). This chapter will focus on investigating how deep learning architectures can be applied to the active fire detection of Sentinel-2 imagery, where active fire detection is formulated as a binary semantic segmentation task achieved by three different deep-learning-based models. A Dual-domain Channel-Position Attention (DCPA) network is designed based on the unique

characteristics of active fires in the false-color images to further improve the detection accuracy. A dataset is constructed using the open-access Sentinel-2 products collected over wildfires on the east coast of Australia and the west coast of the United States in 2019 and 2020. Based on the developed active fire detection network, an automated active fire detection framework with a data collection and preprocessing module, a deep-learning-based active fire detection module, and a final product generation module is proposed. This framework can automatedly output active fire detection results in a Geotiff format based on the geographic location and time input by the user. This Geographic Information System (GIS) preferred format is easily visualized and post-processed in software such as ArcGIS (Ormsby et al., 2004), QGIS (QGIS, 2018), etc.

5.2 Test Sites and Datasets

Sentinel-2 constellation is part of the ESA's Copernicus Program providing high-quality multi-spectral imagery of the Earth since 2015. The twin satellites share a sun-synchronous orbit at 786 km altitude with a repeat cycle of 10 days for one satellite and 5 days for two (Bouzinac et al., 2018). Each satellite carries an MSI with 13 bands in the visible, NIR, and SWIR ranges (Spoto et al., 2012). Prior studies indicate that fires with larger sizes and higher temperatures normally exhibit higher reflectances in SWIR bands while negligible contributions in the red band (Kumar and Roy, 2018, Hu et al., 2021). Hence, active fire detection using Sentinel-2 imagery can be realized by filtering out pixels with high values in B12 and B11 and low values in B4 as in the AFD-S2 method. Details of Sentinel-2/MSI bands involved in this study are listed in Table 5.1, where B4, B11, and B12 are employed as inputs of deep neural networks

for active fire detection, and combinations of B12-B11-B4 and B12-B11-B8A are used as an auxiliary for the ground truth annotation and the fire patch sifting.

Table 5.1 List of Sentinel-2/MSI band information

Channel	Sentinel-2A		Sentinel-2B		Spatial resolution
	Central Wavelength	Bandwidth	Central Wavelength	Bandwidth	
4 – Red	664.6 nm	31 nm	664.9 nm	31 nm	10m
8A – Narrow NIR	864.7 nm	21 nm	864.0 nm	22 nm	20m
11 – SWIR 1	1613.7 nm	91 nm	1610.4 nm	94 nm	20m
12 – SWIR 2	2202.4 nm	175 nm	2185.7 nm	185 nm	20m

Since most of the prior studies use TOA reflectance for active fire detection (Murphy et al., 2016, Schroeder et al., 2016, Kumar and Roy, 2018, Hu et al., 2021), in this chapter, Sentinel-2 Level-1C TOA products collected over 2019-2020 wildfires on the east coast of Australia and the west coast of the United States are downloaded from the Sentinel Open Hub (Potin et al., 2019) to build the dataset used for training and testing. This includes 135 Sentinel-2 Military Grid Reference System (MGRS) tiles over the states of New South Wales (NSW), Australia Capital Territory (ACT), and Victoria (VIC) covering areas of about 378,543 km^2 from October 18, 2019, to January 14, 2020, and 178 Sentinel-2 MGRS tiles over the state of California from July 28, 2020, to October 12, 2020, with total coverage of about 457,594 km^2. Figure 5.1(a) and Figure 5.1(b) illustrate the Sentinel-2 coverages in Australia and the United States, respectively.

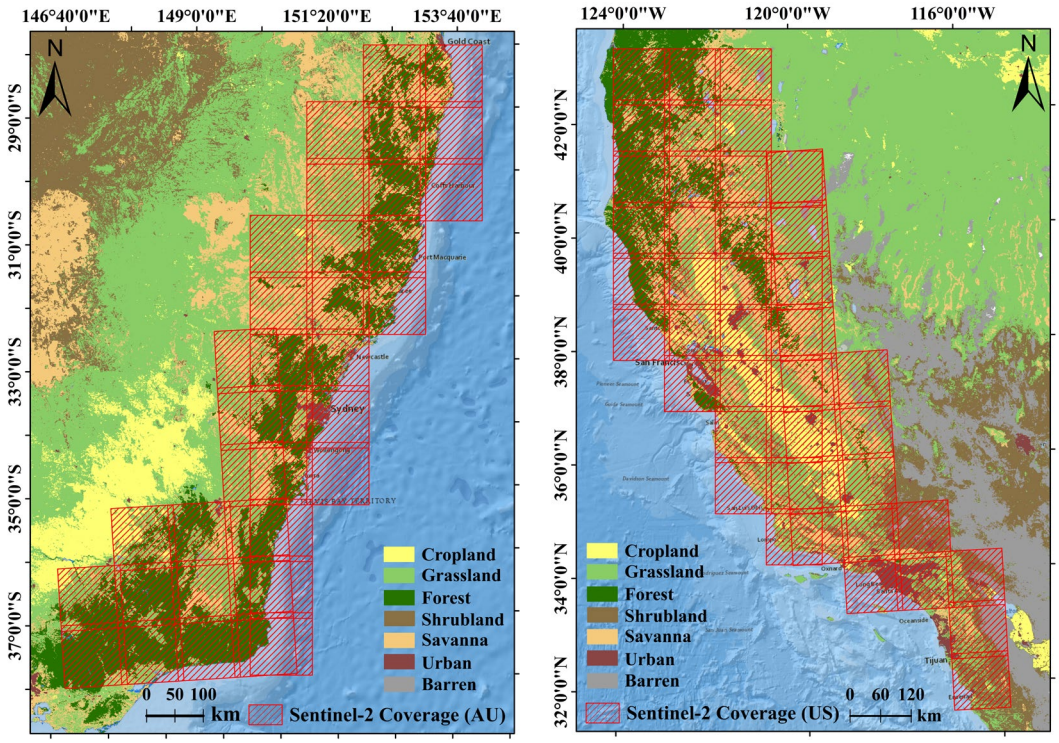


Figure 5.1 Sentinel-2 image coverage in Australia and the United States. Base map: land cover map derived from the 2019 Version 6 MODIS Land Cover Type (MCD12Q1) product (Sulla-Menashe et al., 2019)

The downloaded Sentinel visible bands are with spatial resolutions of 10 m and associated NIR and SWIR bands are with spatial resolutions of 20 m. Hence, visible bands are downsampled to the same spatial resolutions as NIR and SWIR bands before the dataset construction. Related annotations are generated by the synergy of the AFD-S2 method and the visual inspection of Sentinel-2 band combinations. In AFS-S2, a single criterion in Equation (5.1) is applied for the active fire detection in temperate conifer forests on the east coast of Australia, whereas multiple criteria in Equation (5.2) are used in the Mediterranean forests, woodlands, and scrubs on the west coast of the United States.

$$0.504 * \rho_{SWIR_2} - \rho_{Red} \geq a \tag{5.1}$$

$$0.734 * \rho_{SWIR_2} - \rho_{Red} \geq b,\ \rho_{SWIR_2} \geq c,\ \rho_{SWIR_1} \geq d\ or\ \ \rho_{SWIR_2} \geq 1.0 \qquad (5.2)$$

A crucial part is to determine a correct threshold to differentiate the active fire and the background pixels. In this study, a, b, c, and d are initialized as 0.198, 0.068, 0.355, and 0.475 as presented in the AFS-S2 method, where a fine-tuning is further applied referring to the visual inspection of B12-B11-B4 and B12-B11-B8A false-color combinations and a correction is also performed on the obvious mislabeling to reduce the OEs and CEs in the generated annotation map.

A dataset consisting of false-color patches and associated dense pixel annotations of 2 classes (active fire and background) is constructed by clipping and sifting Sentinel red and SWIR bands and the associated annotation map above (see section 5.3.1).

5.3 Methodology

5.3.1 Automated active fire detection framework

Details of the proposed automated active fire detection framework are shown in Figure 5.2, along with the data collection and preprocessing, deep-learning-based active fire detection, and product generation modules included in the yellow, green, and blue boxes, respectively.

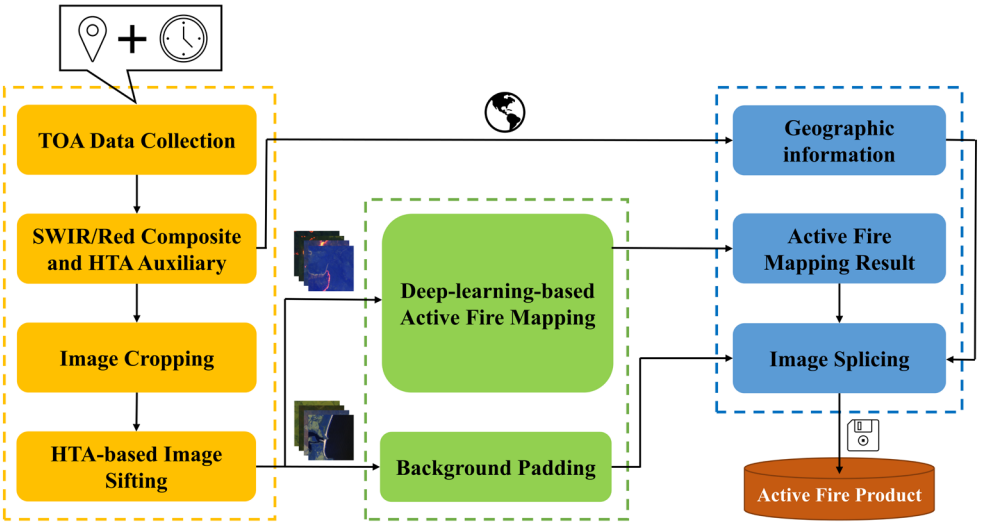


Figure 5.2 Illustration of the automated active fire detection framework implemented in this research

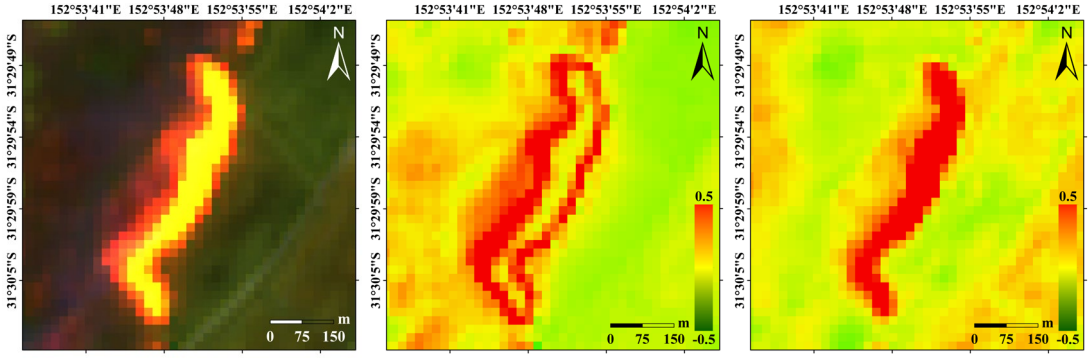
With the time and location information provided by users, Sentinel-2 TOA reflectance products can be automatically downloaded in the proposed framework through the API offered by ESA. Related SWIR and red bands are composited into false-color images with more salient active fires and normalized hotspot index (NHI) images are generated based on Equation 5.3 and Equation 5.4 (Marchese et al., 2019, Genzano et al., 2020) as auxiliaries in the upcoming sifting step to coarsely separate between patches with high-temperature anomalies (HTAs) (wildfires, industrial heat sources, volcanic activities, and land management fires) and the background paddings.

$$NHI_{SWIR} = \frac{SWIR_2 - SWIR_1}{SWIR_2 + SWIR_1} \tag{5.3}$$

$$NHI_{SWNIR} = \frac{SWIR_1 - NIR}{SWIR_1 + NIR} \tag{5.4}$$

Figure 5.3 presents an example of the false-color, NHI_{SWIR}, and NHI_{SWNIR} images in the active fire area. NHI_{SWIR} is more sensitive to high-temperature anomalies with

mid or low intensities, whereas folding values commence when further increase the intensity level. Compared with NHI_{SWIR}, NHI_{SWNIR} is more suitable to detect high-temperature anomalies with high intensity and can be used as compensation for the NHI_{SWIR} derived results. Criteria for the HTA detection are slightly looser than those in (Marchese et al., 2019, Genzano et al., 2020) to make sure cool smoldering fires are included in the generated HTA maps. These HTA maps are further fed into the image cropping step along with the false-color combinations.



(a) (b) (c)

Figure 5.3 Example of false-color, $\boldsymbol{NHI_{SWIR}}$ and $\boldsymbol{NHI_{SWNIR}}$ images in the active fire area

All images are cropped into small patches with identical sizes and sequential identities, which are further allocated to the active fire or the background dataset, respectively. Whether a false-color patch is active or not is determined by if there is a positive value in the related HTA patch. This dataset splitting strategy avoids redundant processing in the background area and thus reduces the time and memory cost of the framework. Besides that, the synergy of image cropping and dataset splitting can effectively reduce the imbalance between the active fire and background proportions in each false-color

patch. In a subsequent step, the active false-color patches are fed into the trained deep neural networks to generate the segmentation results while the background ones only deliver their identities to the padding images for further splicing. The segmentation results and associated background paddings are further projected back into the original frame to generated the final product in Geotiff format, where geographical information extracted from the original data is added into this spliced product.

5.3.2 Deep-learning-based active fire detection

Active fire pixels normally appear in bright orange in a SWIR-red false-color image due to the higher reflectance in SWIR bands and the lower reflectance in the red band (Miller and Thode, 2007, Schroeder et al., 2016, Gargiulo et al., 2019). After igniting, the active fire quickly spread around in the condition of high temperature and strong wind, which turns into local fire clusters of different sizes. The sharp contrast between active fire and background allows the fire detection task to be considered as a binary semantic segmentation of the false-color image, even though special attention needs to be paid to the unique behaviors of the active fire. For example, OEs are usually caused by small and cool fires, CEs are likely to occur in the soil-dominated pixels or highly reflective building rooftops, and proportions of the active fire and background pixels are unbalanced (Palacios et al., 2012).

We first considered active fire detection as a regular binary semantic segmentation and three state-of-the-art (SOTA) deep-learning-based models (DeepLabV3+, Gated-SCNN, and HRNet-V2) are employed to achieve this task. Introductions of these three models have been indicated in several studies (Chen et al., 2017, Sun et al., 2019,

Takikawa et al., 2019, Wang et al., 2020) and won't be repeated here. To better balance the OE and CE in the active fire detection results, a DCPA network is further proposed by connecting a dual-domain feature extractor to the existing semantic segmentation models using a channel-position attention block. Taking the HRNetV2 model as an example, the DCPA+HRNetV2 network is shown in Figure 5.4.

This proposal of the DCPA network is inspired by the principle of the visual inspection process. Modifications mainly focus on the first few layers of the deep neural network which is normally responsible for the extraction of simple features. During active fire detection, human attention is automatically placed on the salient regions that present a sharp contrast with the background. Besides that, boundaries of the active fires inside are outlined based on spatial distributions and interdependences of the associated fire pixels. Attention on the salient regions helps to locate anomalous regions and hence reduce the OEs in the sparse and cool fires, while statistics in the spatial distribution and interdependence provides a better distinction between the active fire area and the surrounding pixels, which help to reduce the CEs caused by the ambiguity between the hot soil and cooler active fire. Therefore, as illustrated in the yellow box of Figure 5.4, a frequency and a spatial feature extractor are employed in our network to obtain the dual-domain features used for segmentation.

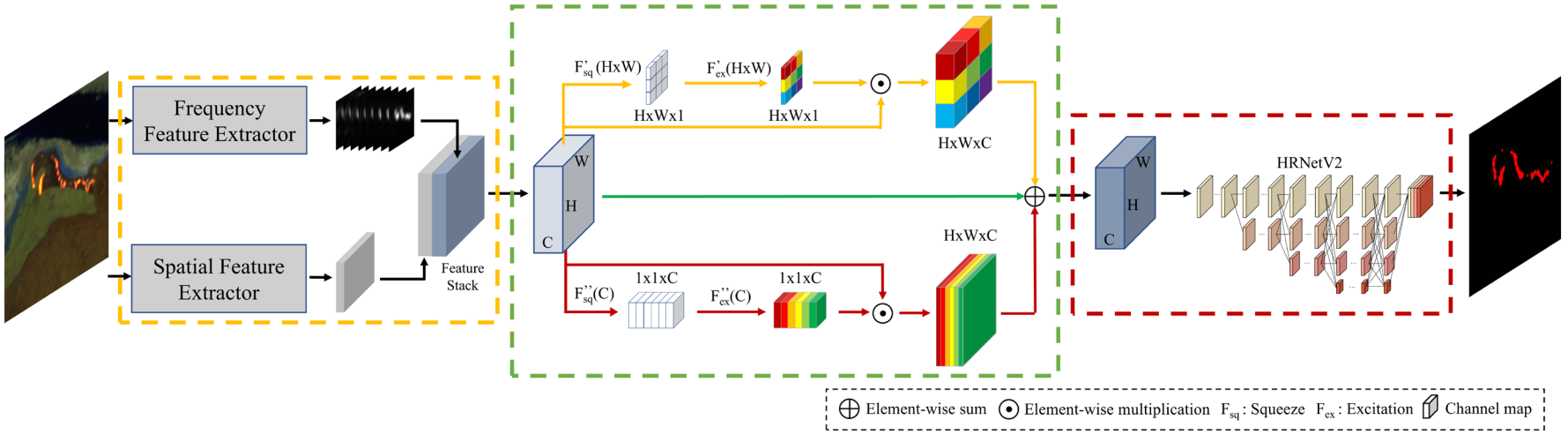


Figure 5.4 Structure of the proposed DCPA+HRNetV2 active fire detection network

On the one hand, a frequency feature extractor is constructed using the visual saliency paradigm in (Li et al., 2012) with the quaternion matrix f re-defined as below.

$$f(n,m) = \omega_0 P_0 + \omega_1 P_1 i + \omega_2 P_2 j + \omega_3 P_3 k \tag{5.5}$$

$$P_1 = \rho_{SWIR_2}, P_2 = \rho_{SWIR_1}, P_3 = \rho_{Red} \tag{5.6}$$

where ω_0, ω_1, ω_2, and ω_3 are weights of the feature matrices P_0, P_1, P_2, and P_3, and i, j, k are imaginary units in the quaternion form (i^2=j^2=k^2=ijk=-1, ij=k, ik=j, jk=i). The motion feature P_0 of a static input is equal to zero and has no contribution to the quaternion matrix f. Weights of spectral feature matrices are set as ω_1=0.2, ω_2=0.4, and ω_3=0.4. Hypercomplex Fourier Transform (HFT) of the quaternion matrix f can then be generated through

$$F_H[u,v] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu 2\pi (mv/M + nu/N)} f(n,m) = A(u,v) e^{\chi \Phi(u,v)} \tag{5.7}$$

where $\mu = (i+j+k)/\sqrt{3}$ is a unit pure quaternion, $A(u,v)$ and $\Phi(u,v)$ are the amplitude and phase spectrums of $F_H[u,v]$, and χ is a pure quaternion matrix as the aigenaxis spectrum (Ell and Sangwine, 2006).

It has been proven that texture-rich salient regions like active fire are more likely to be embedded in the background of the amplitude spectrum whereas the repeated and uniform patterns such as the spatial background are prone to appear as sharp spikes (Chen et al., 2015). Therefore, frequency features of different salient regions can be effectively enhanced by smoothing the amplitude spectrum with a series of Gaussian kernels $\{g_k\}$ at different scales (Li et al., 2012).

$$g_k(u,v) = \frac{1}{\sqrt{2\pi} 2^{k-1} \tau_0} e^{-(u^2+v^2)/(2^{2k-1}\tau_0^2)} \tag{5.8}$$

$$\Lambda_k(u,v) = (g_k(u,v) \star A)(u,v) \tag{5.9}$$

where, $\tau_0 = 0.5$, $k = 1, \dots, \lceil \log_2 min\{H, W\} \rceil + 1$. H and W are the height and width of the image. A sequence of saliency maps $\{s_k \in \mathbb{R}^{H \times W}\}$ can be further generated by the inversed HFT as part of the dual-domain feature stack.

$$s_k = \left\| \frac{1}{\sqrt{MN}} \sum_{v=0}^{M-1} \sum_{u=0}^{N-1} e^{\mu 2\pi(mv/M + nu/N)} \Lambda_k(u,v) e^{\chi \Phi(u,v)} \right\|^2 \tag{5.10}$$

On the other hand, two CNN layers with filter sizes of 3×3 and 1×1 are connected to a Resblock (He et al., 2016) as the spatial feature extractor. The outputs of this spatial feature extractor are the other part of the dual-domain feature stack $S \in \mathbb{R}^{H \times W \times C}$ with C representing the channel depth.

This dual-domain feature stack is further fused by a channel-position self-attention block built on channel-wise and position-wise squeeze-and-excitation (SE) networks (Hu et al., 2018) as presented in the green box of Figure 5.4. In the yellow stream, global channel information is first squeezed into a position descriptor through average pooling in the channel dimension.

$$x = F'_{sq}(S) = \frac{1}{C} \sum_{k=1}^{C} S(:,:,k) \tag{5.11}$$

An excitation operation is further performed on the squeezed output x to capture the position-wise dependencies, which is achieved by a simple gating mechanism with the ReLU activation function $\delta(\cdot)$.

$$x_1 = F'_{ex}(x) = \delta\left(W_1\left(\delta(W_2 x)\right)\right) \tag{5.12}$$

where, $W_1 \in \mathbb{R}^{H\times W\times\frac{H\times W}{d}}$ and $W_2 \in \mathbb{R}^{\frac{H\times W}{d}\times H\times W}$ are the weights of two convolution layers with d set as 256 for the best performance. The output of the yellow stream S' can be obtained by rescaling S with the excitation output x_1.

In the red stream, the average pooling is initially applied in the spatial dimension to obtain the squeezed channel descriptor. This is followed by an excitation operation to obtain the channel-wise dependencies.

$$y = F''_{sq}(S) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} S(i,j,:) \tag{5.13}$$

$$y_1 = F''_{ex}(y) = \sigma\left(W_3\left(\delta(W_4 y)\right)\right) \tag{5.14}$$

where, $\sigma(\cdot)$ is the sigmoid activation function, $W_3 \in \mathbb{R}^{C\times\frac{C}{r}}$ and $W_4 \in \mathbb{R}^{\frac{C}{r}\times C}$ are the weights of two fully connected layers with $r = 16$ for the best performance. Likewise, the output of this stream S'' is obtained by rescaling S with the excitation output y_1.

The original feature stack S is preserved through the green stream in the fusion process to tackle the gradient vanishing issue and to boost feature discriminability. Outputs of these three streams are then weighted summed to generate the segmentation input.

$$\tilde{S} = S + \alpha S' + \beta S'' \tag{5.15}$$

where α and β are initialized as 0 and gradually increase during the training process. The fused feature is then fed into a semantic segmentation model to get the active fire detection result.

5.3.3 Evaluation Metrics

Quantitative metrics that are commonly used in semantic segmentation and active fire detection such as CE, OE, and Intersection over Union (IoU) are calculated to evaluate the performance. All these metrics can be directly derived from the elements in the confusion matrix, which are, the numbers of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) pixels. In an active fire detection task, TP represents the correctly labeled active fire pixel, FP is the background pixel incorrectly labeled as active fire, TN stands for the correctly labeled background pixel, and FN is the active fire pixel incorrectly labeled as background.

Table 5.2 Metric calculation reference

Reference / Classified	Active Fire	Background
Active Fire	TP	FP
Background	FN	TN

Based on Table 5.2, the IoU, CE, and OE can be calculated as

$$IoU = \frac{TP}{TP + FP + FN}, \quad CE = \frac{FP}{TP + FP}, \quad OE = \frac{FN}{TP + FN} \tag{5.16}$$

5.3.4 Implementation details

The deep-learning-based active fire detection module with different segmentation architectures (DeepLabV3+, GS-CNN, HRNetV2, and DCPA+HRNetV2) are trained, validated, and tested on the constructed dataset, respectively. After the clipping and sifting, this dataset has a total of 37,016 active fire samples with 21,342 of them from Sentinel-2 images collected over Australia (AU) and 15,674 of them from the United

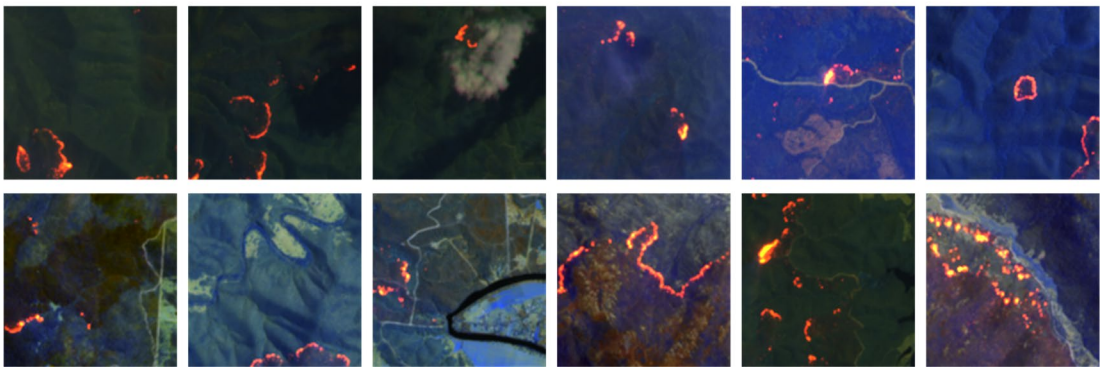
States (US). Samples in the dataset are randomly divided into three subsets with 60% for training, 10% for validation, and 30% for testing. Each sample in the constructed dataset includes a false-color patch in size of 128×128×3 and a dense pixel annotation with two classes in size of 128×128. During the training process, inputs are augmented by random flipping and rotating. An Adam optimizer (Paszke et al., 2019) is initialized with a base learning rate of 0.01 and a weight decay of 0.0005. The poly learning rate policy with the power of 0.9 is employed to drop the learning rate during the training process. All experiments are implemented in the PyTorch framework on two NVIDIA GeForce RTX 2080Ti GPUs with a batch size of 16.

5.4 Experiment Result

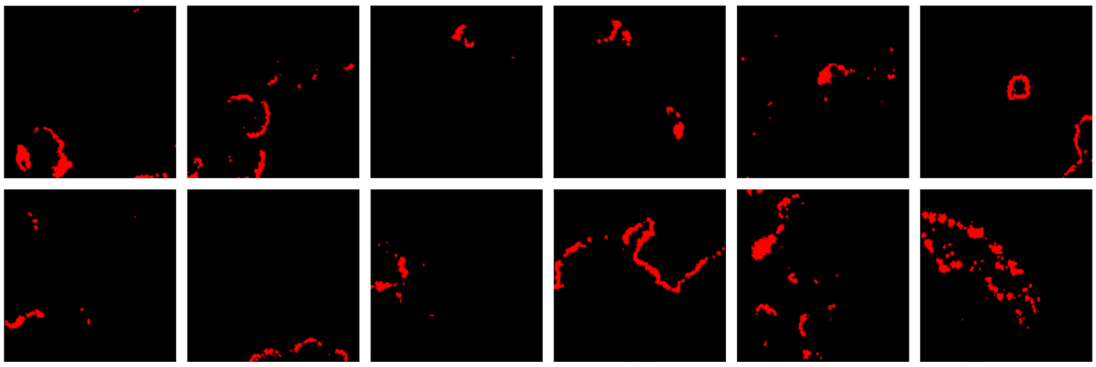
5.4.1 Deep-learning-based active fire detection

By treating the active fire detection as a regular binary semantic segmentation process, results derived from three SOTA deep-learning-based semantic segmentation models (DeepLabV3+, Gated-SCNN, and HRNet-V2) are first shown in Figure **5.5** and Figure **5.6** with related accuracy metrics listed in Table **5.3**.

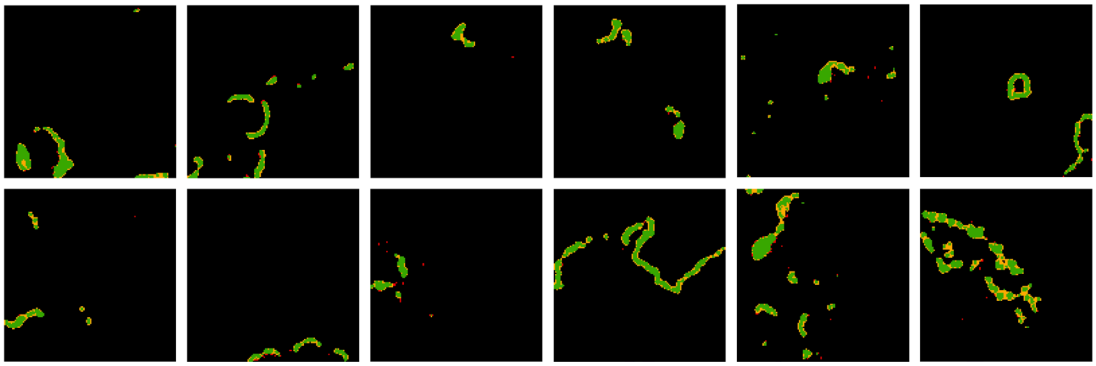
East Coast of Australia



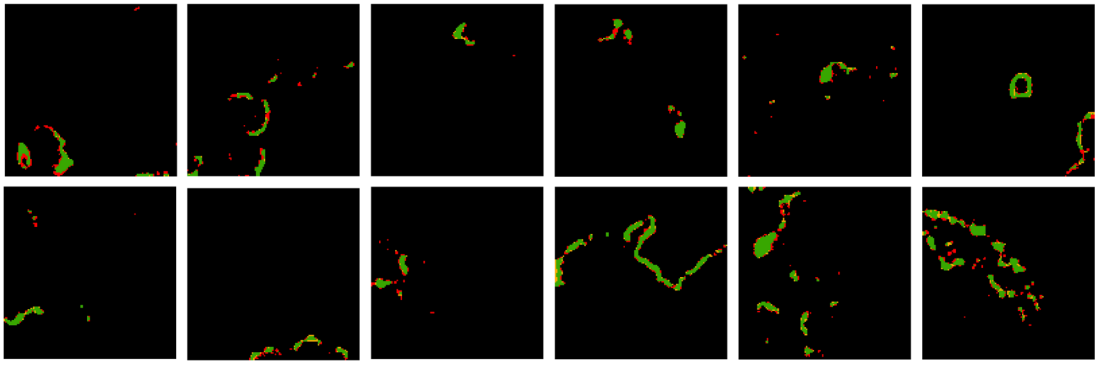
(a) Ground Truth



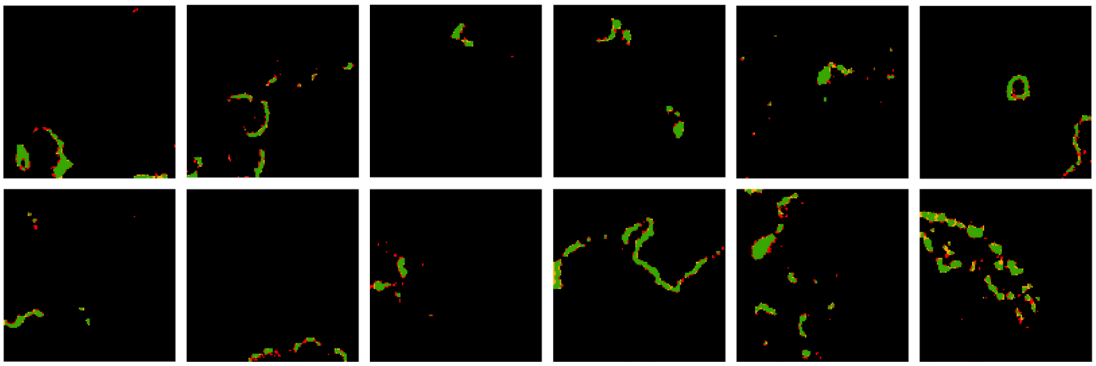
(b) DeepLabV3+



(c) Gated-SCNN



(d) HRNetV2



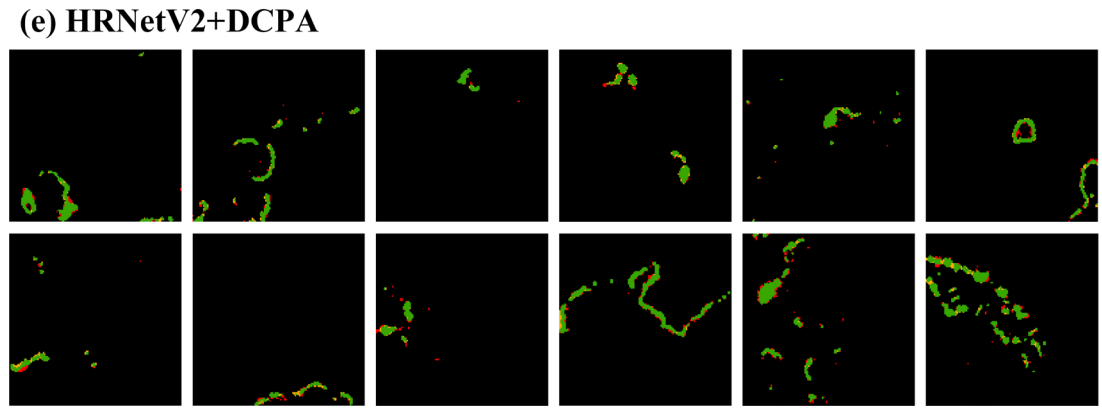
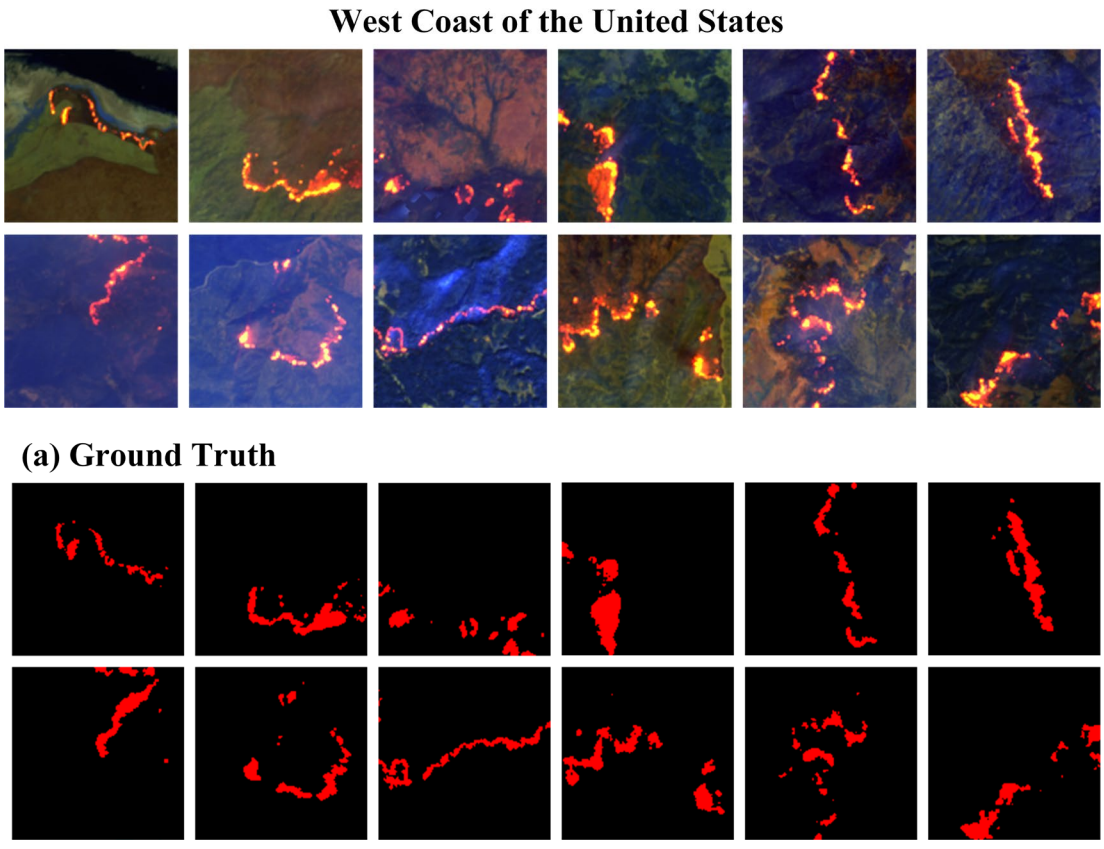
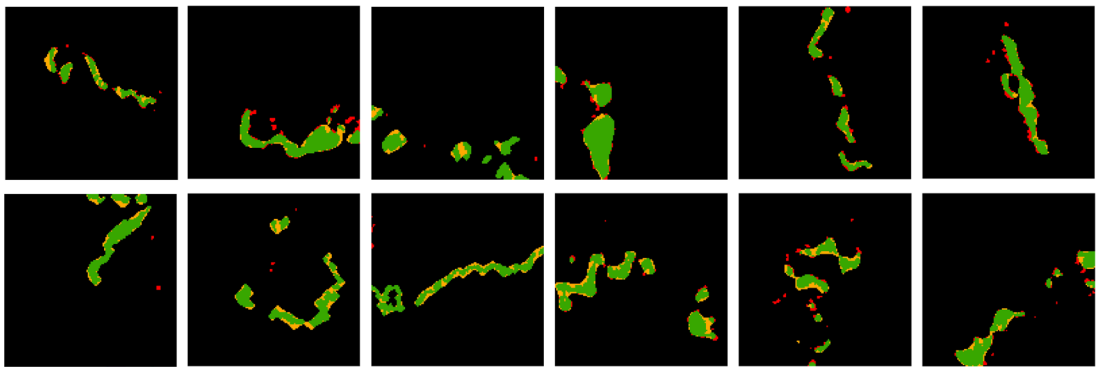


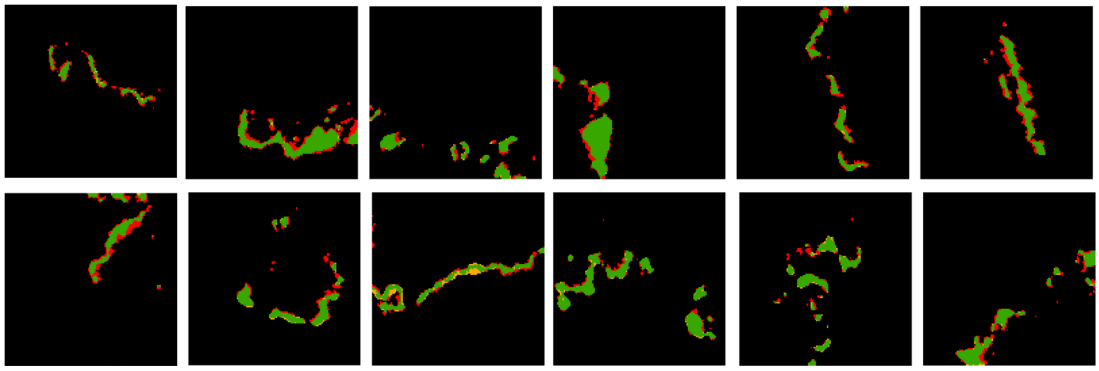
Figure 5.5 Active fire samples (false-color composite and binary annotation) and the associated segmentation results on the east coast of Australia (Green: TP; Red: FN; Yellow: FP): (a) Ground truth; (b) DeepLabV3+; (c) Gated-SCNN; (d) HRNetV2; (e) DCPA+HRNetV2.



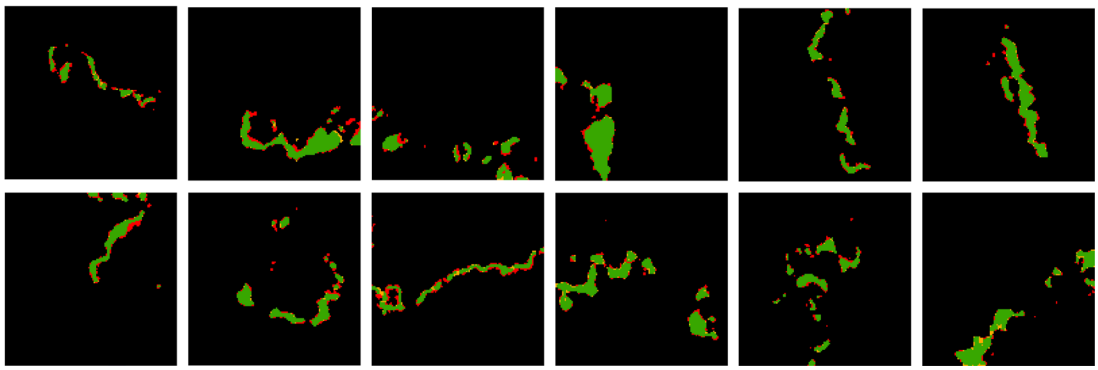
(b) DeepLabV3+



(c) Gated-SCNN



(d) HRNetV2



(e) HRNetV2+DCPA

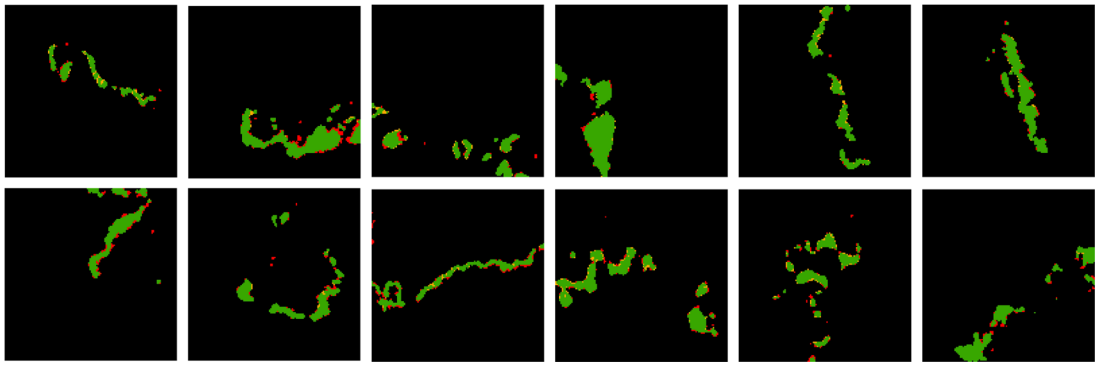


Figure 5.6 Active fire samples (false-color composite and binary annotation) and the associated segmentation results on the west coast of the United States (Green: TP; Red: FN; Yellow: FP), (a) Ground truth; (b) DeepLabV3+; (c) Gated-SCNN; (d) HRNetV2; (e) DCPA+HRNetV2

Table 5.3 Accuracy metrics of the active fire detection result

Method	Backbone	OE [a]	CE [a]	IoU [a]	OE [b]	CE [b]	IoU [b]
DeepLabV3+	Xception-71	7.6%	28.1%	67.8%	9.1%	25.3%	69.5%
Gated-SCNN	WideResNet-38	19.6%	16.0%	69.7%	19.9%	13.2%	71.4%
HRNetV2	HRNetV2-W48	18.5%	14.2%	71.8%	19.0%	11.5%	73.2%
DCPA+HRNetV2	HRNetV2-W48	17.3%	13.1%	73.4%	17.4%	9.2%	76.2%

[a] East coast of Australia (AU); [b] West coast of the United States (US)

Active fire detection results above indicate that Gated-SCNN and HRNetV2 models focus more on the spatial details in the segmentation result by paying more attention to the description of the boundaries and edges or preserving high spatial resolution representations throughout the propagation, which significantly reduces the CEs in the DeepLabV3+ result. The performance of HRNetV2 is slightly better than the Gated-SCNN method. However, too much attention to the description of the active fire and background boundaries will lead to the neglect of cooler fires around and thus higher OEs in related results. It is also evidenced in the results above that the introduction of the DCPA network helps to locate sparse and isolated active fires with only a few pixels and balance the OEs and CEs for a better IoU thanks to the dual-domain feature extraction and the channel-position-wise attention employed.

5.4.2 Automated active fire detection framework

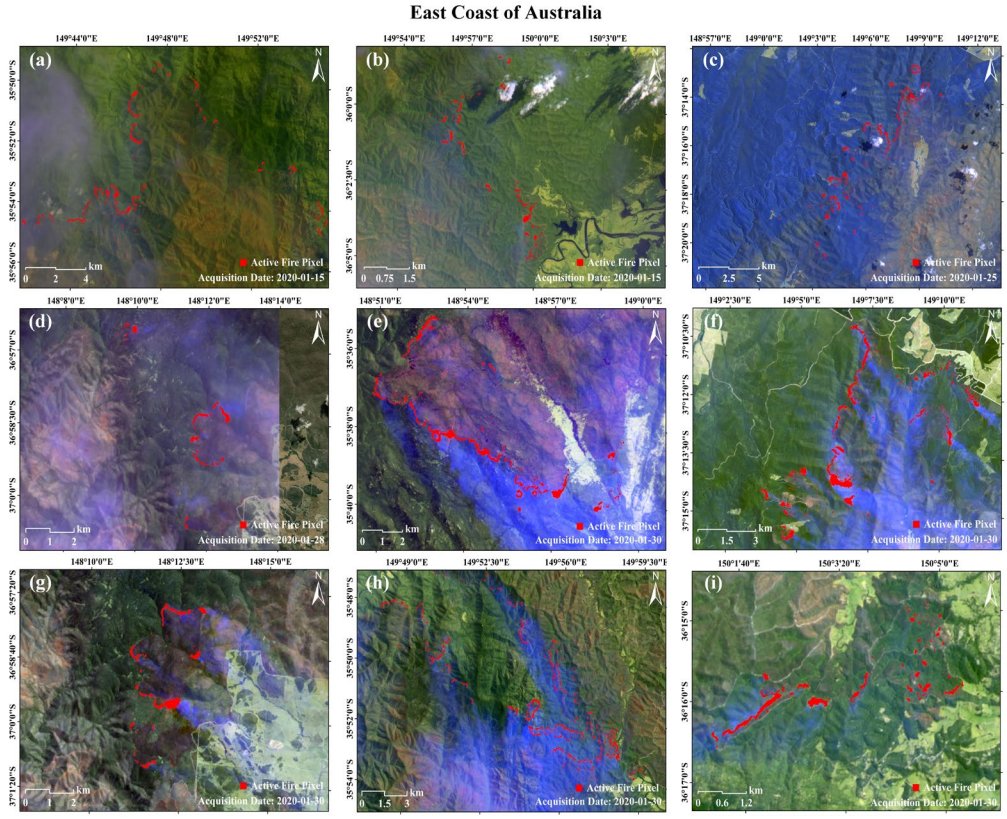
The trained DCPA+HRNetV2 model is assembled back into the automated framework following by a further evaluation of the detection accuracy and processing efficiency. Binary segmentation results are automatically generated in Geotiff format by inputting the locations of the bounding box and the time of the fire event. With the details listed in Table **5**.**4**, tests are performed on the east coast of Australia and the west coast of the United States, respectively. Final products are further zoomed into areas with dense active fire pixels and laid over the false-color base image in ArcMap as shown in Figure **5**.**7**.

Table 5.4 Test of the automated active fire detection framework

East Coast of Australia				
Acquisition Date	2020-01-15	2020-01-25	2020-01-28	2020-01-30
Top Left [a]	-35.5, 148.7	-37.3, 148.5	-36.3, 147.6	-34.8, 148.7
Top Right [a]	-35.5, 150.2	-37.3, 149.2	-36.3, 148.7	-34.8, 150.6
Bottom Left [a]	-36.8, 148.7	-37.8, 148.5	-36.9, 147.6	-37.6, 148.7
Bottom Right [a]	-36.8, 150.2	-37.8, 149.2	-36.9, 148.7	-37.6, 150.6
IoU	71.4%	71.1%	70.2%	68.7%
Time Cost	358s	105s	133s	839s
West Coast of the United States				
Acquisition Date	2020-10-13	2020-10-18	2020-10-28	2020-10-29
Top Left [a]	37.2, -119.4	37.6, -118.9	37.6, -119.4	41.9, -123.7
Top Right [a]	37.2, -118.6	37.6, -118.2	37.6, -118.9	41.9, -123.3
Bottom Left [a]	36.7, -119.4	37.3, -118.9	37.3, -119.4	41.0, -123.7
Bottom Right [a]	36.7, -118.6	37.3, -118.2	37.6, -118.9	41.0, -123.3
IoU	72.6%	71.3%	73.0%	70.8%
Time Cost	58s	51s	46s	132s

[a] Location of the corner pixels (Latitude, Longitude) in UTM projection and WGS84 geodetic system

As indicated in Table 5.4, the automated active fire detection framework achieved average IoUs of 70.4% over Australia and 71.9% over the United States. The most time-consuming parts of the framework are the data collection and active fire detection steps. Results indicated that, for Sentinel-2 input with total coverage of around 12,000 km^2, it takes about 350 seconds to get the final output, where the active fire detection process takes up to 140 seconds. This underlines the necessity of the active fire and background sample splitting in the preprocessing module. All the operations above are undertaken in configurations of Intel Core i7-9700k processer with a base frequency of 3.60GHZ, NVIDIA GeForce RTX 2080Ti GPU, and a system memory of 64 GB.



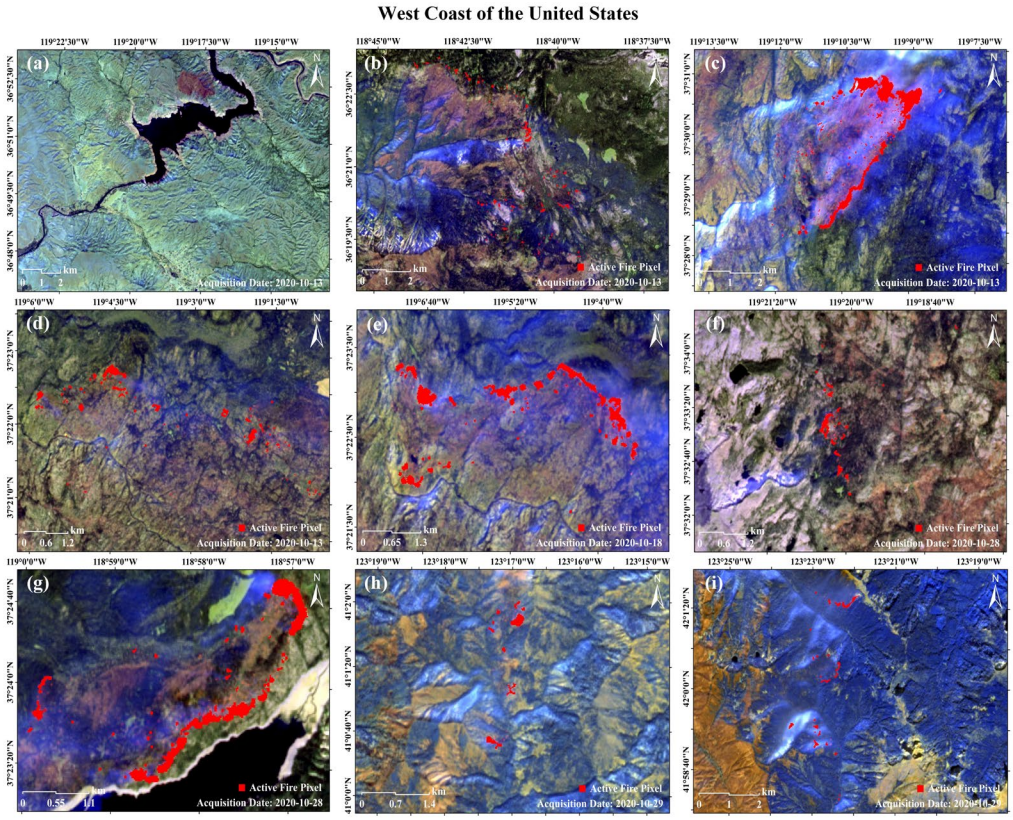
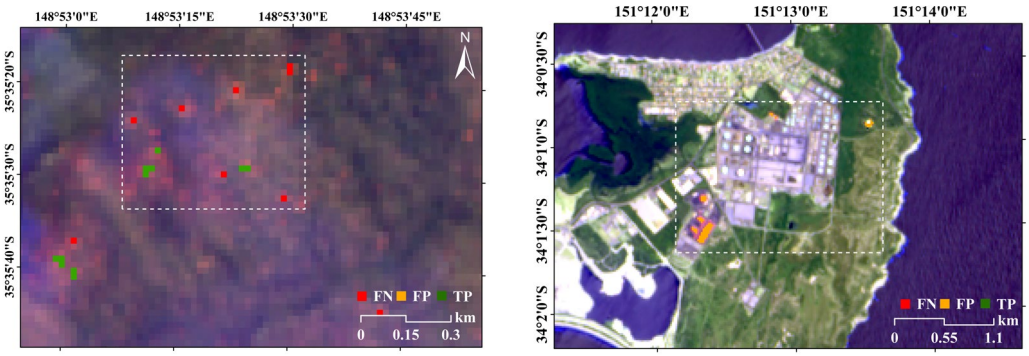


Figure 5.7 Test results of the automated framework in Australia and the United States

5.5 Discussion

In this section, discussions are carried out on the error sources in the generated active fire products including annotation method, data preprocessing step, data source, and segmentation algorithm. Compared with semantic segmentation applications in other fields, active fire detection can be simplified as a binary segmentation task, though of increased difficulty as it deals with a large volume of small fire objects. Annotations in the constructed dataset are derived from the synergy of the AFD-S2 method and the visual inspection of Sentinel-2 composites. Although the time and labor involved are significantly reduced by this semi-supervised operation, there are still some omitted

small and cool fires with lower radiance reflectance or mislabeled soil-dominated pixels and highly reflective building rooftops derived from the AFD-S2 method not corrected by the visual inspection, which may mislead deep-learning-based models to learn wrong recognition patterns in the active fire segmentation. Besides that, though criteria for the HTA detection are slightly loosened during the splitting of the active fire patch and the background padding, there are still some cool smoldering fires that are not included in the generated HTA map, and thus misclassified into the background dataset. Another source of error is intrinsic in SWIR/Red-based active fire detection, which mainly uses the sharp contrast between the high-temperature active fire and the background in the false-color image. However, the existence of heat radiation results in ambiguities between classes affecting the classification accuracy of pixels along the edges, and consequently, introducing errors in the segmentation results. As a way of example, Figure 5.8 zooms over areas with small and cool fires and highly reflective building rooftops to present the OEs and CEs in the results.



(a) (b)

Figure 5.8 Examples of OEs and CEs in the active fire detection result: (a) OE; (b) CE

Despite the aforementioned sources of error, the proposed framework outperforms the current fire monitoring systems described in Section 5.1 in terms of delivering the

spatial details of the active fire, which can help to calibrate or/and improve the current low-resolution results. The detection accuracy of this framework can be further improved through developing the associated deep-learning-based algorithms or using the training datasets with higher spatial resolution.

5.6 Conclusion

This chapter develops an automated active fire detection framework including a data collection and preprocessing, a deep-learning-based active fire detection, and a final product generation module as a prototypical function for the future fire monitoring system. The active fire detection module is the basic part of the proposed framework developed on a specifically designed DCPA+HRNetV2 network, which is trained on the dataset constructed using the SWIR and red bands in Sentinel-2 Level-2C products. Results indicated this DCPA and HRNetV2 combination outperformed DeepLabV3, GS-CNN, and HRNetV2 segmentation models on the active fire detection, and the automated framework can effectively deliver detection results with an average IoU larger than 70%. Though in the current state, the refresh rate of this framework is just 5 days and it is only suitable for inputs of Sentinel-2 acquisitions over the east coast of Australia and the west coast of the United States, the spatial resolution (20m) of the outputs is much higher than the results derived from the current fire systems. With the launch of more and more high-resolution and super-spectral sensors onboard remote sensing satellites, high-quality and near-real-time earth observation will be achieved in the future. At that time, an automated framework like this will offer a powerful tool and cost-efficient resource in support of the governments and fire service agencies that need timely, optimized firefighting plans.

Chapter 6 Deep-learning-based Burned Area Mapping using the Synergy of Sentinel-1 and Sentinel-2 Data

This chapter intends to present the advantages of different satellite remote sensing on the application of burned area (BA) mapping and further investigates if the existing coarse BA mapping products derived from the exclusive use of the SR data can be further improved by the synergy of SR, BS, and/or COR features with higher spatial resolutions. This achievement is largely attributed to prior studies on the physical basis of mapping BA from satellite observations in the solar, mid-infrared, thermal infrared, and microwave domain and the associated BA mapping algorithms and BA products derived (Chuvieco et al., 2019). Experiments are carried out on the SR, BS, and COR features originated from Sentinel-1 and Sentinel-2 constellations, where high-spatial-resolution PlanetScope imagery is used for annotation and validation. A Siamese Self-Attention (SSA) classification strategy is proposed to explore the latent relationships of multi-source attributes and improve their synergies on BA mapping. A multi-source dataset is constructed at the object level for training and testing. Associated results are analyzed by different test sites, feature sources, and classification strategies to appraise the improvements achieved. Related BA products generated in this study can be used as a powerful data resource for governments and fire management agencies to inform post-fire planning and policy.

I have published this work in Remote Sensing of Environment (Zhang et al., 2021a).

Zhang Q, Ge L, Zhang R, et al. Deep-learning-based burned area mapping using the synergy of Sentinel-1&2 data. *Remote Sensing of Environment* 2021; 264, 112575.

6.1 Introduction

Around 350 million hectares of land are affected by wildfires annually due to human activities and/or natural events (Syphard and Keeley, 2015, Schroeder et al., 2016). Prescribed burnings (Furlaud et al., 2018) provide benefits to biodiversity through promoting new plant growth and clearing the natural waste material. However, an uncontrollable fire can rapidly spread over a large scale and turn into a huge threat to the society and environment. The increasing frequency and severity of wildfire events over the past decades urge local governments and fire management agencies to call for reliable, timely, and cost-efficient fire mapping systems (Thompson et al., 2018, Lagadrilliere et al., 2019). Since the invention of the first Infrared camera (Mangeon et al., 2016), many in-situ and airborne tools have been developed for BA mapping. However, the inconsistency in data collection and mapping methods used makes it hard to analyze the performance at regional, continental or global scales (Chuvieco et al., 2019). Compared with the ground-based and airborne or unmanned aerial vehicle (UAV)-based methods, satellite remote sensing provides large-scale and cost-efficient earth observations in short revisit times, which is hence more suitable for developing a systematic and global fire mapping system (Colson et al., 2018).

Prior research (Chuvieco et al., 2008, Hantson et al., 2013, Chuvieco et al., 2019) has evidenced the capabilities of multispectral remote sensing data for hotspot detection and BA mapping. In particular, the high temporal and coarse spatial resolution images

acquired by MODIS have been widely applied to the generation of global BA products (MCD45 (Roy et al., 2008), MCD64 (Giglio et al., 2018), Fire_cci v5.0 (Chuvieco et al., 2018) and Fire_cci v5.1 (Lizundia-Loiola et al., 2020)) in the past years. However, high discrepancies are observed in these products over small and fragmented burns which introduce uncertainties into the BA mapping result (Smith et al., 2007, Chang and Song, 2009, Ramo et al., 2021). Accuracies of these products are further improved using the medium-spatial-resolution images acquired by the Landsat ETM+ and OLI, and Sentinel-2 MSI (Bastarrika et al., 2011, Stroppiana et al., 2012, Warner et al., 2017, Roy et al., 2019).

Subsequent studies prove that the exclusive use of multispectral data for BA mapping is limited by solar illumination, weather condition, and spectral confusion among BA, dark soil, and cast shadow (Stroppiana et al., 2015, Fassnacht et al., 2021). In this regard, the potential of SAR and InSAR techniques on the BA mapping are further explored since the active microwave has better penetration through the cloud coverage and less dependency on the weather condition (Moreira et al., 2013). However, studies indicated that the BS feature over BA depends on many factors including polarization (HH, HV, VH, and VV), frequency (X, C, and L), topography, and soil moisture (Tanase et al., 2010a, Tanase et al., 2013), and lower accuracies are observed in SAR-based BA mapping results than the multispectral-based ones (Tanase et al., 2020). Properties of the repeat-pass COR over BAs have also been investigated at different bands and polarizations (Tanase et al., 2010b) indicating a strong connection with the burned severity in the condition of a stable and dry environment. Synergies of active and passive remote sensing are further explored. Marginal improvements have been

achieved using the synergy of BS and SR features over cloud-free areas (Belenguer-Plomer et al., 2021). However, whether improvements in BA accuracy can be achieved through the synergy of SR, BS, and/or COR data have not yet been investigated.

In recent years, local-adaptive algorithms including support vector machine (SVM) (Cao et al., 2009), random forest (RF) (Ramo and Chuvieco, 2017), and linear regression (LR) (Huang and Pan, 2003) have been widely applied in the global BA mapping, taking advantage of the high processing efficiency and free access to the medium resolution imagery (Chuvieco et al., 2019). These algorithms discriminate the burned and unburned area by maximizing the inter-class and minimizing the intra-class variation based on a set of attributes. Associated results indicated that the feature selection is a critical pre-step when adding extra information to the classification process (Dragozi et al., 2014). However, when a large number of features are involved, feature selection will be a quite time-consuming process and latent relations between features of different sources are hard to define. With the advances in deep-learning-based algorithms, high-level features can be automatedly extracted by CNN (Lo et al., 1995), RNN (Mikolov et al., 2010), and/or self-attention mechanism (Vaswani et al., 2017), etc.

Therefore, this chapter tends to explore whether details in the existing BA product derived from the coarse spatial resolution data can be further improved through the synergy of BS, COR, and/or SR data with higher spatial resolutions. To achieve this, an object dataset is constructed over 12 fire perimeters containing a certain proportion of burned and unburned samples, and a local-adaptive classification method is further

proposed for the automated feature extraction and BA discrimination to ensure this improvement is potentially able to be applied to regional or/and global scales.

6.2 Test sites and datasets

6.2.1 Test sites

According to statistics from the California Department of Forestry and Fire Protection and the National Interagency Fire Center, there were 8,527 wildfires occurred in California in 2018 with an affected area of 1,893,913 acres and economic loss of 12 billion USD (CALFIRE, 2018). This makes it the deadliest and most destructive fire season ever recorded in California history. In 2019, over 7,860 fires were recorded affecting a ground area of about 259,823 acres. In this study, 12 fire events with the largest BAs that occurred in California from 2018 to 2019 were selected as the test sites. Figure 6.1 and Figure 6.2 present the locations of these selected sites and the timelines of the associated fire events, respectively. The proportions of different land covers over each site are summarized from the 2018 Version 6 MODIS Land Cover Type (MCD12Q1) product (Sulla-Menashe and Friedl, 2018) as listed in Table 6.1.

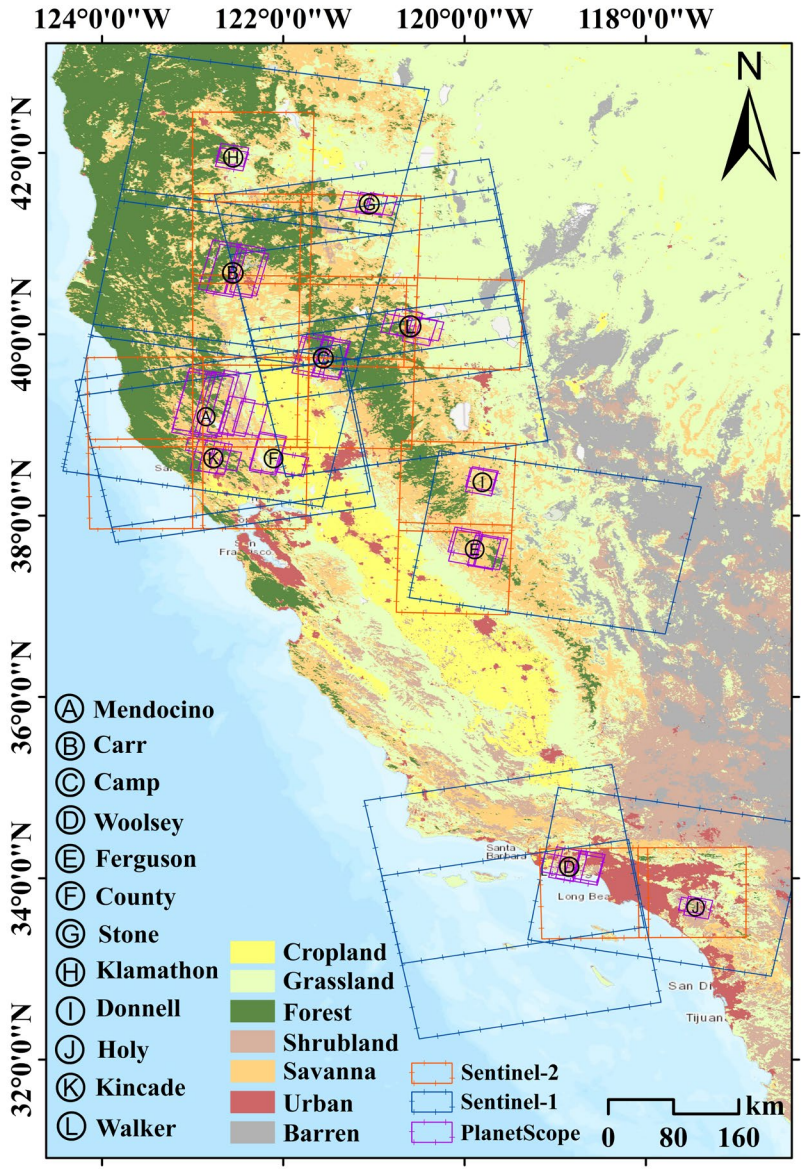


Figure 6.1 Locations of the test sites along with an indication of the satellite image coverage (Sentinel-1, Sentinel-2, and PlanetScope). Base map: land cover map derived from MCD12Q1 product.

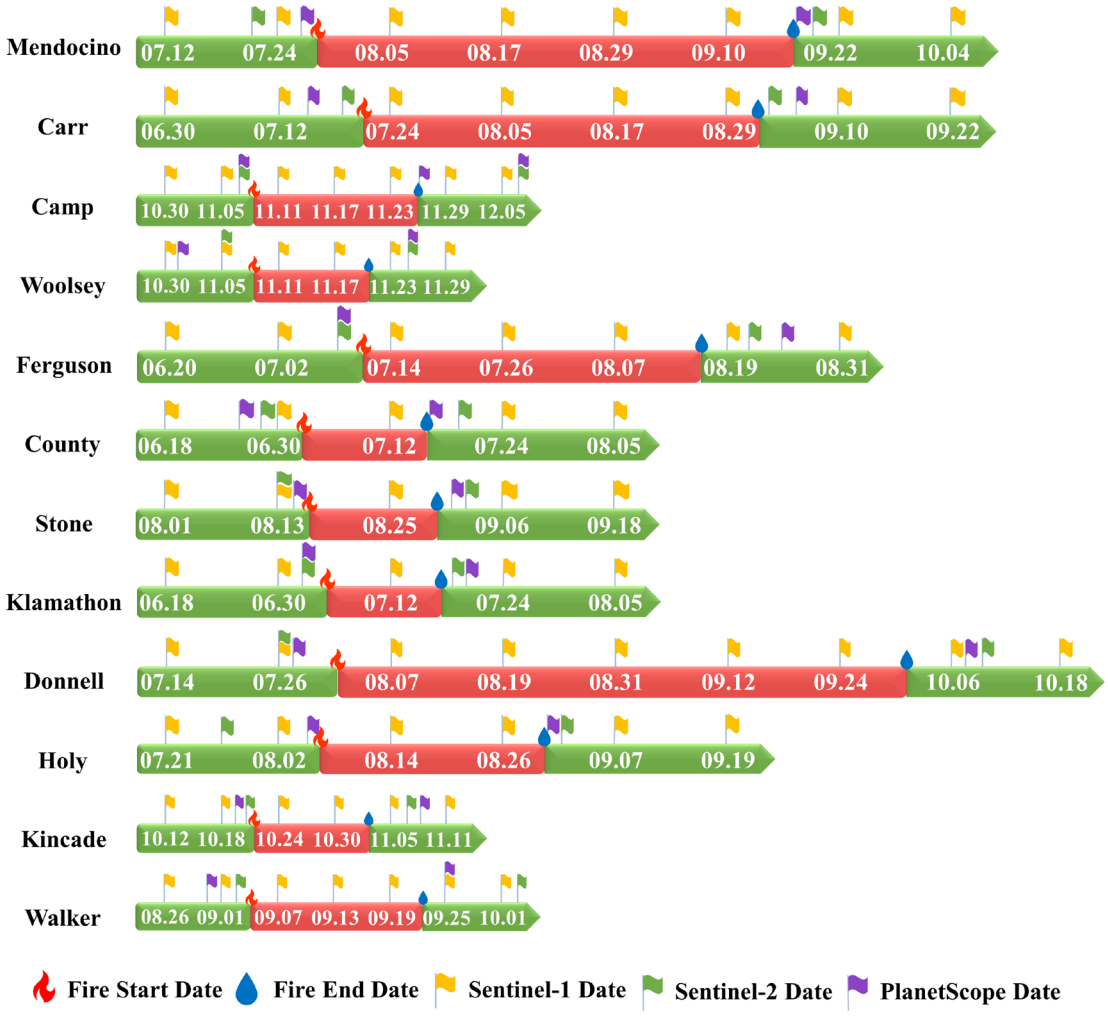


Figure 6.2 Timelines of the fire events.

Table 6.1 List of top 12 wildfire events in California (2018 – 2019)

Name	Location	Start Date	Containment Date
Mendocino	39.24°N, 123.10°W	July 27, 2018	September 18, 2018
Carr	40.65°N, 122.62°W	July 23, 2018	August 30, 2018
Camp	39.81°N, 121.44°W	November 8, 2018	November 25, 2018
Woolsey	34.24°N, 118.70°W	November 8, 2018	November 22, 2018
Ferguson	37.65°N, 119.88°W	July 13, 2018	August 18, 2018
County	38.81°N, 122.18°W	June 30, 2018	July 14, 2018
Stone	41.41°N, 121.06°W	August 15, 2018	August 29, 2018
Klamathon	41.89°N, 122.53°W	July 5, 2018	July 16, 2018
Donnell	38°21′N 119°56′W	August 1, 2018	October 1, 2018
Holy	33.74°N 117.52°W	August 6, 2018	August 30, 2018
Kincade	38.79°N, 122.78°W	October 23, 2019	November 06, 2019
Walker	40.05°N, 120.67°W	September 04, 2019	September 25,2019

Table 6.2 Proportions of different land covers over each site

Name	The proportion of different land cover over BA					Reported
	Forest	Shrubland	Savanna	Grassland	Cropland	BA (Acres)
Mendocino	28.5%	11.2%	42.5%	12.6%	0.1%	459,123
Carr	50.6%	0.1%	39.3%	3.4%	0.2%	229,651
Camp	40.8%	0.1%	35.4%	15.7%	5.6%	153,336
Woolsey	0%	28.8%	10.3%	22.1%	0.3%	96,949
Ferguson	17.9%	2.4%	74.9%	4.7%	0%	96,901
County	1.4%	1.3%	28.0%	46.8%	15.3%	90,288
Stone	1.8%	0%	38.2%	59.8%	0.1%	39,387
Klamathon	5.1%	13.0%	18.7%	62.9%	0%	38,008
Donnell	5.5%	0.1%	74.3%	18.9%	0%	36,450
Holy	8.1%	49.8%	6.4%	25.5%	0%	23,136
Kincade	21.0%	2.7%	70.9%	2.4%	0.1%	77,758
Walker	23.0%	0%	52.7%	23.4%	0%	54,608

6.2.2 Datasets

Spaceborne acquisitions over each site include one pre-fire and one post-fire SR image acquired by Sentinel-2, the time-series (TS) dual polarimetric SLC and Ground Range Detected (GRD) product obtained by Sentinel-1, and the pre- and post-fire visible and NIR images from PlanetScope. Sentinel acquisitions are used to provide multi-source attributes for BA discrimination, while high-resolution PlanetScope images are used to generate associated labels for training and validation. Related details of the original acquisitions are listed in Table 6.3.

Table 6.3 Details of the downloaded multi-sensor acquisitions

Sentinel-1A/B		**Sentinel-2A/B**		**PlanetScope**	
Mode	IW [a]	**Wavelength**	442-2200 nm	**Wavelength**	455-860 nm
Polarization	VH+VV	**Band**	13 [e]	**Band**	4
Spatial Res [b]	5m×20m	**Spatial Res**	10-60 m	**Spatial Res**	3 m
Temp [c] Res	6 days	**Temp Res**	5 days	**Temp Res**	1 day
Product	Level-1 [d]	**Product**	Level-1C	**Product**	Level-3A

[a] Interferometric Wide Swath; [b] Resolution; [c] Temporal; [d] SLC+GRD; [e] The research uses bands 2, 3, 4, 5, 6, 7, 8A, 11, and 12 (visible, NIR, and SWIR)

A total of 140 Sentinel-1 SAR (SLC and GRD) and 24 Sentinel-2 products (see Figure **6.1**) were downloaded from the Sentinel Open Hub (Potin et al., 2019) based on the timeline of the events. Since the minimum repeat cycle is 6 days for the two-satellite constellation, time intervals of the downloaded time-series Sentinel-1 data are set as 6 or 12 days for different fire events. For instance, time intervals are set as 6 days when fire periods are less than a month (Camp, Woolsey, Kincade, Walker) and set as 12 days when fire periods are longer than months (Mendocino, Carr, Ferguson, Donnell). This improves the efficiency of SAR/InSAR processing. Besides that, some of the

time intervals are set as 12 days though the fire periods are short than a month because acquisitions with an exact repeat cycle of 6 days are not available in these areas. Time intervals longer than 12 days are not considered in this research because severe temporal decorrelations in the interferometric coherence will increase difficulties in discriminating fire-affected areas. High-resolution PlanetScope data captured on the pre- and post-fire dates were downloaded through Planet Explorer (Planet, 2020).

6.3 Methodology

6.3.1 Object Dataset Construction

The proposed BA mapping algorithm includes two parts, object dataset construction, and deep-learning-based classification. Figure **6.3** presents the whole process of object dataset construction. The original Sentinel acquisitions are preprocessed, stacked, and segmented into small objects with multi-source attributes, which are further assigned as burned or unburned referring to the PlanetScope data.

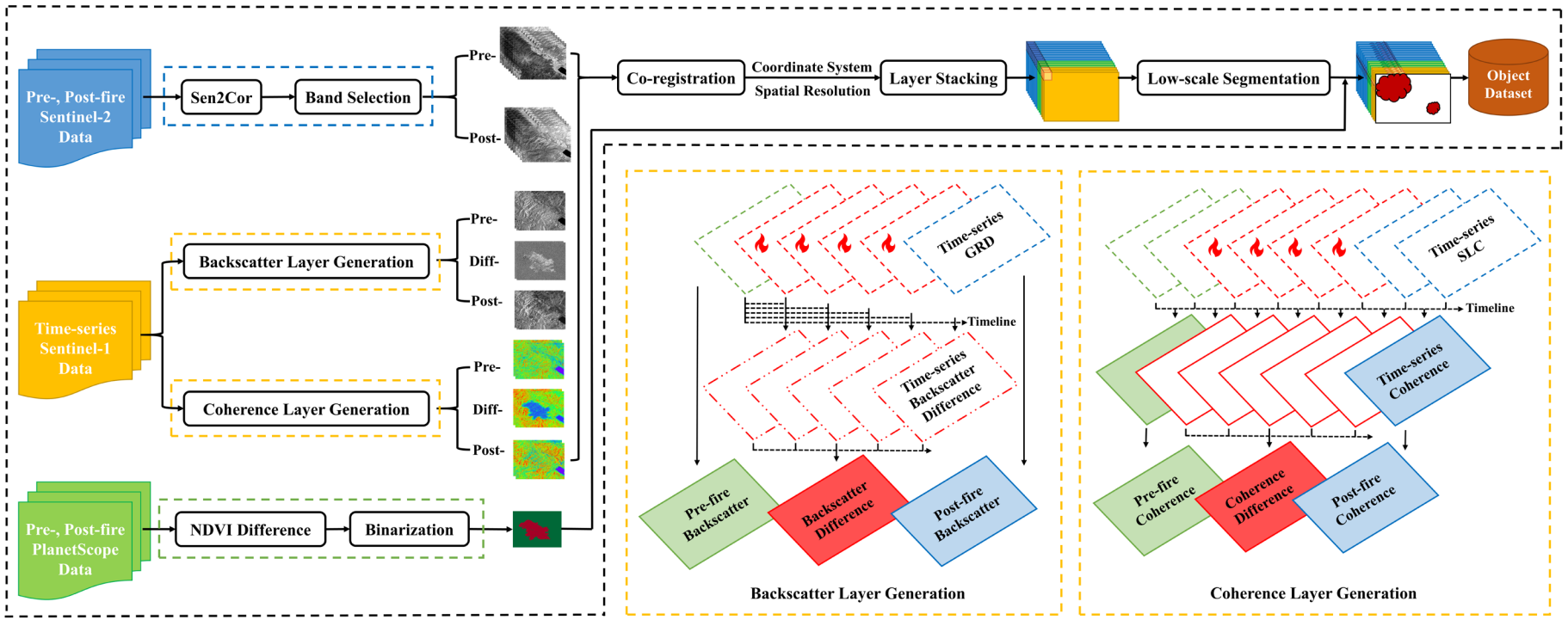


Figure 6.3 Object dataset construction method. Data processing in the black box: Blue-Sentinel-2, Yellow-Sentinel-1, Green-PlanetScope; Sentinel-1 footprint in the yellow boxes: Green-Pre-fire, Red-Wildfire, Blue-Post-fire.

The downloaded Sentinel-2 Level-1C Top-Of-Atmosphere (TOA) reflectances are in the sub-pixel registration to the UTM projection and WGS84 geodetic system. Only Sentinel-2 images without cloud coverage over BAs are selected to avoid the loss of ground information. Atmospheric, terrain and cirrus corrections are performed using the sen2cor module within SNAP to generate the Level-2A products with Bottom of Atmosphere (BOA) reflectances (Louis et al., 2016, Main-Knorn et al., 2017, Shao et al., 2019b). Instead of using all the spectral information, the costal aerosol, water vapor, and cirrus bands are disposed during the band selection because of their lower spatial resolutions. The strategy of using time-series multispectral imagery to enhance the pre- and post-fire BA mapping results is not employed here. This is because, during a wildfire event, images collected in visible and infrared bands are prone to be affected by the thick smoke from hotspots, which may aggravate errors during the classification.

The downloaded SAR data are time-series Level-1 SLC data in slant-range geometry and time-series GRD products in ground-range geometry. Both SLC and GRD data are ranked in chronological order and further processed by the Sentinel-1 toolbox in SNAP (Veci et al., 2017). In terms of GRD data, the radiometric terrain correction (Filipponi, 2019, Truckenbrodt et al., 2019) is performed to produce the time-series gamma0 results (Small, 2011). Associated layover and shadow mask are generated through the SAR simulation module based on the Shuttle Radar Topography Mission (SRTM) DEM (Mondini, 2017). After that, gamma0 results at the pre- and post-fire dates are preserved as pre- and post-fire backscatter layers (BS_{pre} and BS_{post}) and the backscatter difference layer (BS_{diff}) is generated through the pixel-wise subtractions in Equation (6.1).

$$BS_{diff} = Max\{BS_{pre} - BS_{ts_1}, BS_{pre} - BS_{ts_2}, \dots, BS_{pre} - BS_{post}\} \quad (6.1)$$

The quality of the interferometric phase can be quantitated as the magnitude of the interferometric coherence (Bamler and Hartl, 1998). Compared to the unburned area, more fire-derived changes occurred in the scatterers of burned areas, which introduce extra noise into the interferometric phase and cause severe temporal decorrelations in the interferometric coherence. Therefore, coherence information can be used as an important piece of information for the BA discrimination (see section 6.6). During the process, time-series InSAR pairs derived from the chronological SLC data are fed into the TOPS interferometry workflow (Veci, 2015) within SNAP to generate time-series interferometric coherence results. As presented in Figure **6.3**, the pre- and post-fire coherence layers are preserved and the coherence difference layer is produced based on Equation (6.2).

$$COR_{diff} = Min\{COR_{ts_1}, COR_{ts_2}, \dots, COR_{ts_n}\} \quad (6.2)$$

The pre-processed Sentinel images are transformed into the UTM/WGS84 coordinate system and resampled to a uniform spatial resolution of 20 m applying the bilinear interpolation (Baboo and Devi, 2010). This Sentinel stack is made up of 30 dimensions including 18 Sentinel-2 SR layers (9 pre-fire and 9 post-fire), 12 Sentinel-1 dual-polarimetric (VV, VH) SAR layers (2 pre-fire BS layers, 2 BS difference layers, 2 post-fire BS layers, 2 pre-fire COR layers, 2 COR difference layers, and 2 post-fire COR layers). Worth noting that, layover and shadow areas in SAR layers have been masked out by assigning large negative values (-9999).

Considering that BAs only occupy a small part of the whole Sentinel coverage, coarse clipping around the fire perimeters is applied on the layer stack to adjust the proportion of the burned and unburned samples in the object dataset. After clipping, the multi-dimensional layer stacks are segmented by watershed transformation (Vincent and Soille, 1991) and full lambda-schedule (FLS) algorithm (Liu et al., 2011) delivering boundaries of the objects as vector files (shapefiles) and means of the values inside in each dimension as the attributes for classification. This object dataset is named after SR+COR+BS in the following sections for concise. The same operation is further performed on three more multi-source layer stacks to construct the SR+COR, SR+BS, and SR datasets. Object samples are of 30 dimensions in the SR+COR+BS dataset, 24 dimensions in the SR+COR dataset, 24 dimensions in the SR+BS dataset, and 18 dimensions in the SR dataset.

Annotations of the samples in the object dataset are derived from the difference of normalized difference vegetation index (dNDVI) between the pre- and post-fire PlanetScope images. The most difficult part during the binarization is to determine the threshold for each dNDVI map. Binarizing the whole dNDVI map with a consistent threshold is not recommended since BAs often cover a large area where dNDVI values inside vary with heterogeneous landscapes and uneven environments. Hence, dNDVI maps are cropped into small patches with sizes of 256×256, where the binary threshold is manually chosen for each patch. These binarized patches are then integrated over each site and projected back into the same coordinate system as the multi-dimensional layer stack preserving the spatial resolution of 3 m. Labels of the Sentinel objects are

further assigned using binarized values of pixels that fall inside the associated object polygon based on the majority voting rule (Zwicker, 2016).

After the construction of associated object datasets, a training and testing split strategy is further applied. Twelve test sites are divided into ten dependent and two independent sites (Stone and Walker), where each dependent site provides 60% of the objects for training and the other 40% for dependent testing while all objects in the independent sites are used for independent testing. The dependent testing offers to analyze the performance of different classification methods by feature sources, test sites, and land cover types, whereas the independent testing is designed to evaluate the robustness of the developed model over independent fire perimeters.

6.3.2 Siamese Self-Attention Classification

The diversity of the multi-source attributes is the main concern when applying binary classification because prior studies evidenced that adequate inputs have a decisive impact on the BA mapping accuracy (Ramo et al., 2018). Therefore, the self-attention mechanism (Vaswani et al., 2017) is employed here to explore the latent relations among multi-source features and to improve their synergies since it was designed to draw global dependencies of long sequences. Two self-attention modules are further connected by a Siamese structure with pairs of inputs and shared weights (Bromley et al., 1994) to maximize the inter-class and minimize the intra-class variations. This ensures the flexibility of the trained classifier on the whole training dataset. After training, either self-attention classifier can be used for the burned and unburned

discrimination. The details of the SSA classification strategy are presented in Figure 6.4.

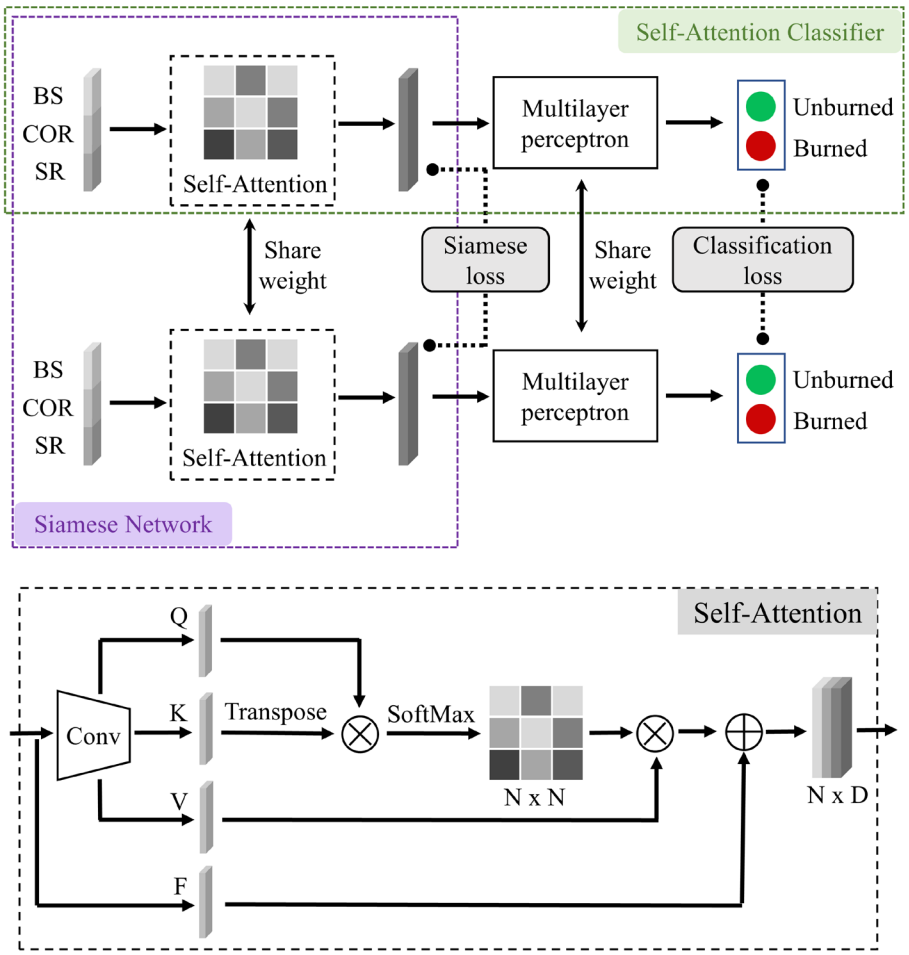


Figure 6.4 Siamese Self-Attention (SSA) classification.

In the proposed SSA network, a pair of multi-source attribute sequences are fed into a convolutional layer to generate the associated query Q, key K, and value $V \in \mathrm{R}^{N\times D}$. A subsequent softmax operator is applied on the dot product of Q and K to get the self-attention mask $M \in \mathrm{R}^{N\times N}$.

$$m_{ij} = \frac{\exp\left(Q_i K_j^T\right)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\exp\left(Q_i K_j^T\right)} \tag{6.3}$$

where, m_{ij} is the element in M standing for the correlation between the ith row of the query and the jth row of the key.

The replicate input sequence F is further added to the masked weighted value MV as the output of the self-attention because this provides the network with a long-range channel-wise context dependency to tackle the gradient vanishing and to boost feature discriminability (Gomez et al., 2017).

$$P = MV + \alpha F \tag{6.4}$$

where α is initialized as 0 and gradually learns to assign more weights in the training process.

To make sure SAR attributes in the layover and shadow areas have no contributes to the binary classification, the large negative values of the input are reset as zero in F during the replication. Whereas these large negative values are retained in the calculation of MV because the softmax operator can convert them into zero in M and hence have no contributions to P.

The output P is then fed into the multilayer perceptron (MLP) with two hidden layers to finalize the classification task, where the activation function is set as Leaky ReLU (Laurent and Brecht, 2018) and the classification loss is set as cross-entropy (Zhang and Sabuncu, 2018).

$$L_c = -\frac{1}{2}\sum_{c=1}^{2}[y_c \log \hat{y}_c + (1 - y_c)\log(1 - \hat{y}_c)] \tag{6.5}$$

where y_c is the binarized label and $\hat{y}_c$ is the predicted probability of class c.

A Siamese loss is also considered in the total loss during the training process, which is defined after the Euclidean distance between the output feature pairs (f_1, f_2) of the self-attention blocks.

$$L_{sia} = (1 - y_{sia})S^2 + y_{sia}\max(0,1 - S)^2 \tag{6.6}$$

$$S = 1 - \exp(-\|f_1 - f_2\|_2) \tag{6.7}$$

where, $y_{sia} = 0$ if the input pair are from the same class and $y_{sia} = 1$ if they are from different classes. Introducing Siamese loss can help the network to learn a nonlinear transformation that maps the input data into a low-dimensional manifold, where the distance between same-class samples is smaller and so the opposite (Yang and Jin, 2006).

The total loss of the whole SSA network is then updated as the sum of classification loss L_c and Siamese loss L_{sia}

$$L_{SSA} = L_c + \lambda L_{sia} \tag{6.8}$$

where λ is a regularization parameter to balance the two losses in the training phase.

6.3.3 Accuracy Metrics

OE, CE, and Dice Coefficient (DC) are used to evaluate the quality of the BA product (Padilla et al., 2015). All these accuracy metrics can be calculated using the confusion matrix with the numbers of TP, FP, TN, and FP samples (Table 6.4). In BA mapping, TPs represent the burned samples that are correctly labeled, FPs are unburned samples incorrectly labeled as burned, TNs are the unburned samples correctly labeled, and FNs represent burned samples incorrectly labeled as unburned.

Table 6.4 Metric calculation reference

Reference / Classified	Burned	Unburned
Burned	TP	FP
Unburned	FN	TN

Based on Table 6.4, the OE, CE and DC can be calculated by

$$OE = \frac{FN}{TP + FN} \tag{6.9}$$

$$CE = \frac{FP}{TP + FP} \tag{6.10}$$

$$DC = \frac{2TP}{2TP + FN + FP} \tag{6.11}$$

6.3.4 Implementation details

The training is carried out independently on the four training datasets (SR+COR+BS, SR+COR, SR+BS, and SR) using two classification methods (SSA and Random forest (RF)) (Breiman, 2001, Çömert et al., 2019). The training of SSA undertakes on two NVIDIA GeForce RTX 2080Ti GPUs under the PyTorch framework (Ketkar, 2017), where synchronized Adam solver (Kingma and Ba, 2014) is used allowing each GPU to process a batch of 32 pairs simultaneously. Layers in the SSA model are randomly initialized by a Gaussian distribution with zero mean and a standard deviation of 0.001. The base learning rate, momentum, and weight decay are set as 0.01, 0.9, and 0.0001. The training of the RF undertakes on the Intel Core i7-9700k processer with a base frequency of 3.60 GHz, where the number of decision trees is initiated as 1000 with a depth of 400.

6.4 Experiment Results

6.4.1 Dependent Testing

After training, four dependent testing datasets (SR+COR+BS, SR+COR, SR+BS, and SR) are fed into four SSA classifiers and four RF classifiers to predict the burned probabilities, where burned probabilities over 0.5 are classified as burned. Predictions are further compared with PlanetScope-derived annotations and associated accuracy metrics are calculated using the confusion matrix. DCs of the SSA and RF classifiers over each dependent site are presented in Figure **6.5** and more details about the OEs and CEs are listed in Table 6.5 and Table 6.6.

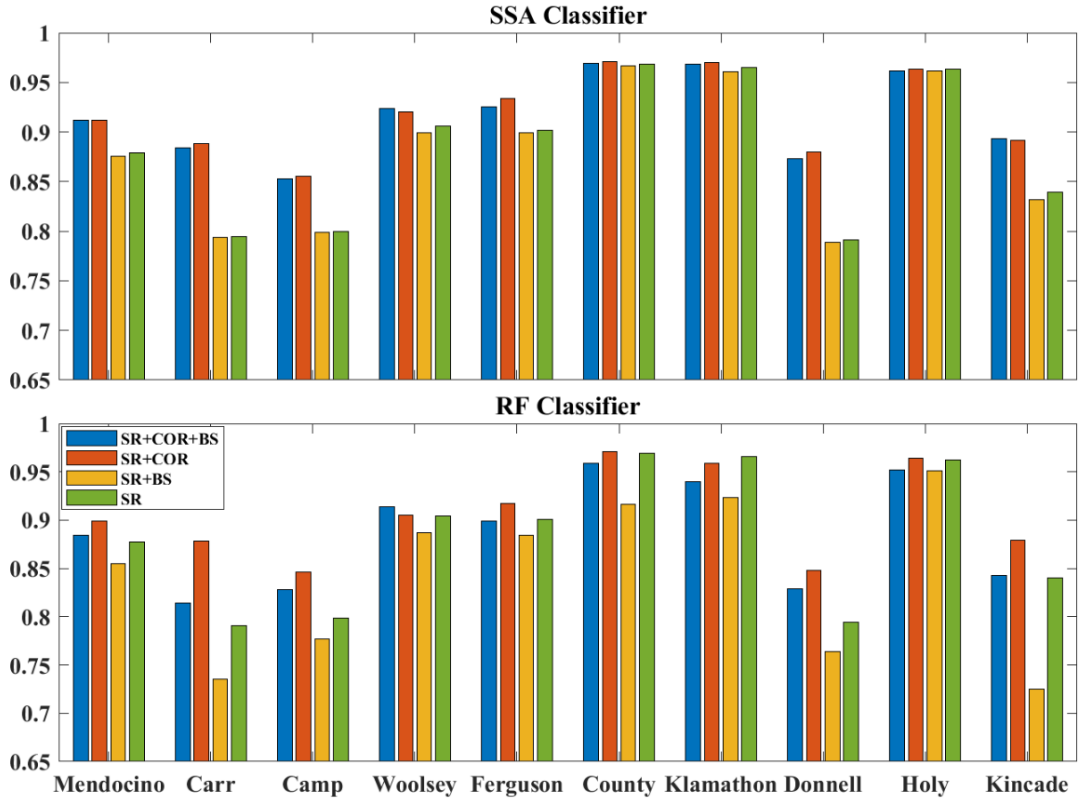


Figure 6.5 DCs of different classifiers over each dependent site.

The SSA classification results are further projected back into the UTM/WGS-1984 coordinate system to present the final BA product. Since dependent testing datasets

only cover 40% of the objects in the dependent sites, these testing results can not give us an integrated BA map over each dependent site. Hence, predictions are also carried out on training datasets using the four SSA classifiers to generate the whole BA maps in Figure 6.6 to Figure 6.15. Worth noting, predictions over the training dataset are not involved in the calculation of any accuracy metrics above.

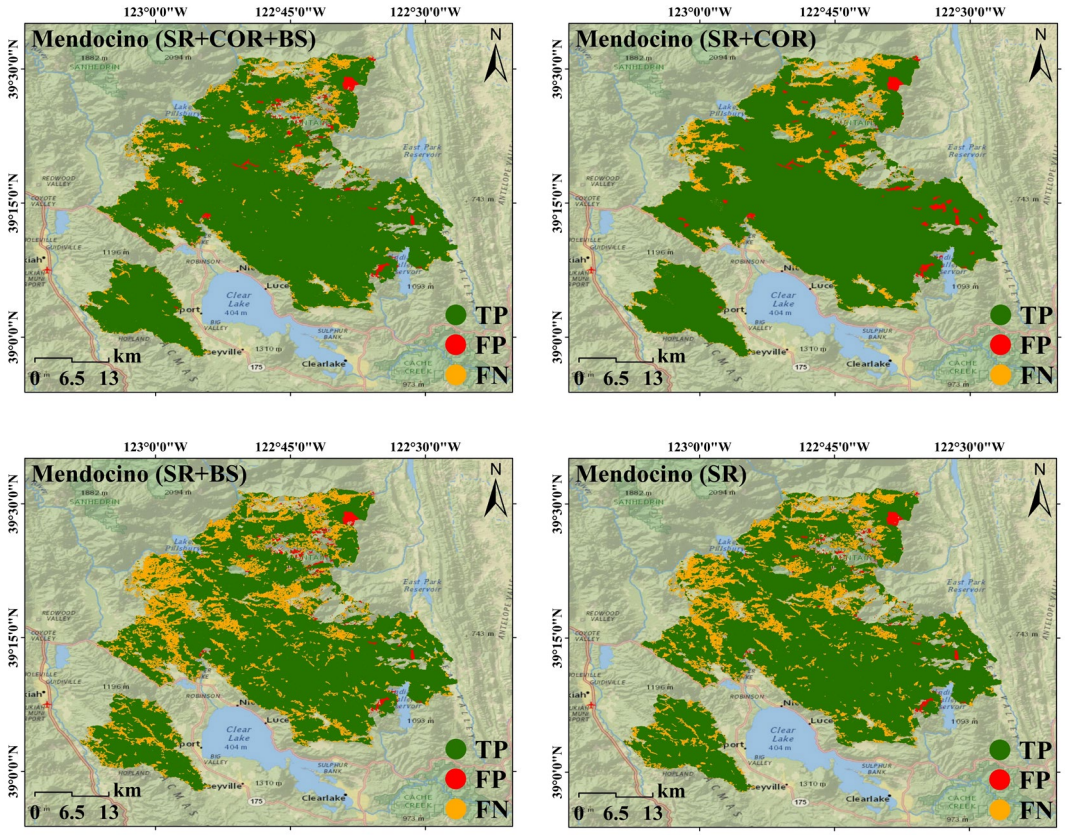


Figure 6.6 Predicted BA maps (SSA) over the Mendocino fire.

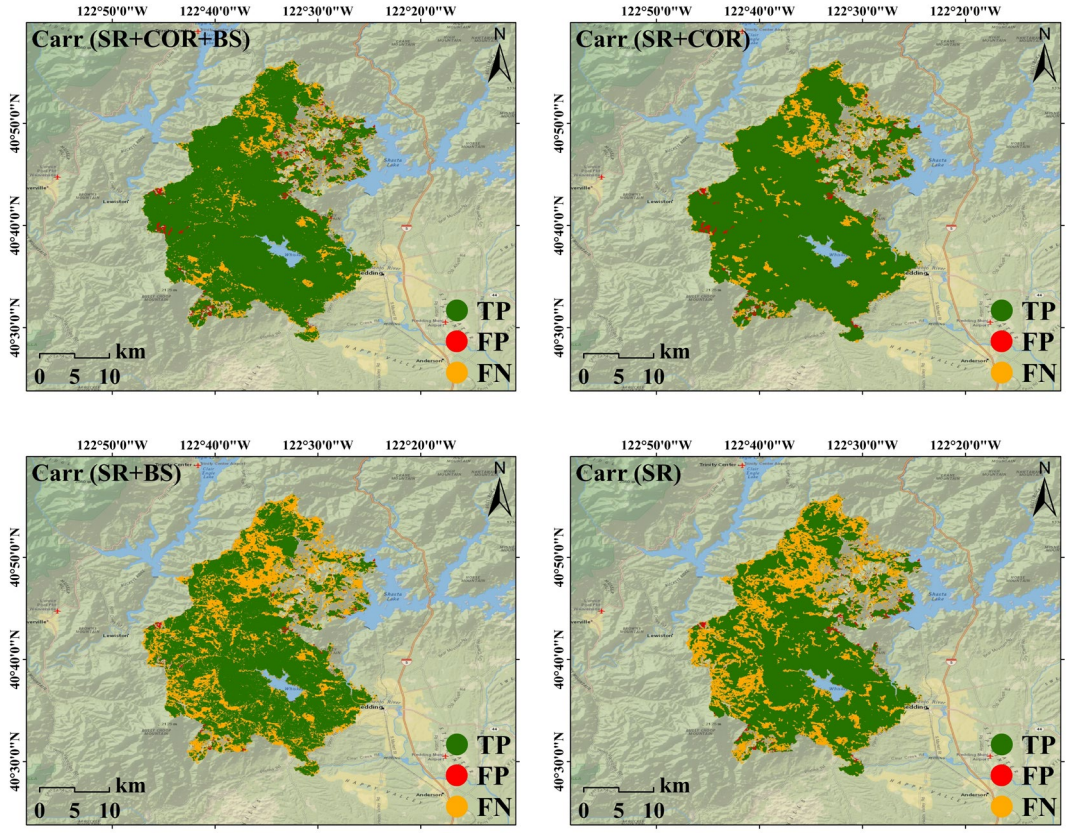
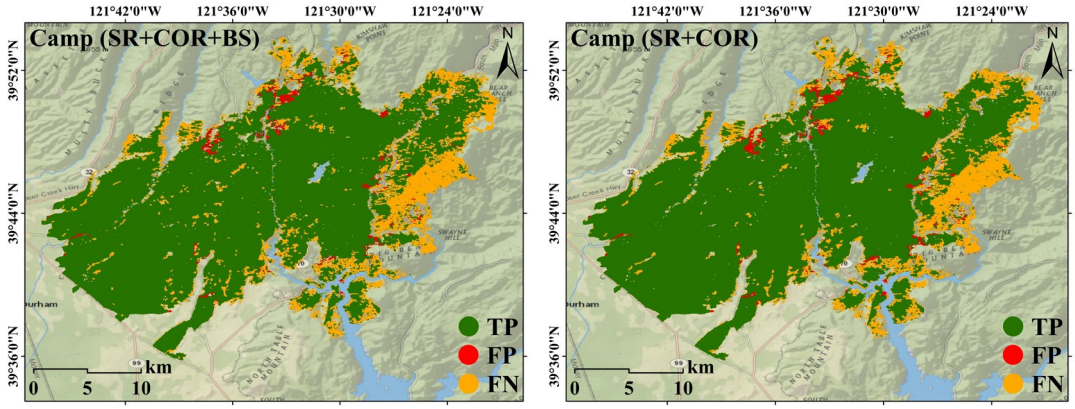


Figure 6.7 Predicted BA maps (SSA) over the Carr fire.



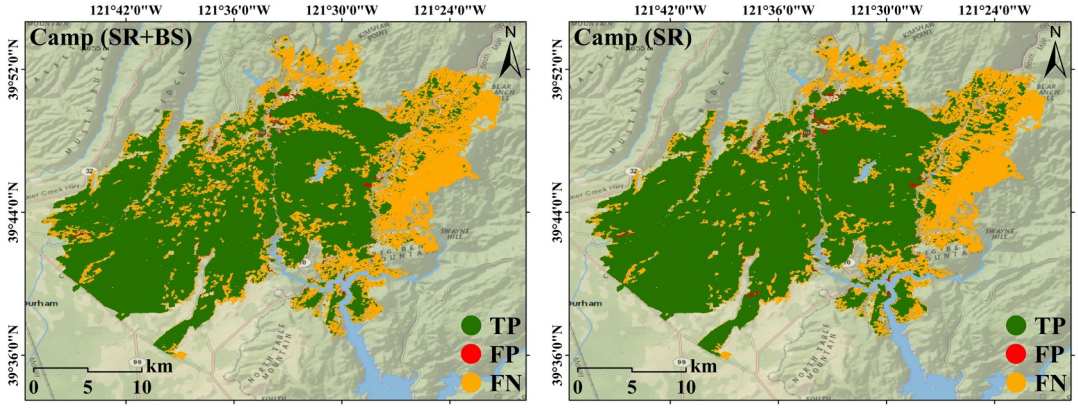


Figure 6.8 Predicted BA maps (SSA) over the Camp fire.

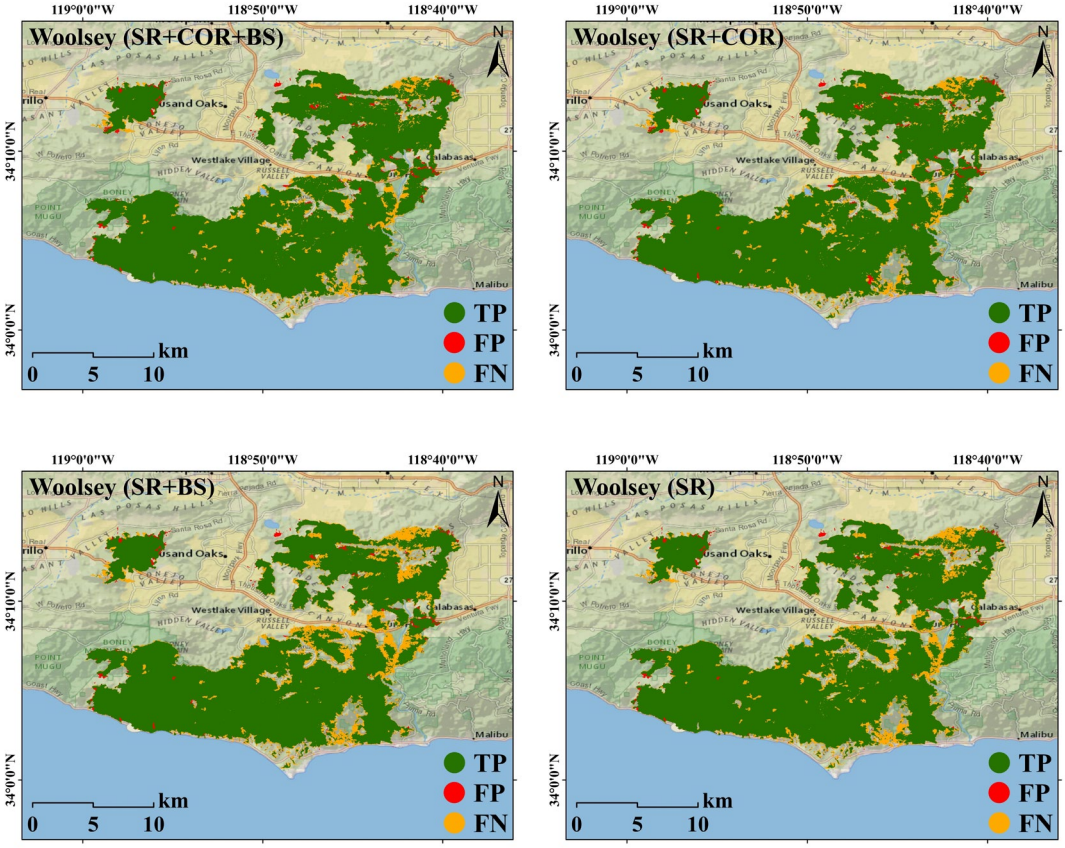


Figure 6.9 Predicted BA maps (SSA) over the Woolsey fire.

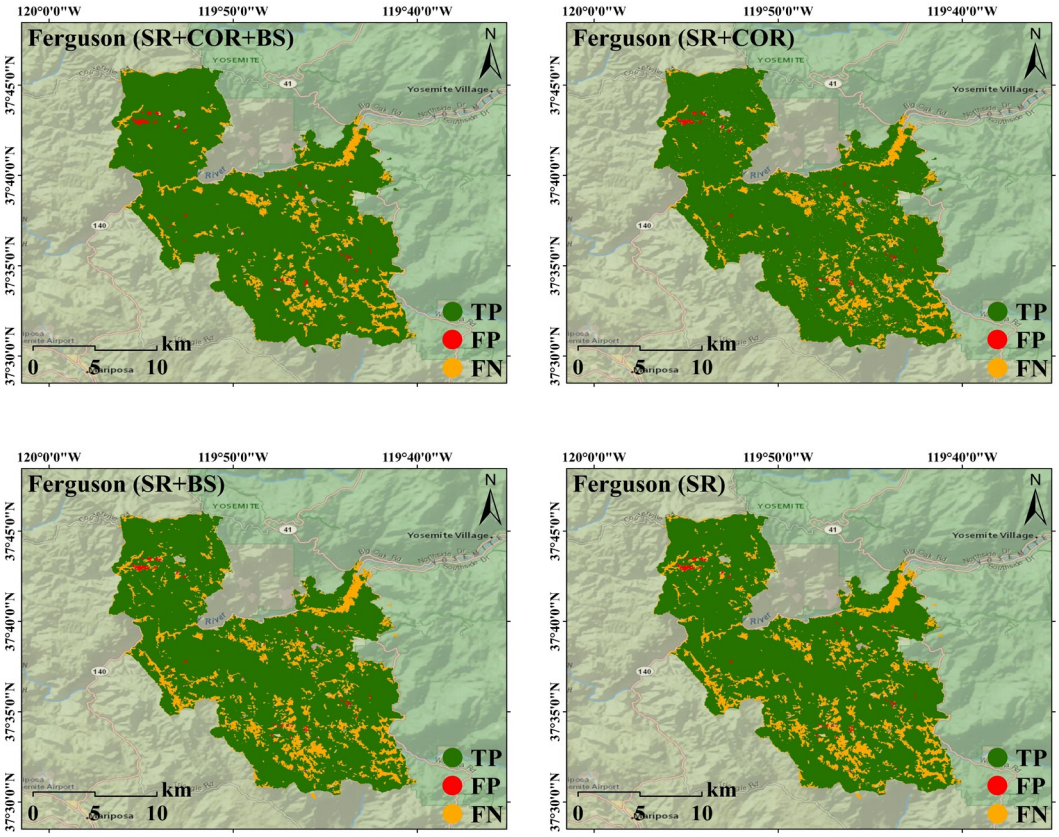
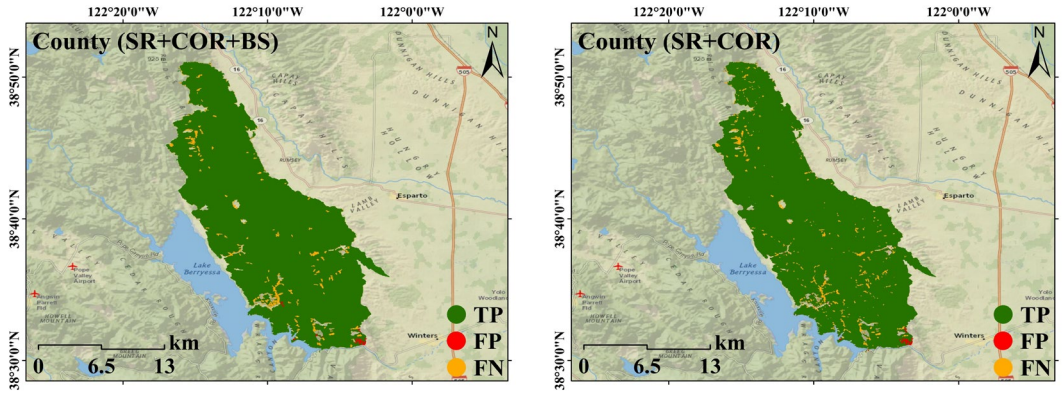


Figure 6.10 Predicted BA maps (SSA) over the Ferguson fire.



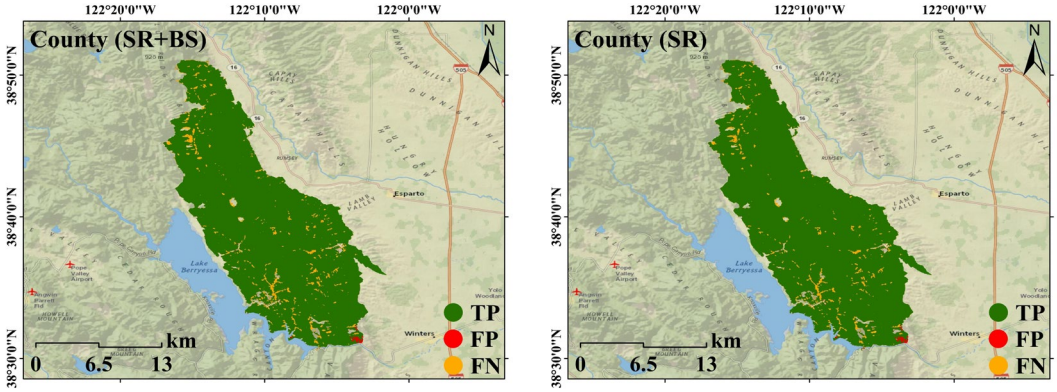


Figure 6.11 Predicted BA maps (SSA) over the County fire.

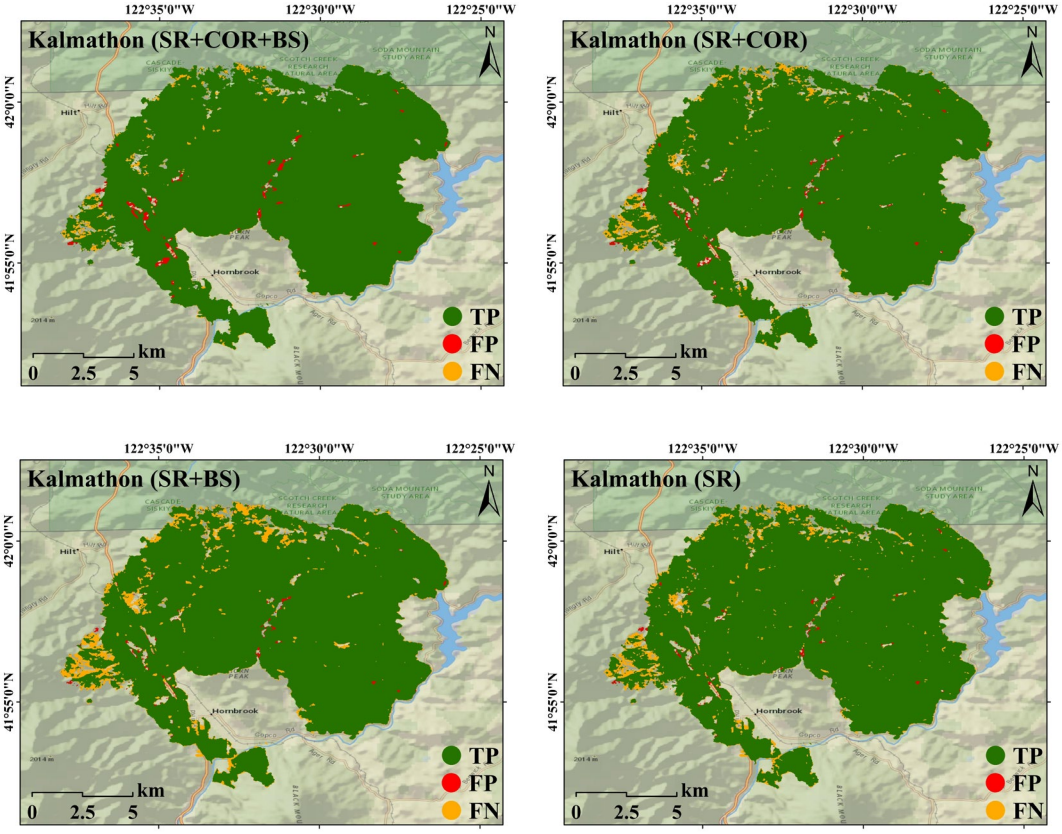


Figure 6.12 Predicted BA maps (SSA) over the Klamathon fire.

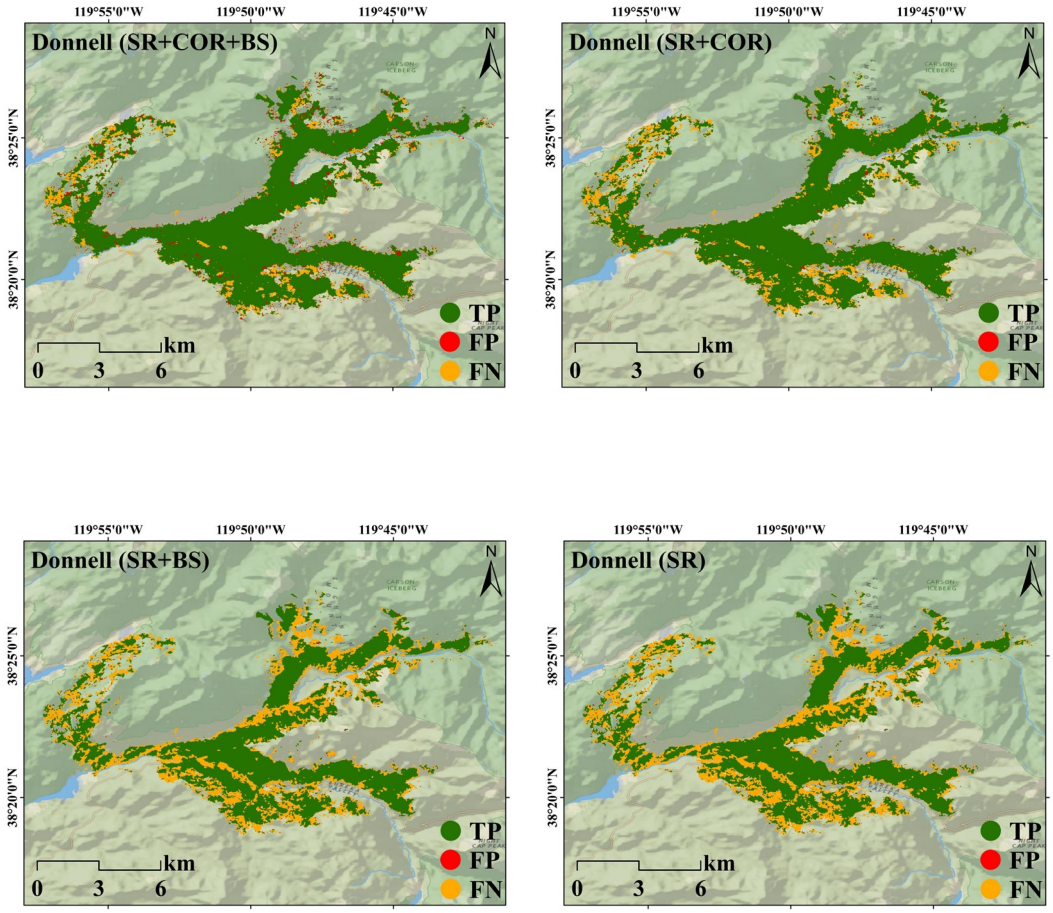
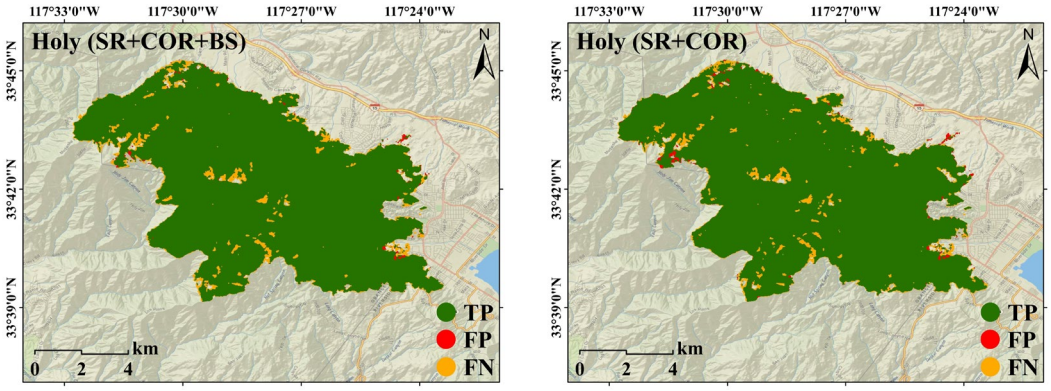


Figure 6.13 Predicted BA maps (SSA) over the Donnell fire.



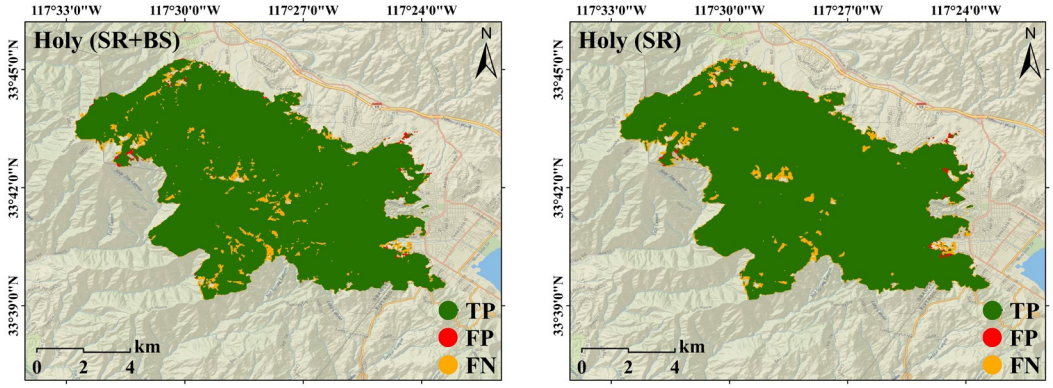


Figure 6.14 Predicted BA maps (SSA) over the Holy fire.

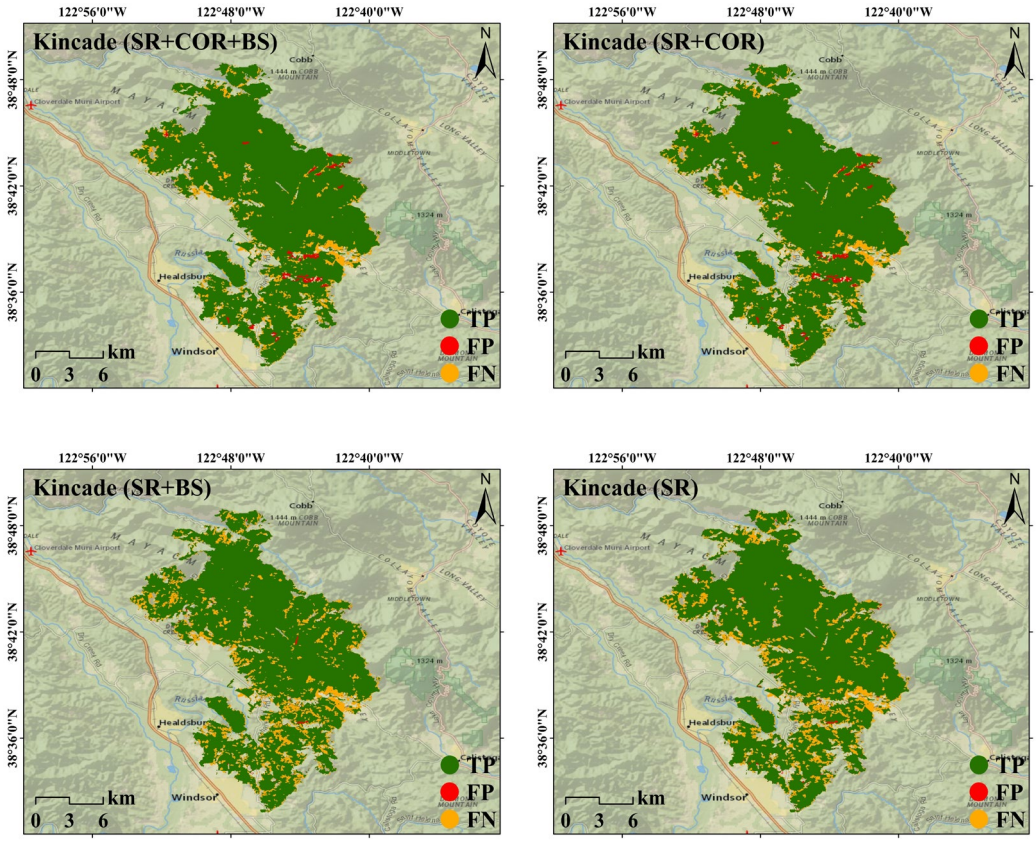


Figure 6.15 Predicted BA maps (SSA) over the Kincade fire.

Table 6.5. Accuracy metrics of SSA classification result over each dependent site

		Mendocino	Carr	Camp	Woolsey	Ferguson	County	Klamathon	Donnell	Holy	Kincade
SR+COR+BS	**OE**	14.3%	18.7%	23.8%	12.1%	13.4%	5.6%	4.2%	20.0%	6.5%	17.7%
	CE	2.6%	3.1%	3.2%	2.6%	0.8%	0.4%	2.2%	3.8%	0.8%	2.5%
	DC	91.2%	88.4%	85.3%	92.4%	92.5%	96.9%	96.8%	87.3%	96.2%	89.3%
SR+COR	**OE**	14.3%	18.1%	23.5%	12.6%	11.7%	5.3%	4.4%	19.1%	6.1%	17.5%
	CE	2.6%	2.9%	3.2%	2.8%	0.8%	0.4%	1.6%	3.4%	1.1%	2.9%
	DC	91.2%	88.8%	85.5%	92.0%	93.4%	97.1%	97.0%	88.0%	96.3%	89.2%
SR+BS	**OE**	20.7%	33.1%	33.1%	16.7%	17.8%	6.1%	6.5%	34.3%	6.4%	28.2%
	CE	2.1%	2.4%	0.9%	2.3%	0.8%	0.4%	1.1%	1.1%	1.0%	1.0%
	DC	87.6%	79.4%	79.9%	89.9%	89.9%	96.7%	96.1%	78.9%	96.2%	83.2%
SR	**OE**	20.1%	32.9%	33.0%	15.8%	17.3%	5.8%	5.9%	34.1%	6.3%	27.1%
	CE	2.3%	2.4%	0.9%	2.0%	0.8%	0.3%	1.1%	1.0%	1.0%	1.1%
	DC	87.9%	79.5%	80.0%	90.6%	90.2%	96.8%	96.5%	79.1%	96.3%	83.9%

Table 6.6. Accuracy metrics of RF classification result over each dependent site

		Mendocino	Carr	Camp	Woolsey	Ferguson	County	Klamathon	Donnell	Holy	Kincade
SR+COR+BS	**OE**	19.0%	30.0%	27.7%	13.8%	17.8%	7.6%	10.3%	27.4%	8.4%	26.4%
	CE	2.6%	2.7%	3.1%	2.8%	0.9%	0.4%	1.3%	3.4%	0.8%	1.2%
	DC	88.4%	81.4%	82.8%	91.4%	89.9%	95.9%	94.0%	82.9%	95.2%	84.3%
SR+COR	**OE**	16.5%	19.8%	24.9%	15.5%	14.7%	5.2%	6.6%	24.2%	6.0%	20.0%
	CE	2.6%	3.0%	3.2%	2.7%	0.8%	0.4%	1.4%	3.6%	1.1%	2.6%
	DC	89.9%	87.8%	84.6%	90.5%	91.7%	97.1%	95.9%	84.8%	96.4%	87.9%
SR+BS	**OE**	24.3%	41.0%	36.2%	19.0%	20.3%	15.3%	13.8%	37.8%	8.7%	43.0%
	CE	1.9%	2.4%	0.7%	2.0%	0.8%	0.2%	0.8%	1.1%	0.7%	0.5%
	DC	85.5%	73.5%	77.7%	88.7%	88.4%	91.6%	92.3%	76.4%	95.1%	72.5%
SR	**OE**	20.5%	33.4%	33.0%	16.0%	17.5%	5.7%	5.9%	33.6%	6.4%	26.8%
	CE	2.2%	2.5%	0.9%	2.1%	0.8%	0.3%	0.8%	1.3%	1.0%	1.4%
	DC	87.7%	79.1%	79.9%	90.4%	90.1%	96.9%	96.6%	79.4%	96.2%	84.0%

Variations of BA mapping accuracies achieved by the four SSA classifiers are further analyzed as a function of the binary threshold applied on the burned probability. As indicated in Figure 6.16, DCs over shrublands and grasslands are slightly affected by the input multi-source data (shrubland: 96.0%-96.5%, grassland: 94.2%-95.4%) with optimum thresholds located between 0.4 and 0.6. More prominent improvements in DCs of multi-source inputs are observed over forests and savannas (forest: 77.7%-90.6%, savanna: 89.1%-92.9%). Accuracies over forests improve when the softmax burned probability threshold was increased (optimum: 0.5-0.7) and conversely over savannas (optimum: 0.3-0.5).

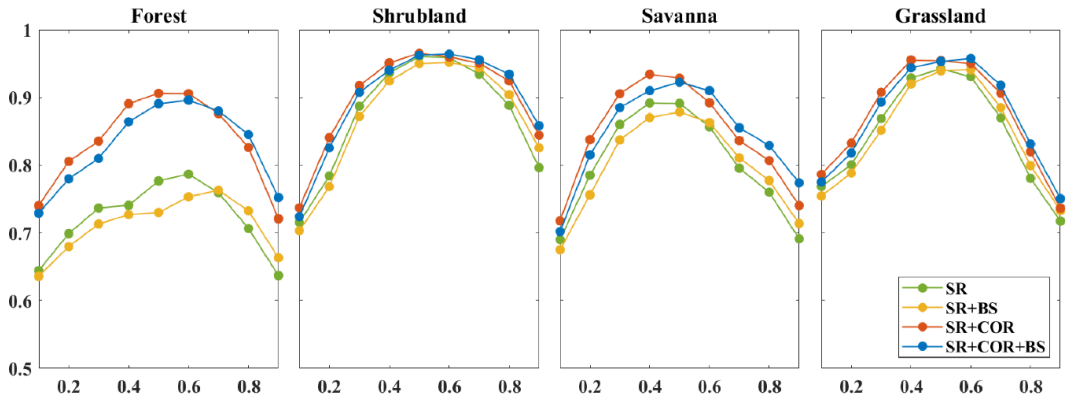


Figure 6.16 Variation of DC (SSA) as a function of the threshold applied on the burned probability.

6.4.2 Independent Testing

After validating the better performance of the proposed method, the robustness of the multi-source (SR+COR+BS) SSA classifier needs to be investigated over independent cases. This is necessary because the main part of the proposed method is underpinned by deep learning networks, which strongly rely on the sufficiency and diversity of the training dataset, although practically it is impossible to collect samples from every

place. Predictions with OE of 8.7%, CE of 2.6%, and DC of 94.2% (Stone) and OE of 18.8%, CE of 3.5%, and DC of 88.2% (Walker) are achieved by the multi-source SSA classifier over the independent testing dataset. Figure 6.17 illustrates the associated reprojected BA maps.

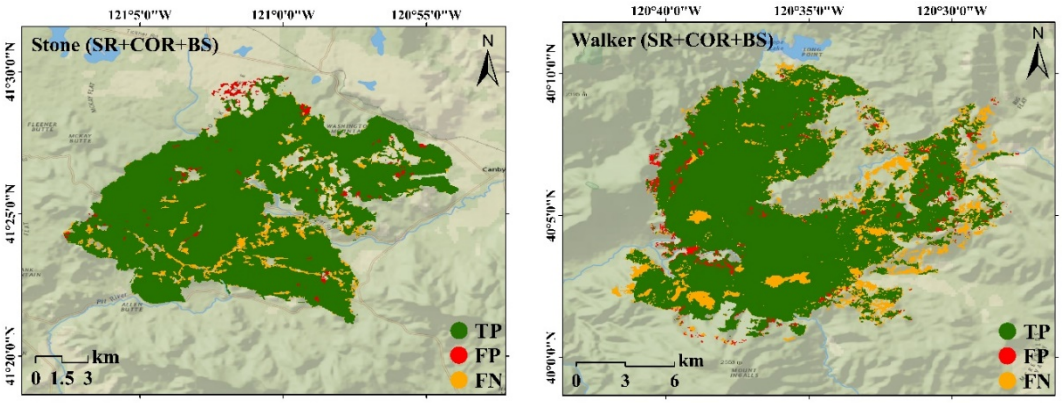


Figure 6.17 Predicted BA maps (SSA) over the Stone and Walker fires.

6.5 Discussion

As presented in Figure 6.5, overall DCs of both SSA and RF classifiers are lower in test sites with steep terrains or/and forest dominant covers such as Camp and Carr than areas with flat topography or/and non-forest covers like County and Klamathon. This directly relates to the characteristics of active and passive remote sensing acquisitions. On the one hand, spectral overlaps among BA and cast shadow originated from steep terrains or/and forests tend to cause misclassification when using SR-derived features (Stroppiana et al., 2009, Fassnacht et al., 2021). On the other hand, although the cloud coverage isn't a problem in SAR-related features due to the penetration of microwave, the unique side-looking imaging geometry tends to cause distortions (foreshortening, layover, and shadow) over steep terrains. Besides that, the SAR data used in our study

are obtained by the Sentinel-1 C-band sensor where the penetration is not as strong as L-band SAR sensors and thus prone to be affected by the absorption and attenuation (Tanase et al., 2010a, Tanase et al., 2010b).

The superiority of using multi-source features for BA discrimination is also indicated in Figure 6.5. Improvements are more prominent in test sites with steep terrains or/and forest covers (Camp: 5.5% for SSA classifier and 4.7% for RF classifier, Carr: 9.3% for SSA and 8.7% for RF) than regions with flat topography or/and non-forest covers (County: 0.3% for SSA and 0.2% for RF), where a large part of SR-derived errors are mitigated through introducing SAR features. Overall DCs achieved over the SR+COR dataset are higher than (3.4% for SSA and 2.4% for RF) those over the SR dataset. Conversely, lower DCs are observed over the SR+BS dataset than the SR dataset (0.3% for SSA and 2.6% for RF) indicating that the accuracy improvement mainly attributes to the COR features rather than the BS features.

Comparison is also carried out between different classifiers, where the performance is evaluated by their flexibility over multi-source features and the classification accuracy achieved. Results in Figure 6.5 indicated that when new features are added to the input attribute, the SSA classifier can automatedly learn to pick up the more critical ones for BA discrimination and is less affected by redundant features through assigning proper weights during the training process. This largely attributes to the self-attention module used to learn the latent relations among inter-source and intra-source features and the Siamese structure employed to maximize the inter-class and minimize the intra-class variation. As to the classification accuracy, higher DCs are achieved by the SSA

classifiers over each dependent site than those from the RF classifiers (see Table 6.5 and Table 6.6).

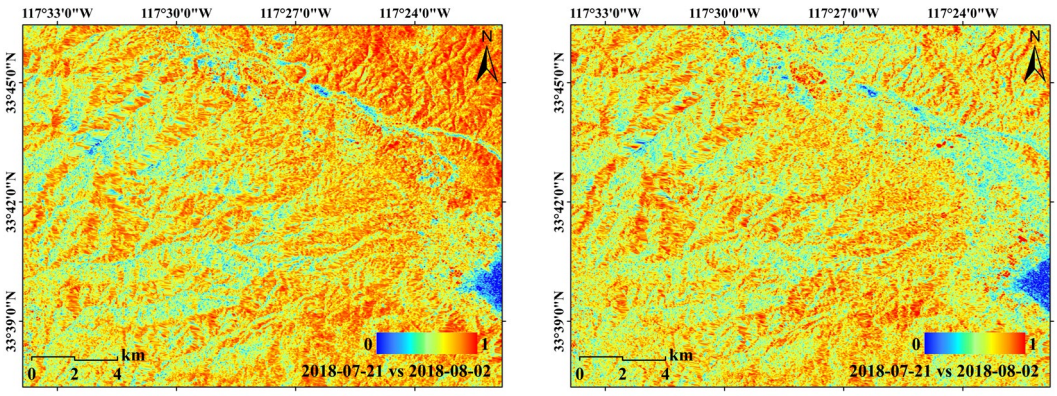
In summary, the effectiveness of the SSA strategy over the synergy of Sentinel-1 COR, BS, and Sentinel-2 SR data has been validated based on the analysis above. Moreover, the robustness of the multi-source SSA classifier is evaluated over two independent sites delivering an average DC of about 90%. Accuracies in independent sites are overall lower than the dependent ones with similar land cover distributions and terrains. This is because less spatial and temporal correlation is shared between the training and independent testing datasets than the training and dependent testing ones although two independent sites are also located in the state of California with fire events that occurred in 2018-2019.

Labels used for the validation are generated by PlanetScope data due to the validation protocol mentioned in (Chuvieco et al., 2019), which focuses on using higher spatial resolution data than the BA product derived. But the high cost of such commercial satellite imagery may restrict the application of this protocol over large geographical areas. An alternative way is to generate the pan-sharpened (10 m) Sentinel-2 products using the sen2res module (Brodu, 2017) in SNAP. Although the spatial resolution of this product is lower than the PlanetScope acquisitions, the reflectance in 13 optical bands acquired by MSI is preserved with open access.

6.6 Supplementary Material

6.6.1 Principle of interferometric coherence for BA mapping

The principle of SR and BS features in BA discrimination has been widely investigated in the past years. This section will focus on exploring how the COR features change with the fire based on a simplified coherence model and why it can be employed as an important auxiliary for BA mapping. Normally, wildfires are more likely to occur in vegetated areas where coherences are overall lower than the permanent scatterers like urban facilities while higher than water bodies (Ferretti et al., 2001). Vegetations are sabotaged by the fast destruction of wildfire causing severe temporal decorrelation in COR features. Afterward, vegetations gradually rehabilitate in burns though this may take months to years depending on the vegetation type. The pre-fire and post-fire dual polarimetric coherence maps over the Holy site are presented in Figure 6.18 with the time-series coherence difference.



(a) (b)

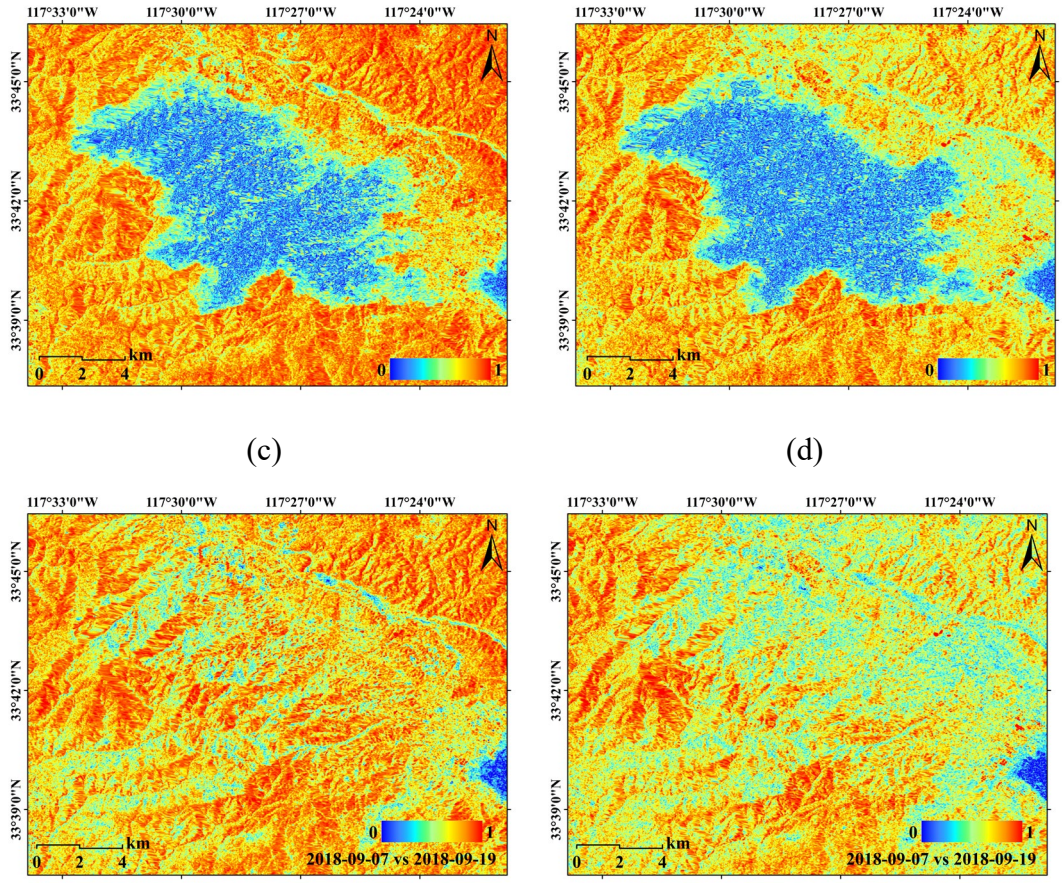


Figure 6.18 Interferometric coherence: a) Pre-fire map in VV polarization; b) Pre-fire map in VH polarization; c) Difference map in VV polarization; d) Difference map in VH polarization; e) Post-fire map in VV polarization; f) Post-fire map in VH polarization.

Practically, after compensating the spectral and SNR decorrelations, the volume and temporal decorrelations become two main components in interferometric coherence (Lavalle, 2009), which are directly related to the internal structure and the temporal stability of the scattering medium. Therefore, analysis of coherence changes is mainly carried out on these two decorrelations using a simplified coherence model in (Lei et al., 2017) given by

$$\gamma=\exp(j\phi_0)\frac{\gamma_{vt}+\gamma_{gt}m(\omega)}{1+m(\omega)} \tag{6.12}$$

where, ϕ_0 is the ground phase, γ_{gt} is a complex value describing the ground temporal decorrelation, γ_{vt} is the coherence component with coupled effects from volumetric scatterers and temporal changes, $m(\omega)$ represents the ground to volume ratio varying with polarization ω.

Unlike the coherence function mentioned in Chapter 3 and Chapter 4, Equation (6.12) accounts for both position and dielectric property changes in the ground and volume layers. This is because temporal baselines of spaceborne data involved in this chapter (6 and 12 days) are much longer than the airborne acquisitions (less than hours) used before and the fast destruction of the wildfire further accelerates and enhances these changes. Hence, the stable ground and the assumption of only wind-derived temporal change no longer stay. In this condition, the volumetric-temporal coherence function can be reformulated as

$$\gamma_{vt}=\frac{\int_0^{h_v}\rho_v(z)\eta_v(z)d_v(z)\exp(jk_zz)dz}{\int_0^{h_v}\rho_v(z)dz} \tag{6.13}$$

where, h_v is the forest height, $\rho_v(z)$ represents the radar backscattering attenuation at height z, $\eta_v(z)$ and $d_v(z)$ are the vertical decorrelation distributions derived from the changes of position and dielectric property in the volume layer, and k_z is the vertical wavenumber.

Coherence magnitudes in Equations (6.12) and (6.13) range from 0 to 1 representing the correlation level between two observations. The behaviors of coherence changes

throughout the fire can be explained by these functions using different decorrelation components.

Before the fire, the dielectric properties of the ground and volume layers are relatively stable and the main decorrelation source in the pre-fire coherence (days before the start date) is the wind-derived random motion of leaves and branches. In this condition, the coherence function can be simplified as (Lavalle and Hensley, 2012)

$$\gamma^{pre} = exp(j\phi_0)\frac{\gamma_{vt}^{pre} + m(\omega)}{1 + m(\omega)} \tag{6.14}$$

$$\gamma_{vt}^{pre} = \frac{\int_0^{h_v} \rho_v(z)\eta_v(z)exp(jk_z z)dz}{\int_0^{h_v} \rho_v(z)dz} \tag{6.15}$$

During the fire, the stability of the scattering medium is damaged by the chemical and physical conversion process, and therefore the magnitudes of γ_{gt}^{ts} and γ_{vt}^{ts} in the time series coherences rapidly drop to zero which further leads to $\left|\gamma^{diff}\right| \approx 0$ in the fire-affected area.

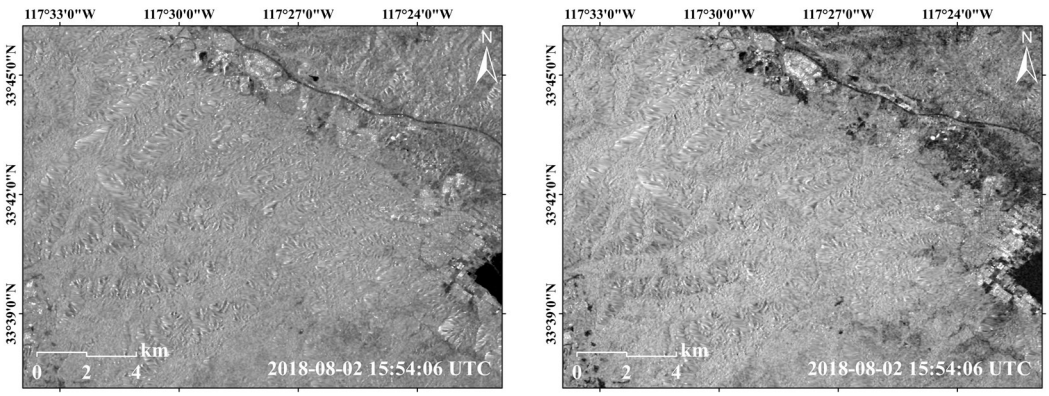
Assuming that the vegetation area is severely burned where most of the leaves and branches are destroyed. In this condition, the decorrelation of volume scattering can be neglected and the ground to volume amplitude ratio becomes very large in the post-fire coherence (days after the contained date and before the vegetation rehabilitation). Besides that, the ground decorrelation is more obvious than in normal conditions since the temperature and moisture of the burned ground change fast after the fire.

$$\gamma^{post} = \exp(j\phi_0)\gamma_g^{post} \tag{6.16}$$

In summary, the magnitudes of the during-fire coherence maps are much smaller than the pre and post-fire ones over the BA as illustrated in Figure 6.18. This means the noise in the during-fire interferometric phase is much larger than the pre and post ones, which applies to both the VV and VH polarizations. This variation of interferometric coherence throughout the fire event can be used as an important piece of information in the fire patch classification.

6.6.2 Illustration of the multi-source data

As mentioned above, the original satellite acquisitions over each site include the pre- and post-fire Sentinel-2 images with 13 bands, time-series Sentinel-1 SAR SLC and GRD images in dual polarizations, and the pre- and post-fire PlanetScope images with 4 bands. Sentinel acquisitions are used to generate multi-source features for the BA discrimination and PlanetScope data are used to provide annotations for training and validation. Figure 6.18 to Figure 6.21 presents the associated pre-processed layers over the Holy site.



(a) (b)

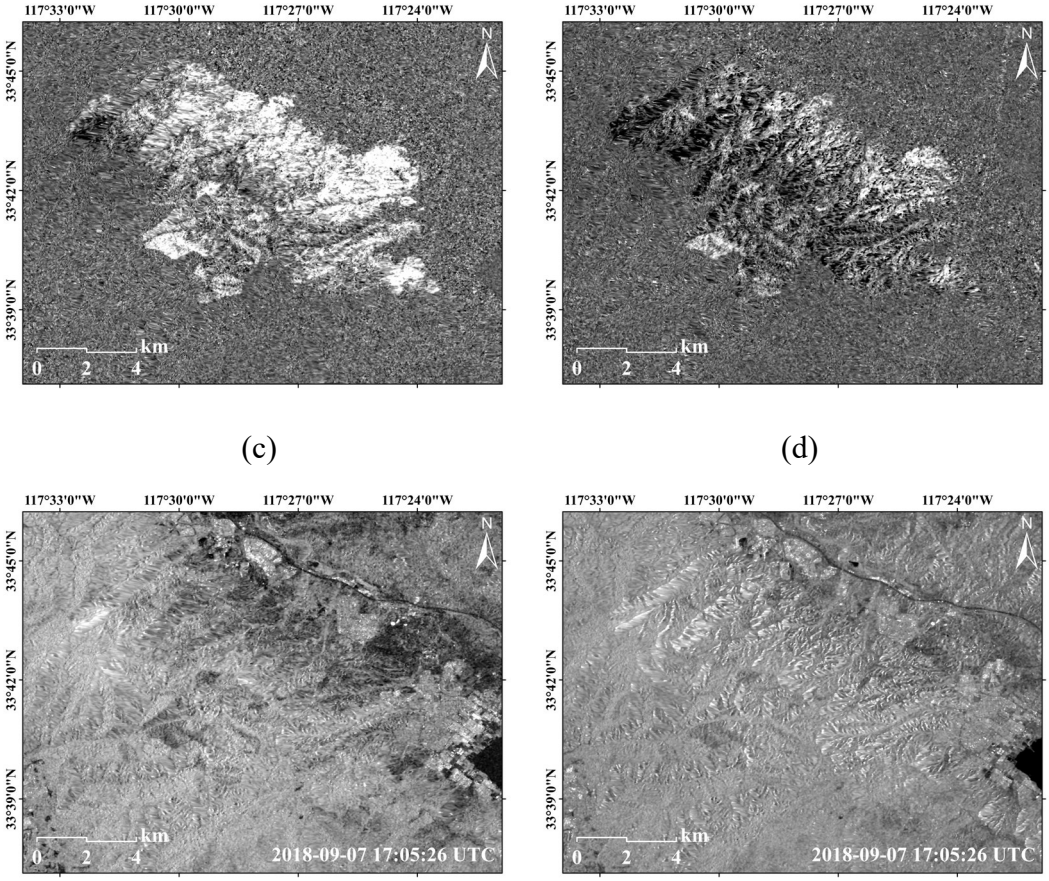


Figure 6.19 Backscatter coefficient: a) Pre-fire map in VV polarization, b) Pre-fire map in VH polarization, c) Difference map in VV polarization, d) Difference map in VH polarization, e) Post-fire map in VV polarization, f) Post-fire map in VH polarization.

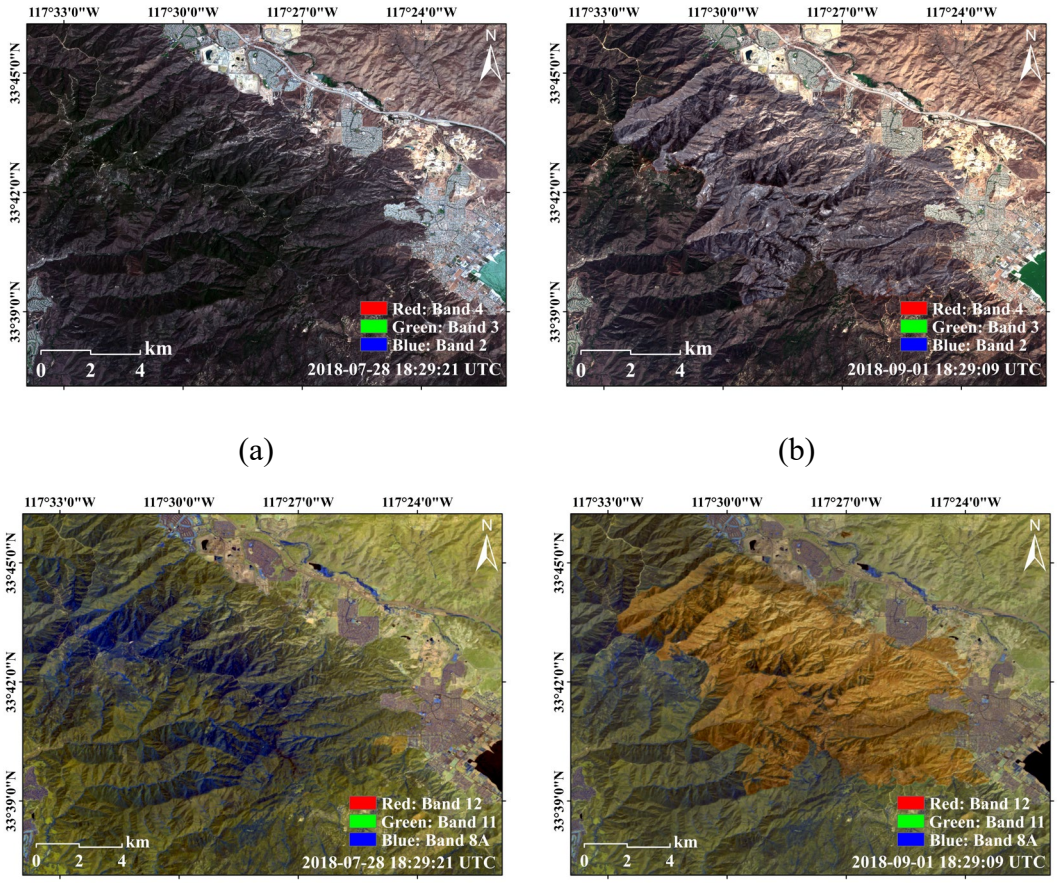


Figure 6.20 Surface reflectance: a) Pre-fire true-color image, b) Post-fire true-color image, c) Pre-fire false-color image, d) Post-fire false-color image

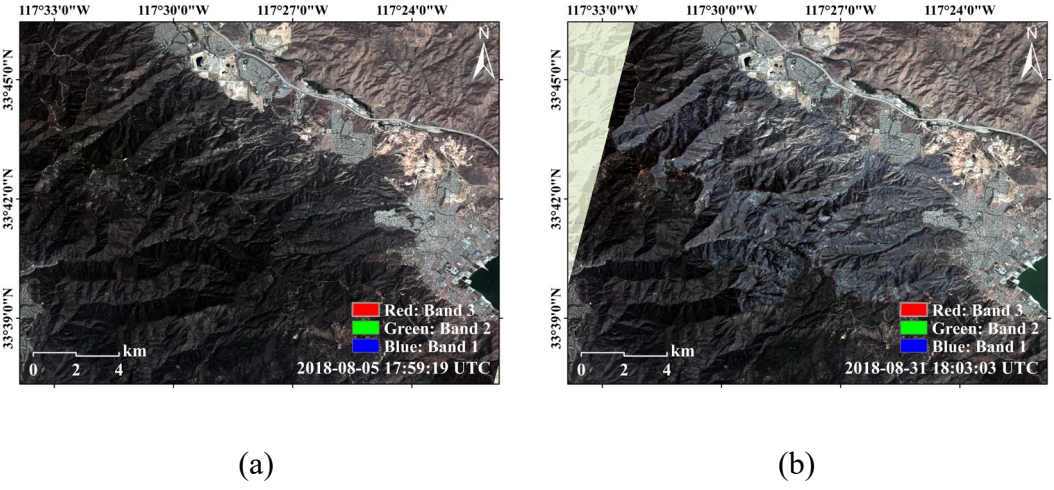


Figure 6.21 Reference data: a) Pre-fire true-color image, b) Post-fire true-color image.

The multi-source remote sensing data have their unique advantages in BA mapping. Compared with the SR and BS features, interferometric coherences are more sensitive to the temporal changes while containing fewer spatial details of the land covers. This is because interferometric coherence is a regional maximum likelihood estimation of the phase standard deviation (Moreira et al., 2013) and also the multi-looking process reduces the spatial details (Nielsen et al., 2020). SR images contain more details of the land cover. As indicated in Figure 6.20, the burned boundaries are more prominent in the SWIR-NIR composite than in the RGB image. Even so, there are still some shadow areas in the Holy site that are hard to tell whether they are burned or not. Luckily, these areas have lower values in the co-fire coherence which offers a useful auxiliary for the BA discrimination.

6.7 Conclusion

This chapter offers a deep-learning-based BA mapping method applied to a synergistic dataset constructed by Sentinel-1 and Sentinel-2 acquisitions. This process relies on the construction of an object dataset following by a supervised SSA classification. Experiment results evidenced that the performance of BA mapping can be improved through the synergy of SR and COR features and the SSA classification method surpasses the RF method on exploring the latent relations among multi-source features and hence the accuracy of BA mapping. The flexibility of the SSA classifier is further evaluated on the independent sites to investigate its potential to apply to the regional

or global fire perimeters. Related BA products generated in this research can be used as a powerful data resource for the governments and fire management agencies to inform post-fire planning and policy.

Chapter 7 Conclusion and Future Work

7.1 Conclusion

The main objective of this dissertation was defined in Chapter 1 as:

"The general goal of this thesis is the optimum exploitation of SAR, InSAR, PolSAR, PolInSAR, LiDAR, and multi-spectral remote sensing for forest characterization and monitoring from three main aspects including the forest height estimation, active fire detection, and burned area mapping."

First of all, contributions to the field of forest height inversion are from two aspects, namely, an improvement in the model-based forest height inversion using PolInSAR data and a proposal of deep-learning-based forest height inversion with a synergy of PolInSAR and LiDAR data.

For the first part, contributions include:

- Extended the RMoG model with homogeneous and heterogeneous scattering attenuation and random motion descriptions and evaluated their performance on forest canopy height mapping at the landscape scale.
- Balanced between the number of PolInSAR acquisitions and forest parameters in the physical model and deployed a dual-baseline repeat-pass PolInSAR data acquisition configuration to address the underdetermined estimation problem.

- Proposed a leveraging strategy in the optimization process to utilize advantages of homogenous and heterogeneous scattering attenuation and random motion properties in the physical model.

As to the second part, contributions include:

- Reformulated the forest height estimation as an image pan-sharpening process between high spatial resolution features derived from PolSAR and PolInSAR and high vertical accuracy height metrics from LiDAR.
- Proposed a specifically designed GAN called PolGAN with one generator and two discriminators (spatial and coherence) to accomplish the forest height pan-sharpening task.
- Deployed a progressive pan-sharpening strategy in the generator to cope with the larger scale factors between spaceborne LiDAR and SAR acquisitions.
- Evaluated the performance of PolGAN with standard or progressive generator in conditions of different downsampling strategies.

SAR data involved in this thesis are collected by UAVSAR as a proxy for the NISAR-like data. LiDAR data are acquired by LVIS, which can be used to simulate the GEDI-like data through downsampling. Therefore, studies on forest height estimation in this thesis will benefit the upcoming spaceborne SAR missions (NISAR, BIOMASS, and Tandem-L) and spaceborne LiDAR missions (GDEI and ICESat-1&2).

From the aspect of active fire detection, contributions include:

- Formulated the multispectral-based active fire detection as a binary semantic segmentation process and evaluated the detection accuracy of different deep-learning-based segmentation models.
- Adjusted the architecture of HRNetV2 with a DCPA network to make it more suitable for active fire detection.
- Constructed a dataset using Sentinel-2 TOA products collected over wildfires that occurred on the east coast of Australia and the west coast of the United States in 2019-2020 for the training and testing.
- Built an automated active fire detection framework for Sentinel-2 imagery as a cost-efficient resource in support of the global fire monitoring system.

The concept in this framework can be further applied to other remote sensing sensors with data acquisitions in SWIR-Red ranges and further serve as a powerful tool to deal with large volumes of high-resolution data for fire monitoring and as a cost-efficient resource in support of the governments and fire service agencies to deploy timely and optimized firefighting plans.

For burned area mapping, contributions include:

- Explored the potential of burned mapping with the synergy of SR, BS, and COR data derived from Sentinel-1 and Sentinel-2 acquisitions.
- Proposed an SSA classification strategy for multi-sensor feature extraction and burned area mapping.
- Constructed a multi-source dataset over 12 California wildfires that occurred in 2018-2019 at the object level for training and testing.

- Evaluated the performance of burned area mapping by different sites, feature sources, and classification strategies to appraise the improvements achieved by the proposed method.

Related burned area mapping products can be used as a powerful data resource for the governments and fire management agencies to evaluate biomass losses, manage post-fire policies, and prevent secondary disasters.

7.2 Future Work

Related studies on forest characterization and monitoring in this thesis are larger than single tree scales. Further efforts will be placed on the recognition and reconstruction of individual trees because global high-resolution remote sensing data are becoming available. Although the analysis of individual trees in remote sensing data has been a well-researched topic for quite a few years, most prior studies are based on the LiDAR, RGB-D, and photogrammetry data (Xie et al., 2020) whereas only a few studies have been conducted with SAR (Schmitt et al., 2015, Magnard et al., 2016). This is because generating high-density point clouds of forest canopies for individual tree recognition and reconstruction requires SAR acquisitions with spatial resolutions in the decimeter level and sharp focusing under most wind conditions and surface scattering. Moreover, sophisticated processing strategies have to be developed based on a thorough understanding of the radar imaging principle because the unique side-looking imaging geometry of SAR may result in severe height, displacement, and occlusion errors for individual trees due to layover and shadowing effects. However, compared with data acquired by LiDAR, RGB-D, and photogrammetry sensors, SAR acquisitions have the potential of wide-swath mapping in a short time and is less limited by the weather condition and solar illustration. Hence, it is necessary to further explore the potential of SAR imagery on the recognition and reconstruction of individual trees.

The most commonly used 3D forest characterization technique is TomoSAR (Reigber and Moreira, 2000). As an extension of multi-baseline SAR interferometry, TomoSAR creates a second synthetic aperture along the elevation direction to retrieve the vertical

scattering profile, which has been further applied to generate point clouds of the forest canopy (Schmitt et al., 2015). An alternative way for 3D forest characterization is the Polarization Coherence Tomography (PCT) (Cloude, 2006), which approximates the reflectivity profile as a weighted series of basis functions and estimates the associated parameters through several PolInSAR coherences. As an extension of PolInSAR, PCT has also been applied to the multi-baseline condition to improve the vertical resolution (Cloude, 2007a, Cloude, 2007b). This method has not been applied to the point cloud generation yet.

TomoSAR and PCT methods are normally applied to single-pass acquisitions to avoid the temporal decorrelation effect. However, in this condition, only a small number of acquisitions in irregular passes are available for the TomoSAR and PCT processing, which raises the difficulty of generating a point cloud with high vertical accuracy. One possible way to increase the number of acquisitions is to use repeat-pass data. To cope with the temporal decorrelation effect in TomoSAR, differential TomoSAR has been proposed by producing spatial and temporal signatures of the scatterers in a SAR cell and decoupling nuisance temporal components through spectral analysis (Lombardini and Cai, 2013, Aghababaee et al., 2019). As to PCT, a possible way is to simplify the dynamic forest canopy as random scatterers in Brownian motion and integrate it into the coherence function to compensate for the temporal decorrelation. Besides that, the fusion between multi-source cloud points can also be employed to improve the quality of cloud points (Zhang, 2010, Li et al., 2021). The generated 3D cloud points can be further processed by the traditional or deep-learning-based 3D shape classification, 3D

object detection and tracking, and 3D point cloud segmentation methods for different applications (Nguyen and Le, 2013, Guo et al., 2020b).

Reference

ABDALATI, W., ZWALLY, H. J., BINDSCHADLER, R., CSATHO, B., FARRELL, S. L., FRICKER, H. A., HARDING, D., KWOK, R., LEFSKY, M. & MARKUS, T. 2010. The ICESat-2 laser altimetry mission. *Proceedings of the IEEE,* 98**,** 735-751.

AGHABABAEE, H., FERRAIOLI, G. & SCHIRINZI, G. 2019. Differential SAR tomography reconstruction robust to temporal decorrelation effects. *IEEE transactions on geoscience and remote sensing,* 57**,** 9071-9080.

AGHABABAEI, H., FERRAIOLI, G., FERRO-FAMIL, L., HUANG, Y., D'ALESSANDRO, M. M., PASCAZIO, V., SCHIRINZI, G. & TEBALDINI, S. 2020. Forest SAR tomography: principles and applications. *IEEE geoscience and remote sensing magazine,* 8**,** 30-45.

AHMED, R., SIQUEIRA, P., HENSLEY, S., CHAPMAN, B. & BERGEN, K. 2011. A survey of temporal decorrelation from spaceborne L-Band repeat-pass InSAR. *Remote Sensing of Environment,* 115**,** 2887-2896.

AIAZZI, B., ALPARONE, L., BARONTI, S. & GARZELLI, A. 2002. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transactions on geoscience and remote sensing,* 40**,** 2300-2312.

ANDERSON, J., MARTIN, M., SMITH, M., DUBAYAH, R., HOFTON, M., HYDE, P., PETERSON, B., BLAIR, J. & KNOX, R. 2006. The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. *Remote Sensing of Environment,* 105**,** 248-261.

AUSTIN, K. F. 2020. Degradation and disease: Ecologically unequal exchanges cultivate emerging pandemics. *World Development,* 137**,** 105163.

BABOO, S. S. & DEVI, M. R. 2010. An analysis of different resampling methods in Coimbatore, District. *Global Journal of Computer Science and Technology*.

BADRINARAYANAN, V., KENDALL, A. & CIPOLLA, R. 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence,* 39**,** 2481-2495.

BAMLER, R. & HARTL, P. 1998. Synthetic aperture radar interferometry. *Inverse problems,* 14**,** R1.

BASTARRIKA, A., CHUVIECO, E. & MARTÍN, M. P. 2011. Mapping burned areas from Landsat TM/ETM+ data with a two-phase algorithm: Balancing omission and commission errors. *Remote Sensing of Environment,* 115**,** 1003-1012.

BELENGUER-PLOMER, M. A., TANASE, M. A., CHUVIECO, E. & BOVOLO, F. 2021. CNN-based burned area mapping using radar and optical data. *Remote Sensing of Environment,* 260**,** 112468.

BERNDES, G., ABT, B., ASIKAINEN, A., COWIE, A., DALE, V., EGNELL, G., LINDNER, M., MARELLI, L., PARÉ, D. & PINGOUD, K. 2016. Forest biomass, carbon neutrality and climate change mitigation. *From science to policy,* 3**,** 3-27.

BESSHO, K., DATE, K., HAYASHI, M., IKEDA, A., IMAI, T., INOUE, H., KUMAGAI, Y., MIYAKAWA, T., MURATA, H. & OHNO, T. 2016. An introduction to Himawari-8/9—Japan's new-generation geostationary

meteorological satellites. *Journal of the Meteorological Society of Japan. Ser. II,* 94**,** 151-183.
BINZAID, Y. I.-Y. S. 2020. GLOBAL WARMING, CO2, FOSSIL FUELS AND LIFE: A PHOTOTHERMAL ALTERNATIVE SOLUTION FOR SOLAR ENERGY ON-DEMAND. *LIFE SCIENCES LEAFLETS,* 130**,** 19-32.
BJORKMAN, A. D., MYERS-SMITH, I. H., ELMENDORF, S. C., NORMAND, S., RÜGER, N., BECK, P. S., BLACH-OVERGAARD, A., BLOK, D., CORNELISSEN, J. H. C. & FORBES, B. C. 2018. Plant functional trait change across a warming tundra biome. *Nature,* 562**,** 57-62.
BLAIR, J. B., RABINE, D. L. & HOFTON, M. A. 1999. The Laser Vegetation Imaging Sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing,* 54**,** 115-122.
BOUZINAC, C., LAFRANCE, B., PESSIOT, L., TOULI, D., JUNG, M., MASSERA, S., NEVEU-VANMALLE, M., ESPESSET, A., FRANCESCONI, B. & CLERC, S. Sentinel-2 level-1 calibration and validation status from the mission performance centre. IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, 2018. IEEE, 4347-4350.
BOYKOV, Y., VEKSLER, O. & ZABIH, R. 2001. Fast approximate energy minimization via graph cuts. *IEEE Transactions on pattern analysis and machine intelligence,* 23**,** 1222-1239.
BREIMAN, L. 2001. Random forests. *Machine learning,* 45**,** 5-32.
BRODU, N. 2017. Super-resolving multiresolution images with band-independent geometry of multispectral pixels. *IEEE Transactions on Geoscience and Remote Sensing,* 55**,** 4610-4617.
BROMLEY, J., GUYON, I., LECUN, Y., SÄCKINGER, E. & SHAH, R. Signature verification using a" siamese" time delay neural network. Advances in neural information processing systems, 1994. 737-744.
BYEON, W., BREUEL, T. M., RAUE, F. & LIWICKI, M. Scene labeling with lstm recurrent neural networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015. 3547-3555.
CAI, J. & HUANG, B. 2020. Super-resolution-guided progressive pansharpening based on a deep convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing,* 59**,** 5206-5220.
CALFIRE 2018. 2018 Statistics and Events. https://www.fire.ca.gov/stats-events/ (accessed 21.04.20)
CAO, X., CHEN, J., IMURA, H. & HIGASHI, O. 2009. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. *Remote Sensing of Environment,* 113**,** 2205-2209.
CARREIRAS, J. M., QUEGAN, S., LE TOAN, T., MINH, D. H. T., SAATCHI, S. S., CARVALHAIS, N., REICHSTEIN, M. & SCIPAL, K. 2017. Coverage of high biomass forests by the ESA BIOMASS mission under defense restrictions. *Remote Sensing of Environment,* 196**,** 154-162.
CHANG, D. & SONG, Y. 2009. Comparison of L3JRC and MODIS global burned area products from 2000 to 2007. *Journal of Geophysical Research: Atmospheres,* 114.
CHAPMAN, B., SIQUEIRA, P., SAATCHI, S., SIMARD, M. & KELLNDORFER, J. Initial results from the 2019 NISAR Ecosystem Cal/Val Exercise in the SE

USA. IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, 2019. IEEE, 8641-8644.
CHEN, B., SHU, H., COATRIEUX, G., CHEN, G., SUN, X. & COATRIEUX, J. L. 2015. Color image analysis by quaternion-type moments. *Journal of mathematical imaging and vision,* 51**,** 124-144.
CHEN, L.-C., PAPANDREOU, G., KOKKINOS, I., MURPHY, K. & YUILLE, A. L. 2014. Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv preprint arXiv:1412.7062*.
CHEN, L.-C., PAPANDREOU, G., KOKKINOS, I., MURPHY, K. & YUILLE, A. L. 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence,* 40**,** 834-848.
CHEN, L.-C., ZHU, Y., PAPANDREOU, G., SCHROFF, F. & ADAM, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the European conference on computer vision (ECCV), 2018. 801-818.
CHUVIECO, E., LIZUNDIA-LOIOLA, J., PETTINARI, M. L., RAMO, R., PADILLA, M., TANSEY, K., MOUILLOT, F., LAURENT, P., STORM, T. & HEIL, A. 2018. Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies. *Earth System Science Data,* 10**,** 2015-2031.
CHUVIECO, E., MOUILLOT, F., VAN DER WERF, G. R., SAN MIGUEL, J., TANASE, M., KOUTSIAS, N., GARCÍA, M., YEBRA, M., PADILLA, M. & GITAS, I. 2019. Historical background and current developments for mapping burned area from satellite Earth observation. *Remote Sensing of Environment,* 225**,** 45-64.
CHUVIECO, E., OPAZO, S., SIONE, W., VALLE, H. D., ANAYA, J., BELLA, C. D., CRUZ, I., MANZO, L., LÓPEZ, G. & MARI, N. 2008. Global burned ‐ land estimation in Latin America using MODIS composite data. *Ecological Applications,* 18**,** 64-79.
CLOUDE, S. 2007a. Multibaseline polarization coherence tomography. *Science and Applications of SAR Polarimetry and Polarimetric Interferometry,* 644**,** 8.
CLOUDE, S. & PAPATHANASSIOU, K. 2003. Three-stage inversion process for polarimetric SAR interferometry. *IEE Proceedings-Radar, Sonar and Navigation,* 150**,** 125-134.
CLOUDE, S. R. Robust parameter estimation using dual baseline polarimetric SAR interferometry. IEEE International Geoscience and Remote Sensing Symposium, 2002. IEEE, 838-840.
CLOUDE, S. R. 2006. Polarization coherence tomography. *Radio Science,* 41.
CLOUDE, S. R. 2007b. Dual-baseline coherence tomography. *IEEE Geoscience and Remote Sensing Letters,* 4**,** 127-131.
CLOUDE, S. R. & POTTIER, E. 1996. A review of target decomposition theorems in radar polarimetry. *IEEE transactions on geoscience and remote sensing,* 34**,** 498-518.
CLOUDE, S. R. & POTTIER, E. 1997. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE transactions on geoscience and remote sensing,* 35**,** 68-78.

COLSON, D., PETROPOULOS, G. P. & FERENTINOS, K. P. 2018. Exploring the potential of Sentinels-1 & 2 of the Copernicus Mission in support of rapid and cost-effective wildfire assessment. *International journal of applied earth observation and geoinformation,* 73**,** 262-276.

ÇÖMERT, R., MATCI, D. K. & AVDAN, U. 2019. Object Based Burned Area Mapping with Random Forest Algorithm. *International Journal of Engineering and Geosciences,* 4**,** 78-87.

CRESPO-PEREMARCH, P., FOURNIER, R. A., NGUYEN, V.-T., VAN LIER, O. R. & RUIZ, L. Á. 2020. A comparative assessment of the vertical distribution of forest components using full-waveform airborne, discrete airborne and discrete terrestrial laser scanning data. *Forest Ecology and Management,* 473**,** 118268.

CRESWELL, A., WHITE, T., DUMOULIN, V., ARULKUMARAN, K., SENGUPTA, B. & BHARATH, A. A. 2018. Generative adversarial networks: An overview. *IEEE Signal Processing Magazine,* 35**,** 53-65.

CUI, X., GUO, X., WANG, Y., WANG, X., ZHU, W., SHI, J., LIN, C. & GAO, X. 2019. Application of remote sensing to water environmental processes under a changing climate. *Journal of Hydrology,* 574**,** 892-902.

CURLANDER, J. C. & MCDONOUGH, R. N. 1991. *Synthetic aperture radar*, Wiley, New York.

DARGAN, S., KUMAR, M., AYYAGARI, M. R. & KUMAR, G. 2020. A survey of deep learning and its applications: a new paradigm to machine learning. *Archives of Computational Methods in Engineering,* 27**,** 1071-1092.

DE ALMEIDA PEREIRA, G. H., FUSIOKA, A. M., NASSU, B. T. & MINETTO, R. 2021. Active fire detection in Landsat-8 imagery: A large-scale dataset and a deep-learning study. *ISPRS Journal of Photogrammetry and Remote Sensing,* 178**,** 171-186.

DENBINA, M. & SIMARD, M. Kapok: An open source Python library for PolInSAR forest height estimation using UAVSAR data. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2017. IEEE, 4314-4317.

DENBINA, M., SIMARD, M. & HAWKINS, B. 2018. Forest height estimation using multibaseline PolInSAR and sparse lidar data fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11**,** 3415-3433.

DHANACHANDRA, N., MANGLEM, K. & CHANU, Y. J. 2015. Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science,* 54**,** 764-771.

DOU, W., CHEN, Y., LI, X. & SUI, D. Z. 2007. A general framework for component substitution image fusion: An implementation using the fast image fusion method. *Computers & Geosciences,* 33**,** 219-228.

DRAGOZI, E., GITAS, I. Z., STAVRAKOUDIS, D. G. & THEOCHARIS, J. B. 2014. Burned area mapping using support vector machines and the FuzCoC feature selection method on VHR IKONOS imagery. *Remote Sensing,* 6**,** 12005-12036.

DUBAYAH, R., BLAIR, J. B., GOETZ, S., FATOYINBO, L., HANSEN, M., HEALEY, S., HOFTON, M., HURTT, G., KELLNER, J. & LUTHCKE, S. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of remote sensing,* 1**,** 100002.

DUNCANSON, L., NEUENSCHWANDER, A., HANCOCK, S., THOMAS, N., FATOYINBO, T., SIMARD, M., SILVA, C. A., ARMSTON, J., LUTHCKE, S. B. & HOFTON, M. 2020. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sensing of Environment,* 242**,** 111779.

EL MOUSSAWI, I., HO TONG MINH, D., BAGHDADI, N., ABDALLAH, C., JOMAAH, J., STRAUSS, O. & LAVALLE, M. 2019. L-band UAVSAR tomographic imaging in dense forests: Gabon forests. *Remote Sensing,* 11**,** 475.

ELL, T. A. & SANGWINE, S. J. 2006. Hypercomplex Fourier transforms of color images. *IEEE Transactions on image processing,* 16**,** 22-35.

ENGLHART, S., KEUCK, V. & SIEGERT, F. 2011. Aboveground biomass retrieval in tropical forests—The potential of combined X-and L-band SAR data use. *Remote sensing of environment,* 115**,** 1260-1271.

FAO, U. 2020. The State of the World's Forests 2020. In brief. Forests, biodiversity and people.: Rome.

FASBENDER, D., RADOUX, J. & BOGAERT, P. 2008. Bayesian data fusion for adaptable image pansharpening. *IEEE Transactions on Geoscience and Remote Sensing,* 46**,** 1847-1857.

FASSNACHT, F. E., SCHMIDT-RIESE, E., KATTENBORN, T. & HERNÁNDEZ, J. 2021. Explaining Sentinel 2-based dNBR and RdNBR variability with reference data from the bird's eye (UAS) perspective. *International Journal of Applied Earth Observation and Geoinformation,* 95**,** 102262.

FATOYINBO, T., ARMSTON, J., SIMARD, M., SAATCHI, S., DENBINA, M., LAVALLE, M., HOFTON, M., TANG, H., MARSELIS, S. & PINTO, N. 2021. The NASA AfriSAR campaign: Airborne SAR and lidar measurements of tropical forest structure and biomass in support of current and future space missions. *Remote Sensing of Environment,* 264**,** 112533.

FERRETTI, A., PRATI, C. & ROCCA, F. 2000. Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. *IEEE Transactions on geoscience and remote sensing,* 38**,** 2202-2212.

FERRETTI, A., PRATI, C. & ROCCA, F. 2001. Permanent scatterers in SAR interferometry. *IEEE Transactions on geoscience and remote sensing,* 39**,** 8-20.

FERRO-FAMIL, L. & POTTIER, E. 2016. Synthetic Aperture Radar Imaging. *Microwave Remote Sensing of Land Surface.* Elsevier.

FILIPPONI, F. Sentinel-1 GRD preprocessing workflow. Multidisciplinary Digital Publishing Institute Proceedings, 2019. 11.

FITCH, J. P. 2012. *Synthetic aperture radar*, Springer Science & Business Media.

FORE, A. G., CHAPMAN, B. D., HAWKINS, B. P., HENSLEY, S., JONES, C. E., MICHEL, T. R. & MUELLERSCHOEN, R. J. 2015. UAVSAR polarimetric calibration. *IEEE Transactions on Geoscience and Remote Sensing,* 53**,** 3481-3491.

FREEMAN, A. & DURDEN, S. L. 1998. A three-component scattering model for polarimetric SAR data. *IEEE transactions on geoscience and remote sensing,* 36**,** 963-973.

FU, J., LIU, J., TIAN, H., LI, Y., BAO, Y., FANG, Z. & LU, H. Dual attention network for scene segmentation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 3146-3154.

FU, W., GUO, H., SONG, P., TIAN, B., LI, X. & SUN, Z. 2017. Combination of PolInSAR and LiDAR techniques for forest height estimation. *IEEE Geoscience and Remote Sensing Letters,* 14**,** 1218-1222.

FURLAUD, J. M., WILLIAMSON, G. J. & BOWMAN, D. M. 2018. Simulating the effectiveness of prescribed burning at altering wildfire behaviour in Tasmania, Australia. *International Journal of Wildland Fire,* 27**,** 15-28.

GARESTIER, F., DUBOIS-FERNANDEZ, P. C. & CHAMPION, I. 2008a. Forest height inversion using high-resolution P-band Pol-InSAR data. *IEEE Transactions on Geoscience and Remote Sensing,* 46**,** 3544-3559.

GARESTIER, F., DUBOIS-FERNANDEZ, P. C. & PAPATHANASSIOU, K. P. 2008b. Pine forest height inversion using single-pass X-band PolInSAR data. *IEEE Transactions on Geoscience and Remote Sensing,* 46**,** 59-68.

GARESTIER, F. & LE TOAN, T. Vegetation modelling for height inversion using InSAR/Pol-InSAR data. 2007 IEEE International Geoscience and Remote Sensing Symposium, 2007. IEEE, 2322-2325.

GARESTIER, F. & LE TOAN, T. 2009. Forest modeling for height inversion using single-baseline InSAR/Pol-InSAR data. *IEEE Transactions on Geoscience and Remote Sensing,* 48**,** 1528-1539.

GARGIULO, M., DELL'AGLIO, D. A. G., IODICE, A., RICCIO, D. & RUELLO, G. A CNN-Based Super-Resolution Technique for Active Fire Detection on Sentinel-2 Data. 2019 PhotonIcs & Electromagnetics Research Symposium-Spring (PIERS-Spring), 2019. IEEE, 418-426.

GENZANO, N., PERGOLA, N. & MARCHESE, F. 2020. A Google Earth Engine tool to investigate, map and monitor volcanic thermal anomalies at global scale by means of mid-high spatial resolution satellite data. *Remote Sensing,* 12**,** 3232.

GHAHREMANI, M. & GHASSEMIAN, H. 2015. Remote-sensing image fusion based on curvelets and ICA. *International Journal of Remote Sensing,* 36**,** 4131-4143.

GHASEMI, N., TOLPEKIN, V. & STEIN, A. 2018. A modified model for estimating tree height from PolInSAR with compensation for temporal decorrelation. *International journal of applied earth observation and geoinformation,* 73**,** 313-322.

GHIASI, G. & FOWLKES, C. C. Laplacian pyramid reconstruction and refinement for semantic segmentation. European conference on computer vision, 2016. Springer, 519-534.

GIGLIO, L., BOSCHETTI, L., ROY, D. P., HUMBER, M. L. & JUSTICE, C. O. 2018. The Collection 6 MODIS burned area mapping algorithm and product. *Remote sensing of environment,* 217**,** 72-85.

GIGLIO, L., CSISZAR, I., RESTÁS, Á., MORISETTE, J. T., SCHROEDER, W., MORTON, D. & JUSTICE, C. O. 2008. Active fire detection and characterization with the advanced spaceborne thermal emission and reflection radiometer (ASTER). *Remote sensing of environment,* 112**,** 3055-3063.

GIGLIO, L., SCHROEDER, W. & JUSTICE, C. O. 2016. The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment,* 178**,** 31-41.

GIRI, C., OCHIENG, E., TIESZEN, L. L., ZHU, Z., SINGH, A., LOVELAND, T., MASEK, J. & DUKE, N. 2011. Status and distribution of mangrove forests of

the world using earth observation satellite data. *Global Ecology and Biogeography,* 20**,** 154-159.
GOLDSTEIN, R. M. & ZEBKER, H. 1987. Interferometric radar measurement of ocean surface currents. *Nature,* 328**,** 707-709.
GOMEZ, A. N., REN, M., URTASUN, R. & GROSSE, R. B. The reversible residual network: Backpropagation without storing activations. Advances in neural information processing systems, 2017. 2214-2224.
GOODFELLOW, I., POUGET-ABADIE, J., MIRZA, M., XU, B., WARDE-FARLEY, D., OZAIR, S., COURVILLE, A. & BENGIO, Y. 2014. Generative adversarial nets. *Advances in neural information processing systems,* 27.
GOSWAMI, B., KUMAR, K. M., SINGH, R. K., KUMAR, R., VERMA, S. K., KUMARI, P. & SHARMA, G. S. 2020. Remote sensing in forestry. *International Journal of Computational Biology and Bioinformatics,* 6**,** 10-17.
GREAVES, H. E., VIERLING, L. A., EITEL, J. U., BOELMAN, N. T., MAGNEY, T. S., PRAGER, C. M. & GRIFFIN, K. L. 2016. High-resolution mapping of aboveground shrub biomass in Arctic tundra using airborne lidar and imagery. *Remote sensing of environment,* 184**,** 361-373.
GUO, Q., SU, Y., HU, T., GUAN, H., JIN, S., ZHANG, J., ZHAO, X., XU, K., WEI, D. & KELLY, M. 2020a. Lidar Boosts 3D Ecological Observations and Modelings: A review and perspective. *IEEE Geoscience and Remote Sensing Magazine***,** 0-0.
GUO, Y., WANG, H., HU, Q., LIU, H., LIU, L. & BENNAMOUN, M. 2020b. Deep learning for 3d point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence*.
HAJNSEK, I., KUGLER, F., LEE, S.-K. & PAPATHANASSIOU, K. P. 2009. Tropical-forest-parameter estimation by means of Pol-InSAR: The INDREX-II campaign. *IEEE transactions on Geoscience and Remote Sensing,* 47**,** 481-493.
HALL, J. V., ZHANG, R., SCHROEDER, W., HUANG, C. & GIGLIO, L. 2019. Validation of GOES-16 ABI and MSG SEVIRI active fire products. *International Journal of Applied Earth Observation and Geoinformation,* 83**,** 101928.
HANCOCK, S., ARMSTON, J., HOFTON, M., SUN, X., TANG, H., DUNCANSON, L. I., KELLNER, J. R. & DUBAYAH, R. 2019. The GEDI simulator: A large ‐ footprint waveform lidar simulator for calibration and validation of spaceborne missions. *Earth and Space Science,* 6**,** 294-310.
HANTSON, S., PADILLA, M., CORTI, D. & CHUVIECO, E. 2013. Strengths and weaknesses of MODIS hotspots to characterize global fire occurrence. *Remote Sensing of Environment,* 131**,** 152-159.
HARDIMAN, B. S., BOHRER, G., GOUGH, C. M., VOGEL, C. S. & CURTIS, P. S. 2011. The role of canopy structural complexity in wood net primary production of a maturing northern deciduous forest. *Ecology,* 92**,** 1818-1827.
HE, K., ZHANG, X., REN, S. & SUN, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 770-778.
HENSLEY, S., AHMED, R., CHAPMAN, B., HAWKINS, B., LAVALLE, M., PINTO, N., PARDINI, M., PAPATHANASSIOU, K., SIQUERIA, P. & TREUHAFT, R. A Comparison of L-band and S-band Interferometry and

Tomography of the BERMS Borel Forest with UAVSAR and F-SAR Datasets. EUSAR 2021; 13th European Conference on Synthetic Aperture Radar, 2021. VDE, 1-4.
HENSLEY, S., MICHEL, T., NEUMANN, M., LAVALLE, M., AHMED, R., MUELLERSCHOEN, R. & CHAPMAN, B. A Comparison of Multi-Baseline Polarimetric Inteferometry at La Amistad and La Selva, Costa Rica with a Modified PolSARProSim Scattering Tool. EUSAR 2014; 10th European Conference on Synthetic Aperture Radar, 2014. VDE, 1-4.
HENSLEY, S., MICHEL, T., NUEMANN, M., LAVALLE, M., MUELLERSCHOEN, R., CHAPMAN, B., JONES, C., AHMED, R., LOMBARDINI, F. & SIQUEIRA, P. Some first polarimetric-interferometric multi-baseline and tomographic results at Harvard forest using UAVSAR. 2012 IEEE International Geoscience and Remote Sensing Symposium, 2012. IEEE, 5202-5205.
HENSLEY, S., WHEELER, K., SADOWY, G., JONES, C., SHAFFER, S., ZEBKER, H., MILLER, T., HEAVEY, B., CHUANG, E. & CHAO, R. The UAVSAR instrument: Description and first results. 2008 IEEE Radar Conference, 2008. IEEE, 1-6.
HIGUERA, P. E. & ABATZOGLOU, J. T. 2021. Record‐setting climate enabled the extraordinary 2020 fire season in the western United States. *Global change biology,* 27**,** 1-2.
HIJMANS, R. J., CAMERON, S. E., PARRA, J. L., JONES, P. G. & JARVIS, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society,* 25**,** 1965-1978.
HU, J., SHEN, L. & SUN, G. Squeeze-and-excitation networks. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018. 7132-7141.
HU, X., BAN, Y. & NASCETTI, A. 2021. Sentinel-2 MSI data for active fire detection in major fire-prone biomes: A multi-criteria approach. *International Journal of Applied Earth Observation and Geoinformation,* 101**,** 102347.
HUA, L. & SHAO, G. 2017. The progress of operational forest fire monitoring with infrared remote sensing. *Journal of forestry research,* 28**,** 215-229.
HUANG, X. & PAN, W. 2003. Linear regression and two-class classification with gene expression data. *Bioinformatics,* 19**,** 2072-2078.
IANDOLA, F., MOSKEWICZ, M., KARAYEV, S., GIRSHICK, R., DARRELL, T. & KEUTZER, K. 2014. Densenet: Implementing efficient convnet descriptor pyramids. *arXiv preprint arXiv:1404.1869*.
JONES, M. W., SMITH, A., BETTS, R., CANADELL, J. G., PRENTICE, I. C. & LE QUÉRÉ, C. 2020. Climate change increases risk of wildfires. *ScienceBrief Review*.
KASS, M., WITKIN, A. & TERZOPOULOS, D. 1988. Snakes: Active contour models. *International journal of computer vision,* 1**,** 321-331.
KEMTER, M., FISCHER, M., LUNA, L., SCHÖNFELDT, E., VOGEL, J., BANERJEE, A., KORUP, O. & THONICKE, K. 2021. Cascading hazards in the aftermath of Australia's 2019/2020 Black Summer wildfires. *Earth's Future***,** e2020EF001884.
KETKAR, N. 2017. Introduction to pytorch. *Deep learning with python.* Springer.

KHAN, S., NASEER, M., HAYAT, M., ZAMIR, S. W., KHAN, F. S. & SHAH, M. 2021. Transformers in vision: A survey. *arXiv preprint arXiv:2101.01169*.

KIAT, P. E., MALEK, M. & SHAMSUDDIN, S. 2020. Net carbon stocks change in biomass from wood removal of tropical forests in Sarawak, Malaysia. *Journal of King Saud University-Science,* 32**,** 1096-1099.

KILLIAN, D. 2020. Black Summer in the Shoalhaven, NSW: what was it like and what could we do better? A resident's perspective. *Public Health Research & Practice,* 30.

KINGMA, D. P. & BA, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

KROGAGER, E. 1990. New decomposition of the radar target scattering matrix. *Electronics Letters,* 26**,** 1525-1527.

KUGLER, F., LEE, S.-K., HAJNSEK, I. & PAPATHANASSIOU, K. P. 2015. Forest height estimation by means of Pol-InSAR data inversion: The role of the vertical wavenumber. *IEEE Transactions on Geoscience and Remote Sensing,* 53**,** 5294-5311.

KUMAR, S. S. & ROY, D. P. 2017. Global operational land imager Landsat-8 reflectance-based active fire detection algorithm. *International Journal of Digital Earth,* 11**,** 154-178.

KUMAR, S. S. & ROY, D. P. 2018. Global operational land imager Landsat-8 reflectance-based active fire detection algorithm. *International Journal of Digital Earth,* 11**,** 154-178.

KWARTENG, P. & CHAVEZ, A. 1989. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. *Photogramm. Eng. Remote Sens,* 55**,** 339-348.

LAGADRILLIERE, P.-A., NOLDE, M., RIEDLINGER, T., RUPPERT, T. & SCHWINGER, M. DLR new functionalities for the European Forest Fire Information System. Proceedings of the International Astronautical Congress, IAC, 2019. 1-7.

LAI, W.-S., HUANG, J.-B., AHUJA, N. & YANG, M.-H. 2018. Fast and accurate image super-resolution with deep laplacian pyramid networks. *IEEE transactions on pattern analysis and machine intelligence,* 41**,** 2599-2613.

LAURENT, T. & BRECHT, J. The multilinear structure of ReLU networks. International Conference on Machine Learning, 2018. PMLR, 2908-2916.

LAVALLE, M. 2009. *Full and compact polarimetric radar interferometry for vegetation remote sensing.* Université Rennes 1.

LAVALLE, M., HAWKINS, B. & HENSLEY, S. Tomographic imaging with UAVSAR: Current status and new results from the 2016 AfriSAR campaign. 2017 IEEE international geoscience and remote sensing symposium (IGARSS), 2017. IEEE, 2485-2488.

LAVALLE, M. & HENSLEY, S. Demonstration of repeat-pass POLINSAR using UAVSAR: The RMOG model. Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, 2012. IEEE, 5876-5879.

LAVALLE, M. & HENSLEY, S. 2015. Extraction of structural and dynamic properties of forests from polarimetric-interferometric SAR data affected by temporal decorrelation. *IEEE Transactions on Geoscience and Remote Sensing,* 53**,** 4752-4767.

LAVALLE, M., SIMARD, M. & HENSLEY, S. 2012. A temporal decorrelation model for polarimetric radar interferometers. *IEEE Transactions on Geoscience and Remote Sensing,* 50**,** 2880-2888.

LECHNER, A. M., FOODY, G. M. & BOYD, D. S. 2020. Applications in Remote Sensing to Forest Ecology and Management. *One Earth,* 2**,** 405-412.

LECUN, Y. & BENGIO, Y. 1995. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks,* 3361**,** 1995.

LECUN, Y., BENGIO, Y. & HINTON, G. 2015. Deep learning. *nature,* 521**,** 436-444.

LEE, J.-S., GRUNES, M. R., AINSWORTH, T. L., DU, L.-J., SCHULER, D. L. & CLOUDE, S. R. 1999. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier. *IEEE Transactions on Geoscience and Remote Sensing,* 37**,** 2249-2258.

LEE, J.-S. & POTTIER, E. 2017. *Polarimetric radar imaging: from basics to applications*, CRC press.

LEE, S.-K., KUGLER, F., PAPATHANASSIOU, K. & HAJNSEK, I. Quantification and compensation of temporal decorrelation effects in polarimetric SAR interferometry. Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, 2012. IEEE, 3106-3109.

LEI, Y. & SIQUEIRA, P. 2014. Estimation of forest height using spaceborne repeat-pass L-band InSAR correlation magnitude over the US state of Maine. *Remote Sensing,* 6**,** 10252-10285.

LEI, Y., SIQUEIRA, P. & TREUHAFT, R. 2017. A physical scattering model of repeat-pass InSAR correlation for vegetation. *Waves in Random and Complex Media,* 27**,** 129-152.

LI, H., XIONG, P., AN, J. & WANG, L. 2018. Pyramid attention network for semantic segmentation. *arXiv preprint arXiv:1805.10180*.

LI, J., LEVINE, M. D., AN, X., XU, X. & HE, H. 2012. Visual saliency based on scale-space analysis in the frequency domain. *IEEE transactions on pattern analysis and machine intelligence,* 35**,** 996-1010.

LI, S., GE, X., HU, H. & ZHU, Q. 2021. Laplacian fusion approach of multi-source point clouds for detail enhancement. *ISPRS Journal of Photogrammetry and Remote Sensing,* 171**,** 385-396.

LIANG, X., KUKKO, A., HYYPPÄ, J., LEHTOMÄKI, M., PYÖRÄLÄ, J., YU, X., KAARTINEN, H., JAAKKOLA, A. & WANG, Y. 2018. In-situ measurements from mobile platforms: An emerging approach to address the old challenges associated with forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing,* 143**,** 97-107.

LIANG, X., SHEN, X., FENG, J., LIN, L. & YAN, S. Semantic object parsing with graph lstm. European Conference on Computer Vision, 2016. Springer, 125-143.

LIN, T.-Y., DOLLÁR, P., GIRSHICK, R., HE, K., HARIHARAN, B. & BELONGIE, S. Feature pyramid networks for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. 2117-2125.

LIN, Z., CHEN, F., NIU, Z., LI, B., YU, B., JIA, H. & ZHANG, M. 2018. An active fire detection algorithm based on multi-temporal FengYun-3C VIRR data. *Remote Sensing of Environment,* 211**,** 376-387.

LIU, L., WEN, X., GONZALEZ, A., TAN, D., DU, J., LIANG, Y., LI, W., FAN, D., SUN, K. & DONG, P. 2011. An object-oriented daytime land-fog-detection

approach based on the mean-shift and full lambda-schedule algorithms using EOS/MODIS data. *International journal of remote sensing,* 32**,** 4769-4785.
LIU, Q., ZHOU, H., XU, Q., LIU, X. & WANG, Y. 2020a. PSGAN: A generative adversarial network for remote sensing image pan-sharpening. *IEEE Transactions on Geoscience and Remote Sensing*.
LIU, X., LIU, Q. & WANG, Y. 2020b. Remote sensing image fusion based on two-stream fusion network. *Information Fusion,* 55**,** 1-15.
LIU, Z., LI, X., LUO, P., LOY, C.-C. & TANG, X. Semantic image segmentation via deep parsing network. Proceedings of the IEEE international conference on computer vision, 2015. 1377-1385.
LIZUNDIA-LOIOLA, J., OTÓN, G., RAMO, R. & CHUVIECO, E. 2020. A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data. *Remote Sensing of Environment,* 236**,** 111493.
LO, S.-C. B., CHAN, H.-P., LIN, J.-S., LI, H., FREEDMAN, M. T. & MUN, S. K. 1995. Artificial convolution neural network for medical image pattern recognition. *Neural networks,* 8**,** 1201-1214.
LOMBARDINI, F. & CAI, F. 2013. Temporal decorrelation-robust SAR tomography. *IEEE Transactions on Geoscience and Remote Sensing,* 52**,** 5412-5421.
LONG, J., SHELHAMER, E. & DARRELL, T. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 3431-3440.
LOUIS, J., DEBAECKER, V., PFLUG, B., MAIN-KNORN, M., BIENIARZ, J., MUELLER-WILM, U., CADAU, E. & GASCON, F. Sentinel-2 Sen2Cor: L2A processor for users. Proceedings Living Planet Symposium 2016, 2016. Spacebooks Online, 1-8.
LUO, L., WANG, X., GUO, H., LASAPONARA, R., ZONG, X., MASINI, N., WANG, G., SHI, P., KHATTELI, H. & CHEN, F. 2019. Airborne and spaceborne remote sensing for archaeological and cultural heritage applications: A review of the century (1907–2017). *Remote sensing of environment,* 232**,** 111280.
MA, J., YU, W., CHEN, C., LIANG, P., GUO, X. & JIANG, J. 2020. Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion. *Information Fusion,* 62**,** 110-120.
MAGNARD, C., MORSDORF, F., SMALL, D., STILLA, U., SCHAEPMAN, M. E. & MEIER, E. 2016. Single tree identification using airborne multibaseline SAR interferometry data. *Remote Sensing of Environment,* 186**,** 567-580.
MAIN-KNORN, M., PFLUG, B., LOUIS, J., DEBAECKER, V., MÜLLER-WILM, U. & GASCON, F. Sen2Cor for sentinel-2. Image and Signal Processing for Remote Sensing XXIII, 2017. International Society for Optics and Photonics, 1042704.
MANGEON, S., FIELD, R., FROMM, M., MCHUGH, C. & VOULGARAKIS, A. 2016. Satellite versus ground-based estimates of burned area: A comparison between MODIS based burned area and fire agency reports over North America in 2007. *The Anthropocene Review,* 3**,** 76-92.
MARCHESE, F., GENZANO, N., NERI, M., FALCONIERI, A., MAZZEO, G. & PERGOLA, N. 2019. A Multi-channel algorithm for mapping volcanic thermal anomalies by means of Sentinel-2 MSI and Landsat-8 OLI data. *Remote Sensing,* 11**,** 2876.

MARTINO, A. J., NEUMANN, T. A., KURTZ, N. T. & MCLENNAN, D. ICESat-2 mission overview and early performance. Sensors, Systems, and Next-Generation Satellites XXIII, 2019. International Society for Optics and Photonics, 111510C.

MASEK, J. G., WULDER, M. A., MARKHAM, B., MCCORKEL, J., CRAWFORD, C. J., STOREY, J. & JENSTROM, D. T. 2020. Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment,* 248**,** 111968.

MASI, G., COZZOLINO, D., VERDOLIVA, L. & SCARPA, G. 2016. Pansharpening by convolutional neural networks. *Remote Sensing,* 8**,** 594.

METSARANTA, J. M. & LIEFFERS, V. J. 2008a. A fifty-year reconstruction of annual changes in the spatial distribution of Pinus banksiana stands: does pattern fit competition theory? *Plant Ecology,* 199**,** 137-152.

METSARANTA, J. M. & LIEFFERS, V. J. 2008b. Inequality of size and size increment in Pinus banksiana in relation to stand dynamics and annual growth rate. *Annals of Botany,* 101**,** 561-571.

METTE, T., PAPATHANASSIOU, K. & HAJNSEK, I. Biomass estimation from polarimetric SAR interferometry over heterogeneous forest terrain. IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium, 2004a. IEEE, 511-514.

METTE, T., PAPATHANASSIOU, K., HAJNSEK, I., PRETZSCH, H. & BIBER, P. Applying a common allometric equation to convert forest height from Pol-InSAR data to forest biomass. IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium, 2004b. IEEE.

METTE, T., PAPATHANASSIOU, K., HAJNSEK, I. & ZIMMERMANN, R. Forest biomass estimation using polarimetric SAR interferometry. Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International, 2002. IEEE, 817-819.

MIKOLOV, T., KARAFIÁT, M., BURGET, L., ČERNOCKÝ, J. & KHUDANPUR, S. Recurrent neural network based language model. Eleventh annual conference of the international speech communication association, 2010.

MILLER, C., GRIFFITH, P., GOETZ, S., HOY, E., PINTO, N., MCCUBBIN, I., THORPE, A., HOFTON, M., HODKINSON, D. & HANSEN, C. 2019. An overview of ABoVE airborne campaign data acquisitions and science opportunities. *Environmental Research Letters,* 14**,** 080201.

MILLER, J. D. & THODE, A. E. 2007. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sensing of Environment,* 109**,** 66-80.

MITCHARD, E. T., SAATCHI, S. S., WHITE, L., ABERNETHY, K., JEFFERY, K. J., LEWIS, S. L., COLLINS, M., LEFSKY, M. A., LEAL, M. E. & WOODHOUSE, I. H. 2012. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: overcoming problems of high biomass and persistent cloud. *Biogeosciences,* 9**,** 179-191.

MOBAHI, H., RAO, S. R., YANG, A. Y., SASTRY, S. S. & MA, Y. 2011. Segmentation of natural images by texture and boundary compression. *International journal of computer vision,* 95**,** 86-98.

MONDINI, A. C. 2017. Measures of spatial autocorrelation changes in multitemporal SAR images for event landslides detection. *Remote Sensing,* 9**,** 554.

MONTESANO, P. M., NEIGH, C., SUN, G., DUNCANSON, L., VAN DEN HOEK, J. & RANSON, K. J. 2017. The use of sun elevation angle for stereogrammetric boreal forest height in open canopies. *Remote sensing of environment,* 196**,** 76-88.

MOREIRA, A., HAJNSEK, I., KRIEGER, G., PAPATHANASSIOU, K., EINEDER, M., DE ZAN, F., YOUNIS, M. & WERNER, M. 2009. Tandem-L: Monitoring the earth's dynamics with InSAR and Pol-InSAR. *Proceedings of PolInSAR 2009***,** 1-5.

MOREIRA, A., KRIEGER, G., HAJNSEK, I., PAPATHANASSIOU, K., YOUNIS, M., LOPEZ-DEKKER, P., HUBER, S., VILLANO, M., PARDINI, M. & EINEDER, M. 2015. Tandem-L: A highly innovative bistatic SAR mission for global observation of dynamic processes on the Earth's surface. *IEEE Geoscience and Remote Sensing Magazine,* 3**,** 8-23.

MOREIRA, A., PRATS-IRAOLA, P., YOUNIS, M., KRIEGER, G., HAJNSEK, I. & PAPATHANASSIOU, K. P. 2013. A tutorial on synthetic aperture radar. *IEEE Geoscience and remote sensing magazine,* 1**,** 6-43.

MURPHY, S. W., DE SOUZA FILHO, C. R., WRIGHT, R., SABATINO, G. & PABON, R. C. 2016. HOTMAP: Global hot target detection at moderate spatial resolution. *Remote sensing of environment,* 177**,** 78-88.

NAJMAN, L. & SCHMITT, M. 1994. Watershed of a continuous function. *Signal Processing,* 38**,** 99-112.

NAYEGANDHI, A., BROCK, J. C., WRIGHT, C. W. & O'CONNELL, M. J. 2006. Evaluating a small footprint, waveform-resolving lidar over coastal vegetation communities. *Photogrammetric Engineering & Remote Sensing,* 72**,** 1407-1417.

NEUMANN, M., FERRO-FAMIL, L. & POTTIER, E. A general model-based polarimetric decomposition scheme for vegetated areas. Proceedings of the 4th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (ESRIN), Frascati, Italy, 2009a. Citeseer, 26-30.

NEUMANN, M., FERRO-FAMIL, L. & REIGBER, A. 2009b. Estimation of forest structure, ground, and canopy layer characteristics from multibaseline polarimetric interferometric SAR data. *IEEE Transactions on Geoscience and Remote Sensing,* 48**,** 1086-1104.

NGUYEN, A. & LE, B. 3D point cloud segmentation: A survey. 2013 6th IEEE conference on robotics, automation and mechatronics (RAM), 2013. IEEE, 225-230.

NIELSEN, A. A., CONNETABLE, P. J., CONRADSEN, K., SKRIVER, H. & KROGAGER, E. Change detection in single-and multi-look polarimetric SAR data. IEEE Radar Conference 2020, 2020. IEEE.

NOCK, R. & NIELSEN, F. 2004. Statistical region merging. *IEEE Transactions on pattern analysis and machine intelligence,* 26**,** 1452-1458.

NOH, H., HONG, S. & HAN, B. Learning deconvolution network for semantic segmentation. Proceedings of the IEEE international conference on computer vision, 2015. 1520-1528.

OHLANDER, R., PRICE, K. & REDDY, D. R. 1978. Picture segmentation using a recursive region splitting method. *Computer Graphics and Image Processing,* 8**,** 313-333.

ORMSBY, T., NAPOLEON, E., BURKE, R., GROESSL, C. & FEASTER, L. 2004. *Getting to know ArcGIS desktop: basics of ArcView, ArcEditor, and ArcInfo*, ESRI, Inc.

OTSU, N. 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics,* 9**,** 62-66.

PADILLA, M., STEHMAN, S. V., RAMO, R., CORTI, D., HANTSON, S., OLIVA, P., ALONSO-CANAS, I., BRADLEY, A. V., TANSEY, K. & MOTA, B. 2015. Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation. *Remote sensing of environment,* 160**,** 114-121.

PALACIOS, A., MUÑOZ, M., DARBRA, R. & CASAL, J. 2012. Thermal radiation from vertical jet fires. *Fire safety journal,* 51**,** 93-101.

PAPATHANASSIOU, K. & CLOUDE, S. R. The effect of temporal decorrelation on the inversion of forest parameters from PoI-InSAR data. International Geoscience and Remote Sensing Symposium, 2003. III: 1429-1431.

PAPATHANASSIOU, K., REIGBER, A., SCHEIBER, R., HORN, R., MOREIRA, A. & CLOUDE, S. Airborne polarimetric SAR interferometry. Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS'98. 1998 IEEE International, 1998. IEEE, 1901-1903.

PAPATHANASSIOU, K. P. & CLOUDE, S. R. 2001. Single-baseline polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing,* 39**,** 2352-2363.

PARDINI, M., KUGLER, F., LEE, S.-K., SAUER, S., TORANO-CAICOYA, A. & PAPATHANASSIOU, K. 2011. Biomass estimation from forest vertical structure: Potentials and challenges for multi-baseline Pol-InSAR techniques. *ESA PolInSAR.*

PASZKE, A., GROSS, S., MASSA, F., LERER, A., BRADBURY, J., CHANAN, G., KILLEEN, T., LIN, Z., GIMELSHEIN, N. & ANTIGA, L. 2019. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*.

PAUSE, M., SCHWEITZER, C., ROSENTHAL, M., KEUCK, V., BUMBERGER, J., DIETRICH, P., HEURICH, M., JUNG, A. & LAUSCH, A. 2016. In situ/remote sensing integration to assess forest health—A review. *Remote Sensing,* 8**,** 471.

PETRIE, G. & TOTH, C. K. 2018. Introduction to laser ranging, profiling, and scanning. *Topographic laser ranging and scanning.* CRC Press.

PICHIERRI, M. & HAJNSEK, I. Comparing performances of RVoG and OVoG crop height inversion schemes from multi-frequency SAR data. Proceedings of EUSAR 2016: 11th European Conference on Synthetic Aperture Radar, 2016. VDE, 1-4.

PLANET 2020. Planet imagery product specifications. Planet Team, San Francisco, CA, USA.

PLATH, N., TOUSSAINT, M. & NAKAJIMA, S. Multi-class image segmentation using conditional random fields and global classification. Proceedings of the 26th Annual International Conference on Machine Learning, 2009. 817-824.

POTIN, P., ROSICH, B., MIRANDA, N., GRIMONT, P., SHURMER, I., O'CONNELL, A., KRASSENBURG, M. & GRATADOUR, J.-B. Copernicus Sentinel-1 Constellation Mission Operations Status. IGARSS 2019-2019

IEEE International Geoscience and Remote Sensing Symposium, 2019. IEEE, 5385-5388.
POTTER, C. 2018. Recovery rates of wetland vegetation greenness in severely burned ecosystems of Alaska derived from satellite image analysis. *Remote Sensing,* 10**,** 1456.
POTTIER, E. & LEE, J.-S. Application of the «H/A/alpha» polarimetric decomposition theorem for unsupervised classification of fully polarimetric SAR data based on the wishart distribution. SAR workshop: CEOS Committee on Earth Observation Satellites, 2000. 335.
POURSHAMSI, M., GARCIA, M., LAVALLE, M. & BALZTER, H. 2018. A machine-learning approach to PolInSAR and LiDAR data fusion for improved tropical forest canopy height estimation using NASA AfriSAR Campaign data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11**,** 3453-3463.
PRADHAN, P. S., KING, R. L., YOUNAN, N. H. & HOLCOMB, D. W. 2006. Estimation of the number of decomposition levels for a wavelet-based multiresolution multisensor image fusion. *IEEE Transactions on Geoscience and Remote Sensing,* 44**,** 3674-3686.
PRAKS, J., KUGLER, F., PAPATHANASSIOU, K. P., HAJNSEK, I. & HALLIKAINEN, M. 2007. Height estimation of boreal forest: Interferometric model-based inversion at L-and X-band versus HUTSCAT profiling scatterometer. *IEEE Geoscience and remote sensing letters,* 4**,** 466-470.
PRESTON, C. M., BHATTI, J. S. & NORRIS, C. E. 2014. Chemical quality of aboveground litter inputs for jack pine and black spruce stands along the Canadian Boreal Forest Transect Case Study. *Écoscience,* 21**,** 202-216.
QGIS 2018. Qgis: A Free and Open Source Geographic Information System. *Version 3.8. 0, Computer software*.
QI, W. & DUBAYAH, R. O. 2016. Combining Tandem-X InSAR and simulated GEDI lidar observations for forest structure mapping. *Remote sensing of Environment,* 187**,** 253-266.
QI, W., LEE, S.-K., HANCOCK, S., LUTHCKE, S., TANG, H., ARMSTON, J. & DUBAYAH, R. 2019. Improved forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data. *Remote sensing of environment,* 221**,** 621-634.
QU, Y., BAGHBADERANI, R. K., QI, H. & KWAN, C. 2020. Unsupervised pansharpening based on self-attention mechanism. *IEEE Transactions on Geoscience and Remote Sensing,* 59**,** 3192-3208.
RAMO, R. & CHUVIECO, E. 2017. Developing a random forest algorithm for MODIS global burned area classification. *Remote Sensing,* 9**,** 1193.
RAMO, R., GARCÍA, M., RODRÍGUEZ, D. & CHUVIECO, E. 2018. A data mining approach for global burned area mapping. *International journal of applied earth observation and geoinformation,* 73**,** 39-51.
RAMO, R., ROTETA, E., BISTINAS, I., VAN WEES, D., BASTARRIKA, A., CHUVIECO, E. & VAN DER WERF, G. R. 2021. African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data. *Proceedings of the National Academy of Sciences,* 118.

REIGBER, A. & MOREIRA, A. 2000. First demonstration of airborne SAR tomography using multibaseline L-band data. *IEEE Transactions on Geoscience and Remote Sensing,* 38**,** 2142-2152.

REIGBER, A. & SCHEIBER, R. 2003. Airborne differential SAR interferometry: First results at L-band. *IEEE Transactions on Geoscience and Remote Sensing,* 41**,** 1516-1520.

RICHARDS, J. A. 2009. *Remote sensing with imaging radar*, Springer.

ROBERTS, G. & WOOSTER, M. 2014. Development of a multi-temporal Kalman filter approach to geostationary active fire detection & fire radiative power (FRP) estimation. *Remote sensing of environment,* 152**,** 392-412.

RONNEBERGER, O., FISCHER, P. & BROX, T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention, 2015. Springer, 234-241.

ROSEN, P. A., HENSLEY, S., JOUGHIN, I. R., LI, F. K., MADSEN, S. N., RODRIGUEZ, E. & GOLDSTEIN, R. M. 2000. Synthetic aperture radar interferometry. *Proceedings of the IEEE,* 88**,** 333-382.

ROY, D. P., BOSCHETTI, L., JUSTICE, C. & JU, J. 2008. The collection 5 MODIS burned area product—Global evaluation by comparison with the MODIS active fire product. *Remote sensing of Environment,* 112**,** 3690-3707.

ROY, D. P., HUANG, H., BOSCHETTI, L., GIGLIO, L., YAN, L., ZHANG, H. H. & LI, Z. 2019. Landsat-8 and Sentinel-2 burned area mapping-A combined sensor multi-temporal change detection approach. *Remote Sensing of Environment,* 231**,** 111254.

SCHMITT, M., SHAHZAD, M. & ZHU, X. X. 2015. Reconstruction of individual trees from multi-aspect TomoSAR data. *Remote Sensing of Environment,* 165**,** 175-185.

SCHNEIDER, F. D., LEITERER, R., MORSDORF, F., GASTELLU-ETCHEGORRY, J.-P., LAURET, N., PFEIFER, N. & SCHAEPMAN, M. E. 2014. Simulating imaging spectrometer data: 3D forest modeling based on LiDAR and in situ data. *Remote Sensing of Environment,* 152**,** 235-250.

SCHOWENGERDT, R. A. 2006. *Remote sensing: models and methods for image processing*, Elsevier.

SCHROEDER, W., OLIVA, P., GIGLIO, L. & CSISZAR, I. A. 2014. The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment,* 143**,** 85-96.

SCHROEDER, W., OLIVA, P., GIGLIO, L., QUAYLE, B., LORENZ, E. & MORELLI, F. 2016. Active fire detection using Landsat-8/OLI data. *Remote sensing of environment,* 185**,** 210-220.

SCHUTZ, B. E., ZWALLY, H. J., SHUMAN, C. A., HANCOCK, D. & DIMARZIO, J. P. 2005. Overview of the ICESat mission. *Geophysical Research Letters,* 32.

SEXTON, J. O., BAX, T., SIQUEIRA, P., SWENSON, J. J. & HENSLEY, S. 2009. A comparison of lidar, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America. *Forest Ecology and Management,* 257**,** 1136-1147.

SHANMUGAPRIYA, P., RATHIKA, S., RAMESH, T. & JANAKI, P. 2019. Applications of remote sensing in agriculture-A Review. *Int. J. Current Microbiol. Appl. Sci,* 8**,** 2270-2283.

SHAO, Z., LU, Z., RAN, M., FANG, L., ZHOU, J. & ZHANG, Y. 2019a. Residual encoder–decoder conditional generative adversarial network for pansharpening. *IEEE Geoscience and Remote Sensing Letters,* 17**,** 1573-1577.

SHAO, Z., PAN, Y., DIAO, C. & CAI, J. 2019b. Cloud detection in remote sensing images based on multiscale features-convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing,* 57**,** 4062-4076.

SILVA, C. A., SAATCHI, S., GARCIA, M., LABRIERE, N., KLAUBERG, C., FERRAZ, A., MEYER, V., JEFFERY, K. J., ABERNETHY, K. & WHITE, L. 2018. Comparison of small-and large-footprint lidar characterization of tropical forest aboveground structure and biomass: a case study from Central Gabon. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11**,** 3512-3526.

SIMARD, M. & DENBINA, M. 2017. An assessment of temporal decorrelation compensation methods for forest canopy height estimation using airborne L-band same-day repeat-pass polarimetric SAR interferometry. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11**,** 95-111.

SIMARD, M. & DENBINA, M. 2018. An Assessment of Temporal Decorrelation Compensation Methods for Forest Canopy Height Estimation Using Airborne L-Band Same-Day Repeat-Pass Polarimetric SAR Interferometry. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11**,** 95-111.

SMALL, D. 2011. Flattening gamma: Radiometric terrain correction for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing,* 49**,** 3081-3093.

SMITH, A., DRAKE, N., WOOSTER, M., HUDAK, A., HOLDEN, Z. & GIBBONS, C. 2007. Production of Landsat ETM+ reference imagery of burned areas within Southern African savannahs: comparison of methods and application to MODIS. *International Journal of Remote Sensing,* 28**,** 2753-2775.

SPOTO, F., SY, O., LABERINTI, P., MARTIMORT, P., FERNANDEZ, V., COLIN, O., HOERSCH, B. & MEYGRET, A. Overview of Sentinel-2. 2012 IEEE International Geoscience and Remote Sensing Symposium, 2012. IEEE, 1707-1710.

STARCK, J.-L., ELAD, M. & DONOHO, D. L. 2005. Image decomposition via the combination of sparse representations and a variational approach. *IEEE transactions on image processing,* 14**,** 1570-1582.

STROPPIANA, D., AZAR, R., CALÒ, F., PEPE, A., IMPERATORE, P., BOSCHETTI, M., SILVA, J., BRIVIO, P. A. & LANARI, R. 2015. Integration of optical and SAR data for burned area mapping in Mediterranean Regions. *Remote Sensing,* 7**,** 1320-1345.

STROPPIANA, D., BORDOGNA, G., CARRARA, P., BOSCHETTI, M., BOSCHETTI, L. & BRIVIO, P. 2012. A method for extracting burned areas from Landsat TM/ETM+ images by soft aggregation of multiple Spectral Indices and a region growing algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing,* 69**,** 88-102.

STROPPIANA, D., BOSCHETTI, M., ZAFFARONI, P. & BRIVIO, P. A. 2009. Analysis and interpretation of spectral indices for soft multicriteria burned-area mapping in mediterranean regions. *IEEE Geoscience and Remote Sensing Letters,* 6**,** 499-503.

STROPPIANA, D., PINNOCK, S. & GREGOIRE, J.-M. 2000. The global fire product: Daily fire occurrence from April 1992 to December 1993 derived from NOAA AVHRR data. *International Journal of Remote Sensing,* 21**,** 1279-1288.

SULLA-MENASHE, D. & FRIEDL, M. A. 2018. User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product. *USGS: Reston, VA, USA***,** 1-18.

SULLA-MENASHE, D., GRAY, J. M., ABERCROMBIE, S. P. & FRIEDL, M. A. 2019. Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product. *Remote Sensing of Environment,* 222**,** 183-194.

SUMNALL, M. J., HILL, R. A. & HINSLEY, S. A. 2016. Comparison of small-footprint discrete return and full waveform airborne lidar data for estimating multiple forest variables. *Remote Sensing of Environment,* 173**,** 214-223.

SUN, K., XIAO, B., LIU, D. & WANG, J. Deep high-resolution representation learning for human pose estimation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 5693-5703.

SYPHARD, A. D. & KEELEY, J. E. 2015. Location, timing and extent of wildfire vary by cause of ignition. *International Journal of Wildland Fire,* 24**,** 37-47.

TABB, M., ORREY, J., FLYNN, T. & CARANDE, R. Phase diversity: A decomposition for vegetation parameter estimation using polarimetric SAR interferometry. Proc. EUSAR, 2002. 721-724.

TAKIKAWA, T., ACUNA, D., JAMPANI, V. & FIDLER, S. Gated-scnn: Gated shape cnns for semantic segmentation. Proceedings of the IEEE International Conference on Computer Vision, 2019. 5229-5238.

TANASE, M. A., BELENGUER-PLOMER, M. A., ROTETA, E., BASTARRIKA, A., WHEELER, J., FERNÁNDEZ-CARRILLO, Á., TANSEY, K., WIEDEMANN, W., NAVRATIL, P. & LOHBERGER, S. 2020. Burned area detection and mapping: Intercomparison of sentinel-1 and sentinel-2 based algorithms over tropical Africa. *Remote Sensing,* 12**,** 334.

TANASE, M. A., SANTORO, M., APONTE, C. & DE LA RIVA, J. 2013. Polarimetric properties of burned forest areas at C-and L-band. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 7**,** 267-276.

TANASE, M. A., SANTORO, M., DE LA RIVA, J., FERNANDO, P. & LE TOAN, T. 2010a. Sensitivity of X-, C-, and L-band SAR backscatter to burn severity in Mediterranean pine forests. *IEEE Transactions on Geoscience and Remote Sensing,* 48**,** 3663-3675.

TANASE, M. A., SANTORO, M., WEGMÜLLER, U., DE LA RIVA, J. & PÉREZ-CABELLO, F. 2010b. Properties of X-, C-and L-band repeat-pass interferometric SAR coherence in Mediterranean pine forests affected by fires. *Remote Sensing of Environment,* 114**,** 2182-2194.

THOMPSON, M. P., MACGREGOR, D. G., DUNN, C. J., CALKIN, D. E. & PHIPPS, J. 2018. Rethinking the wildland fire management system. *Journal of Forestry,* 116**,** 382-390.

TOLLEFSON, J. 2020. Why deforestation and extinctions make pandemics more likely. *Nature,* 584**,** 175-176.

TRETTIN, C. C., DAI, Z., TANG, W., LAGOMASINO, D., THOMAS, N., LEE, S. K., SIMARD, M., EBANEGA, M. O., STOVAL, A. & FATOYINBO, T. E.

2021. Mangrove carbon stocks in Pongara National Park, Gabon. *Estuarine, Coastal and Shelf Science*, 107432.
TREUHAFT, R. N., MOGHADDAM, M. & VAN ZYL, J. J. 1996. Vegetation characteristics and underlying topography from interferometric radar. *Radio Science,* 31, 1449-1485.
TREUHAFT, R. N. & SIQUEIRA, P. R. 2000. Vertical structure of vegetated land surfaces from interferometric and polarimetric radar. *Radio Science,* 35, 141-177.
TRUCKENBRODT, J., FREEMANTLE, T., WILLIAMS, C., JONES, T., SMALL, D., DUBOIS, C., THIEL, C., ROSSI, C., SYRIOU, A. & GIULIANI, G. 2019. Towards Sentinel-1 SAR analysis-ready data: A best practices assessment on preparing backscatter data for the cube. *Data,* 4, 93.
TSANG, L., KONG, J. A. & SHIN, R. T. 1985. Theory of microwave remote sensing.
TU, T.-M., SU, S.-C., SHYU, H.-C. & HUANG, P. S. 2001. A new look at IHS-like image fusion methods. *Information fusion,* 2, 177-186.
ULABY, F. T., MOORE, R. K. & FUNG, A. K. 1982. *Microwave remote sensing: active and passive. Vol. 2, Radar remote sensing and surface scattering and emission theory*, Addison-Wesley Reading, MA.
VAN KHUC, Q., TRAN, B. Q., MEYFROIDT, P. & PASCHKE, M. W. 2018. Drivers of deforestation and forest degradation in Vietnam: An exploratory analysis at the national level. *Forest policy and economics,* 90, 128-141.
VAN ZYL, J. J. 2011. *Synthetic aperture radar polarimetry*, John Wiley & Sons.
VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, Ł. & POLOSUKHIN, I. Attention is all you need. Advances in neural information processing systems, 2017. 5998-6008.
VECI, L. 2015. TOPS Interferometry Tutorial. *Sentinel-1 Toolbox*.
VECI, L., LU, J., FOUMELIS, M. & ENGDAHL, M. ESA's Multi-mission Sentinel-1 Toolbox. EGU General Assembly Conference Abstracts, 2017. 19398.
VINCENT, L. & SOILLE, P. 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Computer Architecture Letters,* 13, 583-598.
WANG, J., SUN, K., CHENG, T., JIANG, B., DENG, C., ZHAO, Y., LIU, D., MU, Y., TAN, M. & WANG, X. 2020. Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*.
WARNER, T. A., SKOWRONSKI, N. S. & GALLAGHER, M. R. 2017. High spatial resolution burn severity mapping of the New Jersey Pine Barrens with WorldView-3 near-infrared and shortwave infrared imagery. *International Journal of Remote Sensing,* 38, 598-616.
WATT, M. S., PEARSE, G. D., DASH, J. P., MELIA, N. & LEONARDO, E. M. C. 2019. Application of remote sensing technologies to identify impacts of nutritional deficiencies on forests. *ISPRS Journal of Photogrammetry and Remote Sensing,* 149, 226-241.
WITHEY, P., JOHNSTON, C. & GUO, J. 2019. Quantifying the global warming potential of carbon dioxide emissions from bioenergy with carbon capture and storage. *Renewable and Sustainable Energy Reviews,* 115, 109408.
WOOSTER, M. J., XU, W. & NIGHTINGALE, T. 2012. Sentinel-3 SLSTR active fire detection and FRP product: Pre-launch algorithm development and

performance evaluation using MODIS and ASTER datasets. *Remote Sensing of Environment,* 120**,** 236-254.

WULDER, M. A., WHITE, J. C., NELSON, R. F., NÆSSET, E., ØRKA, H. O., COOPS, N. C., HILKER, T., BATER, C. W. & GOBAKKEN, T. 2012. Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment,* 121**,** 196-209.

XIE, Y., SHA, Z. & YU, M. 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of plant ecology,* 1**,** 9-23.

XIE, Y., TIAN, J. & ZHU, X. X. 2020. Linking points with labels in 3D: A review of point cloud semantic segmentation. *IEEE Geoscience and Remote Sensing Magazine,* 8**,** 38-59.

XU, R., YU, P., ABRAMSON, M. J., JOHNSTON, F. H., SAMET, J. M., BELL, M. L., HAINES, A., EBI, K. L., LI, S. & GUO, Y. 2020. Wildfires, Global Climate Change, and Human Health. *New England Journal of Medicine,* 383**,** 2173-2181.

YAMAGUCHI, Y., MORIYAMA, T., ISHIDO, M. & YAMADA, H. 2005. Four-component scattering model for polarimetric SAR image decomposition. *IEEE Transactions on Geoscience and Remote Sensing,* 43**,** 1699-1706.

YAMAGUCHI, Y., SATO, A., BOERNER, W.-M., SATO, R. & YAMADA, H. 2011. Four-component scattering power decomposition with rotation of coherency matrix. *IEEE Transactions on Geoscience and Remote Sensing,* 49**,** 2251-2258.

YANG, J., FU, X., HU, Y., HUANG, Y., DING, X. & PAISLEY, J. PanNet: A deep network architecture for pan-sharpening. Proceedings of the IEEE international conference on computer vision, 2017. 5449-5457.

YANG, L. & JIN, R. 2006. Distance metric learning: A comprehensive survey. *Michigan State Universiy,* 2**,** 4.

YANG, M., YU, K., ZHANG, C., LI, Z. & YANG, K. Denseaspp for semantic segmentation in street scenes. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018. 3684-3692.

YAO, H., QIN, R. & CHEN, X. 2019. Unmanned aerial vehicle for remote sensing applications—A review. *Remote Sensing,* 11**,** 1443.

YUAN, Q., SHEN, H., LI, T., LI, Z., LI, S., JIANG, Y., XU, H., TAN, W., YANG, Q. & WANG, J. 2020. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment,* 241**,** 111716.

YUAN, Q., WEI, Y., MENG, X., SHEN, H. & ZHANG, L. 2018a. A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11**,** 978-989.

YUAN, Y., HUANG, L., GUO, J., ZHANG, C., CHEN, X. & WANG, J. 2018b. Ocnet: Object context network for scene parsing. *arXiv preprint arXiv:1809.00916*.

ZHANG, J. 2010. Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion,* 1**,** 5-24.

ZHANG, L., ZHANG, L. & DU, B. 2016a. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine,* 4**,** 22-40.

ZHANG, Q., GE, L., ZHANG, R., METTERNICHT, G. I., DU, Z., KUANG, J. & XU, M. 2021a. Deep-learning-based burned area mapping using the synergy of Sentinel-1&2 data. *Remote Sensing of Environment,* 264**,** 112575.

ZHANG, Q., GE, L., ZHANG, R., METTERNICHT, G. I., LIU, C. & DU, Z. 2021b. Towards a Deep-Learning-Based Framework of Sentinel-2 Imagery for Automated Active Fire Detection. *Remote Sensing,* 13**,** 4790.

ZHANG, Q., LIU, T., DING, Z., ZENG, T. & LONG, T. 2016b. A modified three-stage inversion algorithm based on R-RVoG model for Pol-InSAR data. *Remote Sensing,* 8**,** 861.

ZHANG, Y., DE BACKER, S. & SCHEUNDERS, P. 2009. Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing,* 47**,** 3834-3843.

ZHANG, Z. & SABUNCU, M. Generalized cross entropy loss for training deep neural networks with noisy labels. Advances in neural information processing systems, 2018. 8778-8788.

ZHAO, H., SHI, J., QI, X., WANG, X. & JIA, J. Pyramid scene parsing network. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. 2881-2890.

ZHENG, S., LU, J., ZHAO, H., ZHU, X., LUO, Z., WANG, Y., FU, Y., FENG, J., XIANG, T. & TORR, P. H. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021. 6881-6890.

ZHOU, H., LIU, Q. & WANG, Y. 2021. PGMAN: An Unsupervised Generative Multiadversarial Network for Pansharpening. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 14**,** 6316-6327.

ZHU, X. X. & BAMLER, R. 2012. A sparse image fusion algorithm with application to pan-sharpening. *IEEE transactions on geoscience and remote sensing,* 51**,** 2827-2836.

ZHU, X. X., GROHNFELDT, C. & BAMLER, R. 2015. Exploiting joint sparsity for pansharpening: The J-SparseFI algorithm. *IEEE Transactions on Geoscience and Remote Sensing,* 54**,** 2664-2681.

ZHU, X. X., TUIA, D., MOU, L., XIA, G.-S., ZHANG, L., XU, F. & FRAUNDORFER, F. 2017. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine,* 5**,** 8-36.

ZWICKER, W. S. 2016. Introduction to the Theory of Voting.