7-12-2016

Evaluation of sensor, environment and operational factors impacting the use of multiple sensor constellations for long term resource monitoring

Rajagopalan Rengarajan
rxr1823@rit.edu

Evaluation of sensor, environment and operational factors impacting the use of multiple sensor constellations for long term resource monitoring

by

Rajagopalan Rengarajan

B.E. Anna University

M.S. Purdue University

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in Imaging Science
in the Chester F. Carlson Center for Imaging Science
College of Science
Rochester Institute of Technology

July 12, 2016

Signature of the Author _____

Accepted by _____
Coordinator Ph.D. Degree Program   Date

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE

COLLEGE OF SCIENCE

ROCHESTER INSTITUTE OF TECHNOLOGY

ROCHESTER, NEW YORK

<u>CERTIFICATE OF APPROVAL</u>

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Rajagopalan Rengarajan
has been examined and approved by the
dissertation committee as satisfactory for the
dissertation required for the
Ph.D. degree in Imaging Science

Dr. John R. Schott, Dissertation Advisor

Dr. Joseph Voelkel

Dr. Emmett Ientilucci

Dr. Carl Salvaggio

Date

Evaluation of sensor, environment and operational factors impacting the use of multiple sensor constellations for long term resource monitoring

by

Rajagopalan Rengarajan

Submitted to the
Chester F. Carlson Center for Imaging Science
in partial fulfillment to the requirements
for the degree of Doctor of Philosophy in Imaging Science
at the Rochester Institute of Technology

Abstract

Moderate resolution remote sensing data offers the potential to monitor the long and short term trends in the condition of the Earth's resources at finer spatial scales and over longer time periods. While improved calibration (radiometric and geometric), free access (Landsat, Sentinel, CBERS), and higher level products in reflectance units have made it easier for the science community to derive the biophysical parameters from these remotely sensed data, a number of issues still affect the analysis of multi-temporal datasets. These are primarily due to sources that are inherent in the process of imaging from single or multiple sensors. Some of these undesired or uncompensated sources of variation include variation in the view angles, illumination angles, atmospheric effects, and sensor effects such as Relative Spectral Response (RSR) variation between different sensors. The complex interaction of these sources of variation would make their study extremely difficult if not impossible with real data, and therefore, a simulated analysis approach is used in this study.

A synthetic forest canopy is produced using the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model and its measured BRDFs are modeled using the RossLi canopy BRDF model. The simulated BRDF matches the real data to within 2% of the reflectance in the red and the NIR spectral bands studied. The BRDF modeling process is extended to model and characterize the defoliation of a forest, which is used in factor sensitivity studies to estimate the effect of each factor for varying environment and sensor conditions. Finally, a factorial experiment is designed to understand the significance of the sources of variation, and regression based analysis are performed to understand the relative importance of the factors. The design of experiment and the sensitivity analysis conclude that the atmospheric attenuation and variations due to the illumination angles are the dominant sources impacting the at-sensor radiance.

Acknowledgements

First and foremost, I would like to express my gratitude to my advisor, Dr. John Schott, for his guidance in this research and also in my career. His expertise is well known, but his patience, dedication, and the encouragement he provides to his students are something that cannot be expressed in words, but am sure all his graduate students would attest to it. I take this opportunity to express my gratitude to my external advisor Dr. Joseph Voelkel, who has been extremely helpful with the design of experiments and analysis techniques used in this research. I also would like to thank Dr. Emmett Ientilucci and Dr. Carl Salvaggio for accepting to be in my committee, and I sincerely appreciate their support, and the valuable feedback and comments they provided.

I would like to extend my thanks to Dr. Adam Goodenough, Dr. Aaron Gerace and Dr. Mike Gartley for their mentorship and assistance in working with the DIRSIG tool. Special thanks to Paul Mezzannini for his tireless effort in keeping the research computers running as the modeling effort in this research would not have been possible without it. I would not have pursued the Imaging Science program at RIT if not for my mentor Jim Storey, and am grateful for his advice.

I cannot thank my spiritual teacher, Swami Yogatmananda, enough for his love, guidance and continued support in helping me understand the real purpose of life and how to set priorities and work diligently. Lastly, none of this work would have been possible without the love and encouragement from my family. I would not be where I am today if not for their sacrifice and support, and it is impossible for me to express my appreciation in words.

Contents

List of Figures

List of Tables

Abbreviations

ANOVA	**AN**alysis **O**f **VA**riance
ASD	**A**nalytical **S**pectral **D**evices
BRDF	**B**i-directional **R**eflectance **D**istribution **F**unction
BRF	**B**i-directional **R**eflectance **F**actor
BTDF	**B**i-directional **T**ransmittance **D**istribution **F**unction
BU	**B**oston **U**niversity
DBH	**D**iameter at **B**reast **H**eight
DIRSIG	**D**igital **I**maging and **R**emote **S**ensing **I**mage **G**eneration
DOE	**D**esign **O**f **E**xperiments
ELM	**E**mprical **L**ine **M**ethod
FOV	**F**ield **O**f **V**iew
FPM	**F**ocal **P**lane **M**odule
GIFOV	**G**round **I**nstantaneous **F**ield **O**f **V**iew
HCRDF	**H**emispherical **C**onical **R**eflectance **D**istribution **F**unction
HCRF	**H**emispherical **C**onical **R**eflectance **F**actor
HDRF	**H**emispherical **D**irectional **R**eflectance **F**actor
IFOV	**I**nstantaneous **F**ield **O**f **V**iew
LAI	**L**eaf **A**rea **I**ndex
LIDAR	**LI**ght **D**etection **A**nd **R**anging
LOOC	**L**eave **O**ne **O**ut **C**count
MODIS	**MOD**erate resolution **I**maging **S**pectroradiometer
MODTRAN	**MOD**erate resolution atmospheric **T**ransmission
MSI	**M**ulti**S**pectral **I**nstrument
NIR	**N**ear **I**nfra **R**ed spectral region
OLI	**O**perational **L**and **I**mager

PCA	**P**rincipal **C**omponent **A**nalysis
RAMI	**RA**diation transfer **M**odel **I**ntercomparison
RMSE	**R**oot **M**ean **S**quare **E**rror
ROI	**R**egion **O**f **I**nterest
RSR	**R**elative **S**pectral **R**esponse
SNR	**S**ignal to **N**oise **R**atio
SWIR	**S**hort **W**ave **I**nfra **R**ed spectral region
UMB	**U**niversity of **M**assachusetts, **B**oston
VIS	**V**isible spectral region
SZN	**S**olar **Z**enith
SAZ	**S**olar **AZ**imuth
X-Track	Cross **T**rack
A-Track	**A**long **T**rack
NED	**N**oise **E**quivalent **D**efoliation
NDVI	**N**ormalized **D**ifference **V**egetation **I**ndex
TOA	**T**op **O**f **A**tmosphere
TOA RAD	**T**op **O**f **A**tmosphere **RAD**iance
TOA REFL	**T**op **O**f **A**tmosphere **REFL**ectance
ELM	**E**mprical **L**ine **M**ethod
STD	**ST**andard **D**eviation
SBAF	**S**pectral **B**and **A**djustment **F**actor

Symbols

$R'(\lambda)$	Relative Spectral Response	$[unitless]$
$L(\lambda)$	Effective spectral radiance	$[Wm^{-2}sr^{-1}\mu m^{-1}]$
$\rho(\lambda)$	Spectral reflectance	$[unitless]$
$r(\lambda)$	Spectral reflectance factor	$[unitless]$
$f(\phi_i, \theta_i, \phi_r, \theta_r, \lambda)$	BRDF	$[sr^{-1}]$
$r(\phi_i, \theta_i, \phi_r, \theta_r, \lambda)$	BRF	$[unitless]$
$R(\theta_i, \theta_v, \phi, \lambda)$	RossLi BRF	$[unitless]$
$E(\theta_i, \phi_i, \lambda)$	Spectral Irradiance	$[Wm^{-2}\mu m^{-1}]$
ΔL	Difference in Radiance	$[Wm^{-2}sr^{-1}\mu m^{-1}]$
f_{iso}	RossLi Isotropic coefficient	$[unitless]$
f_{vol}	RossLi Volumetric coefficient	$[unitless]$
f_{geo}	RossLi Geometric coefficient	$[unitless]$
λ	wavelength	μm
θ	Zenith angle	$radians$
ϕ	Azimuth angle	$radians$
τ	Transmittance through atmosphere	$unitless$

Chapter 1

Introduction

Remote sensing satellites have been monitoring the earth's natural resources for over four decades. The first operational moderate resolution earth remote sensing mission called Earth Resources Technology Satellites Program (ERTS, later changed to Landsat) was initiated in the 1960s with the first satellite, Landsat 1 launched in 1972. Since then Landsat missions have continuously monitored the land surfaces of the earth and produced radiometrically and geometrically corrected product for scientists.

The ready access to moderate resolution remote sensing data (i.e. free Landsat data) offers the potential to monitor the long (inter-annual) and short term (annual) trends in the condition of the Earth's resources at finer spatial scales and over longer time periods than ever before possible. This is possible not only because the data are now available to all users, but also because the entire archive of data is now well calibrated radiometrically and geometrically (Markham and Helder, 2012, Schott et al., 2011, Storey, 2001). In addition, the data are being made available in surface reflectance units allowing more direct comparison of the change in surface condition over time (Masek et al., 2006).

Landsat satellites currently in operation (Landsat 7 and Landsat 8) have a revisit period of 16 days which is adequate for land cover and land use change analysis at coarse time and continental scale resolution. However, with recent advancements in the sensor technology (spectral, spatial, radiometric resolution, detector sensitivity, high data rate, etc.), many new applications have recognized the advantage in the use of remotely sensed data to solve difficult problems. Short term resource monitoring applications at a finer time scale requires frequently acquired dataset

over the same region of earth, driving a necessity to collect data approximately once every day. The sensor's spatial resolution, swath width, sensitivity, orbital parameters, etc., enforces a limitation on the revisit period of a satellite. It is impractical to expect a single mission to satisfy the frequent revisit requirements. A frequent revisit period necessitates operating several satellites with high fidelity (spatial, spectral and radiometric resolution) systems which are cost prohibitive for any country.

Over the last two decades, many remote sensing missions are being operated by different countries, and using similar datasets from other missions could alleviate the problem of data availability. However, there is an inherent issue in using more than one sensor's data due to the differences in the method of measurement, the sensor parameters, and environmental factors. Even two similar sensors operated by the same vendor for the same mission could exhibit differences due to many factors. For example, two sensors may view the same ground on two different days (and time), reflecting a change in illumination angles and atmospheric conditions. The sensor characteristics such as Relative Spectral Response (RSR) of one sensor is likely to be different from the other even if produced by the same vendor as an identical sensor. These effects are illustrated in Figure 1.1 which shows a trend line fit to the data from a single pixel location in a time series of MODerate resolution Imaging Spectroradiometer (MODIS) data over time. The large variability in these data about the trend line and the variability in the fits illustrate the large amount of residual variation not compensated by the radiometric calibration and atmospheric compensation algorithms used in generating MODIS products. Landsat and other mission's radiometric calibration and compensation algorithms are expected to be at the same level of accuracy as MODIS. Moreover, MODIS collects data at a low spatial resolution but can be sampled every day, while Landsat is only available every 8 or 16 days (2 instruments or 1 instrument). This indicates that a similar analysis using two operating Landsat sensor's data will be much sparser than illustrated in Figure 1.1, and the use of more than one sensor to increase temporal coverage will require additional correction for sensor to sensor variation.

Currently, most research focuses on cross calibrating datasets from two different sensors. This is performed by observing the same region on the ground (typically a radiometrically calibrated site) on the same day at about the same time. This still raises a few issues. It is unlikely to have two sensors acquire the same region at about the same time on the same day and have similar view angles. Even on the same day due to different time of observation, illumination angles will change. The view angles to image the site are likely to change if the two sensors

observe on the same day. Also, the surface terrain can induce additional view and illumination angle variation between the sensors. If the two sensors image at about same time, but on two different dates, changes in atmospheric constituents can induce sensor to sensor variation along with variation due to phenological changes on the ground. Typically, some of these issues are compensated by using stable invariant sites which may not exhibit significant Bi-directional Reflectance Distribution Function (BRDF) or phenological changes over a short time period. However, any cross-calibration using such sites limits the compensation of sensor data to similar sites and cannot be applicable for other biomes such as forest canopy, agricultural fields, etc. In this research, variations exhibited due to these factors will be studied.
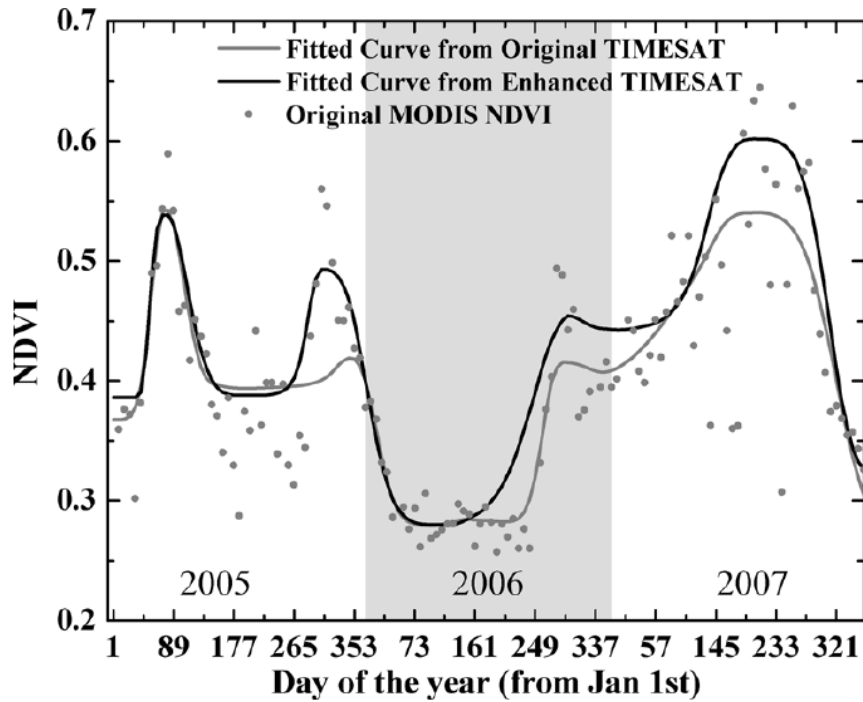
Figure 1.1: An example of fitted NDVI curve over unevenly distributed growing seasons. This sample pixel is located in Southwest US (within tile h09v06). The data period is 2005-2007 (Tan et al., 2011).

Forest canopies and other biomes such as coastal land, agriculture, and marshy lands are considered to exhibit BRDF that are significantly different from Lambertian surfaces. Since most of the earth observing remote sensing satellites are used for natural resource monitoring, it is imperative to use BRDF that closely represent such surfaces. In this research, forest canopies are used as a representative site for understanding the factors influencing sensor variations. The forest canopy reflectance changes over time due to phenologic cycle of the canopy (e.g. leaf off,

budding, leaf emergence, leaf growth characterized by increase in chlorophyll concentration and leaf area index, leaf senescence, and leaf fall). The timing and progress or disturbance of this cycle is of interest (Figure 1.1). Variations not due to changes in the forest condition can be thought of as clutter and ideally needs to be understood and compensated or removed. Some of these undesired or uncompensated sources of variation include:

1. Variation in view angles

2. Variation in illumination angles

3. Variation in atmospheric constituents

4. Variation in RSR between sensors

It is well known that all the sources of variation are real but it is important to know their relative magnitude or how the impact of one or another will change from acquisition to acquisition. Sophisticated scene-sensor-environmental models offer a potential to investigate and simulate the source, magnitude and functional dependencies of these induced variations. The goal of this research effort is to demonstrate the ability to effectively model the relationship between the sources of variation and image derived measurements.

The first phase of this research begins with modeling the forest canopies as accurately as possible. The forest canopy geometries are developed using OnyxTree software (Onyx Computing, 2015) while the ground measured spectral data from the Harvard forest site is used for tree leaves, trunk and ground litter. These models are used in the DIRSIG model whose radiative transfer algorithms are validated against published canopy radiative transfer models as found in Widlowski et al. (2014, 2006). DIRSIG will be used as a virtual goniometer to measure the BRDF of the modeled forest canopy. These measurements are used to fit to an established canopy BRDF model (RossLi BRDF model).

Once the source (forest canopy) is modeled as BRDF, different sources of variations are introduced in DIRSIG. The Landsat 8 and Sentinel 2 sensors are used as representative sensors for studying these variations. Their published sensor parameters are used for RSR, view angle, and illumination angle (based on overlap time) variations. Finally, a factorial design experiment based on Design Of Experiments (DOE) techniques will be designed using DIRSIG as the experimental engine. The factorial design experiment will allow us to independently and

collectively vary all of the sources of variation/clutter described in items 1 through 4 above over all the ranges expected from the sensors of interest. Using these methods, the study seeks to provide quantitative and statistical based methods to identify the significant factors and their relative significance.

Chapter 2

Objectives

As alluded to in the previous chapter, there are numerous factors that are likely to impact remotely sensed data, and it is impractical to study every possible factor. Even identifying and analyzing a few significant factors that affect the remotely sensed data is a complex problem, and is best studied by dividing them into multiple but smaller objectives with well defined tasks for each. The purpose of this chapter is to define these objectives and tasks for better understanding and defining the scope of the research.

This chapter is divided into four sections with Section 2.1 detailing the problem statement. Section 2.2 outlines the main objectives and Section 2.3 describes the tasks that must be accomplished to fulfill each objective. This chapter closes with Section 2.4 summarizing the research work's contribution to the field of remote sensing.

2.1   Problem Statement

The goal of this research effort is to determine the relative significance of various factors that affect the use of remotely sensed data from multiple sensor constellations for long term resource monitoring. Determining the relative significance of factors requires delineation of important and useful factors. Specifically, the sensor factors are limited to two sensors; Landsat 8 Operational Land Imager (OLI) and Sentinel 2 Multi Spectral Instrument (MSI). Since forest canopy is identified as representative for resource monitoring, good forest canopy modeling is

required. Evaluation of each factor's impact requires accurate modeling of forest, sensor, and environmental factors such as atmosphere, terrain and sun position.

2.2 Objectives

1. Design a process to validate DIRSIG for its accuracy in modeling radiative transfer of forest canopies.

2. Build a representative forest canopy scene for DIRSIG that can be placed anywhere in the world.

3. Model the forest canopy using canopy BRDF models in DIRSIG.

4. Model Landsat 8 and Sentinel 2 sensors in DIRSIG to capture their geometric and radiometric characteristics.

5. Identify and model the required environmental factors in DIRSIG.

6. Develop a factorial design experiment to identify and determine the relative significance of factors that affects the apparent reflectance observed by the sensors.

7. Develop a method to define the change in environment variation as a function of variation in sensor reaching radiance.

2.3 Tasks

1. **Design a process to validate DIRSIG for its accuracy in modeling radiative transfer of forest canopies**

   - Validate DIRSIG for appropriate response to shadows, multiple scattering, and nominal reflectance
   - Validation of DIRSIG by comparison with other radiative transfer models' results as published by RAdiation transfer Model Intercomparison (RAMI) (Widlowski et al., 2006)
   - Validation of DIRSIG using more complex forest geometries built by RAMI IV (if available)

2. **Build a representative forest canopy scene for DIRSIG that can be placed anywhere in the world**

   - Identify the type of trees, their size and shape to model based on Harvard forest ground survey campaigns

   - Generate tree models using tree generation software such as OnyxTree, to model trees as three-dimensional geometrical object. Make appropriate changes to tree parameters to mimic natural trees as closely as possible.

   - Identify a strategy for tree placement in the forest

   - Verify the consistency and distribution of trees, and refine as necessary to emulate a real forest site

3. **Model the forest canopy using canopy BRDF models in DIRSIG**

   - Identify a suitable BRDF model for representing forest canopies

   - Define necessary parameters for running DIRSIG simulations to measure virtual BRDF

   - Automate parallel processes to run DIRSIG simulations for BRDF measurements

   - Generate BRDF model coefficients by inversion of model parameters using DIRSIG BRDF measurements

4. **Model Landsat 8 and Sentinel 2 sensors in DIRSIG to capture their geometric and radiometric characteristics**

   - Identify the necessary sensor parameters required for modeling Operational Land Imager (OLI) and MultiSpectral Imager (MSI) in DIRSIG

   - Model the sensor and orbital platform motion for the two sensors in DIRSIG

5. **Identify and model the required environmental factors in DIRSIG**

   - Identify the necessary environmental factors such as sun position, atmosphere, terrain

   - Generate corresponding DIRSIG files that allows for modeling the environmental factors

6. **Develop a factorial design experiment to identify and determine the relative significance of factors that affects the apparent reflectance observed by the sensors**

   - Identify levels for each factors and perform factor screening experiments to identify the significant factors

   - Identify the appropriate design for performing factorial experiments

   - Analyze the result from experiments to determine the relative significance of factors

7. **Develop a method to define the change in environment variation as a function of variation in sensor reaching radiance**

   - Model the variation in the forest as defoliation

   - Generate BRDF for different level of defoliation

   - Estimate the variation in observed signal (i.e. sensor reaching radiance, reflectance and NDVI) for the defoliated forests

2.4   Contribution to Field

This research will make several contributions to the field of remote sensing and the prominent three are summarized below.

1. Firstly, the proposed work is unique for studying the complex factor interactions among environmental and sensor factors using factorial experiment studies in a simulation environment. The published literature has identified some of the key important factors and have assumed some factors to be dominant, but there have been no definite studies on understanding the relationship between these factors and their interaction in a complex environment. In contrast, this research will objectively identify the factors and their relative significance using statistical analysis, providing a strong basis for future research in deriving compensation algorithms for the factor's effect.

2. Secondly, the proposed work has demonstrated the possibility of simulation tools such as DIRSIG to generate an accurate model of a real forest. DIRSIG has been used to

model different types of scenes and materials, but an accurate representation of complex geometries as seen in forest canopy have not been explored extensively in the past. Using DIRSIG as a virtual goniometer can be useful to model any surfaces in a simulated environment. The proposed work shows a novel method in modeling the forest canopy as BRDF by constraining spectra, viewing and illumination angles based on the sensor's RSR and the geographic location of the forest. The measurement from DIRSIG is used for model fitting to generate full hemispherical spectral BRDF. This unique approach reduces the number of measurements required to model BRDF by a factor of 500 or more. The forest canopy's interaction with light is complex and modeled as BRDF in DIRSIG. This approach has significantly reduced the computational complexities and memory requirements. In the future, any modeling of complex geometries in DIRSIG will be simplified using its BRDF as demonstrated in this research. The measured BRDF from the modeled forest in this research can be directly used as a dataset to validate existing and new canopy BRDF models and for evaluating BRDF compensation algorithms.

3. Lastly, the application of design of experiment techniques, though not uncommon in other disciplines, have not been explored in the field of remote sensing extensively. This research has revealed the capabilities of using design of experiment techniques to study complex phenomena and the process established in this research will serve as a template for other studies where more than one factors' interaction and their relationship is of significance. This research further suggests that such studies can be conducted only in a simulated environment as their complex nature prohibits research studies with real-world dataset.

Chapter 3

Background and Theory

This chapter discusses the background and theoretical foundation necessary to accomplish the objectives outlined in Chapter 2. In Section 3.1, the definitions of radiometric terms that are primarily important to comprehend the fundamental principles involved in characterization of sensors and the real-world objects are discussed. Section 3.2 introduces different types of reflectance distribution models that are useful in simplifying the complex canopy interactions. These models are used by the remote sensing community for modeling BRDF. A brief description of the RossLi BRDF model, which has been used in this research to model forest canopy, has been provided followed by a discussion on DIRSIG. DIRSIG has been extensively used in this research to simulate real-world conditions and to study the varying effects of ground, environmental and sensor characteristics. Section 3.3 introduces the radiometric solver algorithms of DIRSIG used in this research. In Section 3.4, an introduction to the imaging systems is presented with an overview of two sensors in particular, namely, OLI and MSI from Landsat 8 and Sentinel 2 respectively. These instruments have been used in this research for modeling the sources of variation due to sensor parameters. Lastly, the chapter introduces the techniques used in designing the experiments, and highlights the advantage that the techniques offer in the analysis of responses from multiple factor interactions.

3.1 Radiometry

The propagation of Electro-Magnetic (EM) energy from a source to ground and/or sensor can be described using radiometry. The radiometric quantities have been used in a variety of disciplines

and as a result, are subjected to varying definitions and symbolization by different authors. Even in the remote sensing community, there have been inconsistencies in their definitions and proper names, particularly in the definition of reflectance quantities (Schaepman-Strub et al., 2006). To eliminate such ambiguities, this section describes some of the important radiometric quantities used in this research.

3.1.1   Relative Spectral Response (RSR)

The responsivity of a detector can be defined as signal measured per unit flux.

$$R(\lambda) = \frac{dS}{d\Phi(\lambda)} \tag{3.1}$$

Its unit can be volts per watt or amps per watt depending on the signal measured by the detector. The responsivity varies as a function of wavelength and hence is referred to as spectral response function. Spectral response function can be different for each detector in a sensor as each detector can exhibit varying responsivity. The relative spectral response (RSR) is a unitless quantity, obtained by normalizing the spectral response function by its peak spectral response.

$$R'(\lambda) = \frac{R(\lambda)}{R_{max}(\lambda)} \tag{3.2}$$

Similar to spectral response function, RSR for each detector in a sensor can be different. Generally, the spectral response for a detector is characterized by illuminating a tungsten source through a monochromator slit and comparing the response of the detector to a well calibrated reference detector (Barsi et al., 2014). The characterization is performed at the instrument level, hence the measured RSRs gives the combined system-level response including the instrument optics, filter transmission and detector sensitivity.

For pushbroom sensors with wide swath like Landsat 8 and Sentinel 2, there are several thousand detectors and it is not practically feasible to measure the spectral response for each detector in every band. During pre-launch calibration, RSR for few hundred detectors are measured, typically at the center of the Focal Plane Modules (FPM) for each spectral band and at the edges of the focal plane. These measurements are averaged to provide a system-level relative spectral response for each spectral band. Metrics such as mean and standard deviation of the

measured RSR for each spectral band in each FPM can be used to approximate the RSR for every detector in the instrument. Such approximations can be useful in sensor simulations where every detector is provided with a unique RSR. For the remainder of this document, RSR represents system-level averaged relative spectral response for a spectral band unless stated otherwise.

3.1.2 Spectral Radiance

The radiance is defined as radiant flux emitted, reflected, transmitted or received by that surface per unit projected area of that surface along that direction per unit solid angle around that direction.

$$L = \frac{d^2\Phi}{dA \ \cos\theta \ d\Omega} \quad [\ Wm^{-2}sr^{-1} \ ] \tag{3.3}$$

Since the radiance changes as a function of wavelength, it is referred as spectral radiance.

$$L(\lambda) = \frac{d^2\Phi(\lambda)}{dA \ \cos\theta \ d\Omega} \quad [\ Wm^{-2}sr^{-1}\mu m^{-1} \ ] \tag{3.4}$$

Often, we are interested in the radiance measured by the detector. When comparing the radiance measured by two different sensors, the bandpass of the their detectors though operating in the same spectral region, could be different. Since the detector has a spectral response, the effective value of the radiance measured is given by weighting the spectral radiance with the detector's RSR function. Thus the effective spectral radiance is given as

$$L_{\lambda eff} = \frac{\int L_\lambda R'(\lambda)d\lambda}{\int R'(\lambda)d\lambda} \tag{3.5}$$

In practice, numerical integration will be used for calculating the effective spectral radiance. For narrow spectral bands with near flat RSR, the effective spectral radiance are good estimates for actual spectral radiance (Schott, 2007). For the reminder of this document, effective subscript is implied, rather than explicitly indicated for radiance measured by the detectors.

3.1.3 Reflectance

With advancements in field, laboratory, and remote sensing measurements and modeling techniques, there has been a widespread use of angular distribution of reflectance. Satellite data providers generate higher level products such as surface reflectance, top-of-atmosphere reflectance, Bi-hemispherical reflectance or albedo products, etc. However, the mode of measurements or the algorithms used for these product generation do not strictly follow the definitions as formulated by United States National Bureau of Standards (1977). While in most cases, these deviations from the original definitions are implicitly understood, their usage could still lead to misinterpretation among the scientific user communities. Schaepman-Strub et al. (2006) has presented the definitions for most commonly used reflectance quantities based on United States National Bureau of Standards (1977) and also provided some examples of the erroneous usage of the reflectance terminology. Following their definitions, reflectance terms used in this document are defined as below. Any changes to the definition are clearly stated here for ease of understanding and to eliminate any misinterpretations.

The total spectral reflectance is defined as the ratio of spectral radiant exitance with the spectral irradiance. The reflectance factor is the ratio of the spectral radiant flux reflected by a surface to that reflected into the same reflected beam geometry by an ideal and diffuse (Lambertian) standard surface irradiated under the same conditions. For measurement purposes, Spectralon panels are used to approximate the ideal diffuse standard surface.

$$\rho(\lambda) = \frac{M(\lambda)}{E(\lambda)} \tag{3.6}$$

$$r(\lambda) = \frac{\rho(\lambda)}{\rho_{spectralon}(\lambda)} \tag{3.7}$$

where,

$\rho(\lambda)$ is the spectral reflectance of an object

$M(\lambda)$ is the spectral exitance from an object

$E(\lambda)$ is the spectral irradiance incident on an object

$r(\lambda)$ is the spectral reflectance factor of an object

$\rho_{spectralon]}(\lambda)$ is the spectral reflectance of a reference such as spectralon

While reflectance ranges from 0 to 1 following the law of conservation of energy, the reflectance factor can have values larger than 1, especially for strong forward reflecting surfaces such as snow (Painter and Dozier, 2004).

In many instances, users are interested in the directional distribution of the reflected flux. Surfaces can exhibit varying directional characteristics of reflectance ranging from perfect specular or mirror-like reflections to completely non-directional characteristics of a perfect diffuser as shown in Figure 3.1.
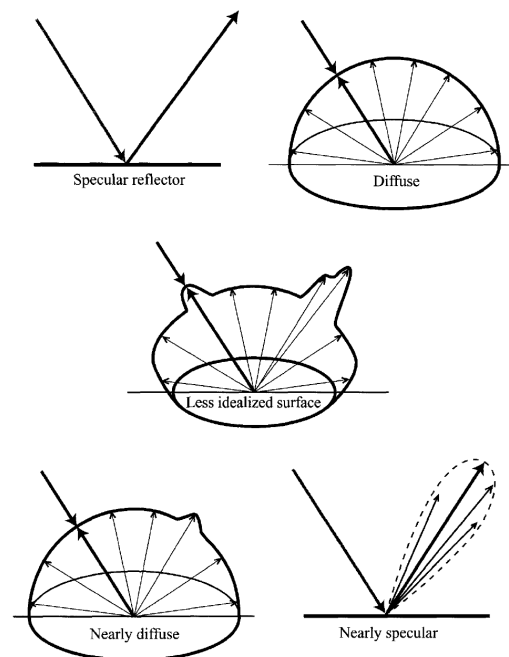
Figure 3.1: Reflectance characteristics of idealized surfaces (Schott, 2007).

United States National Bureau of Standards (1977) has broadly defined nine standard reflectance terms by considering the incident and reflected flux geometries. These geometries are shown in Figure 3.2.

The shaded geometries in Figure 3.2 denote those geometries that are practical, particularly in the remote sensing of ground surfaces using field instruments and satellites. Of these nine geometries, bi-directional and bi-conical reflectance quantities are of importance in this research.
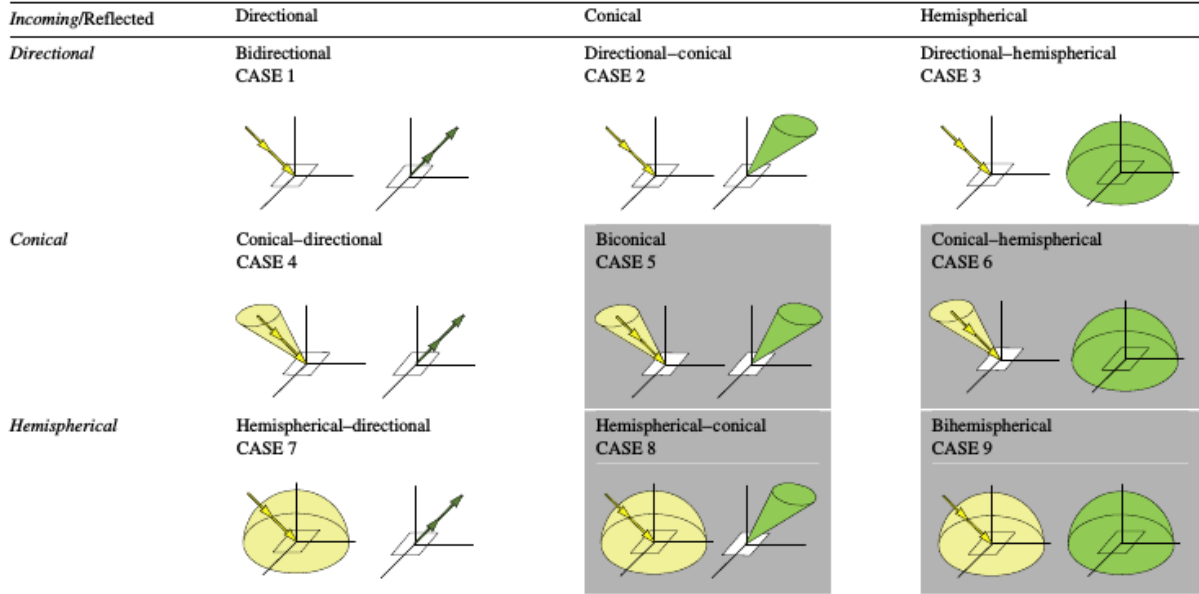
Figure 3.2: Relation of incoming and reflected radiance terminology used to describe reflectance quantities (Schaepman-Strub et al., 2006).

3.1.3.1 Bidirectional Reflectance distribution function (BRDF)

BRDF describes the scattering of a parallel beam of incident light from one direction in the hemisphere into another direction in the hemisphere. It describes the intrinsic reflectance properties of a surface and is expressed in Equation 3.8. The directional component of BRDF is shown in Figure 3.3.

$$f(\phi_i, \theta_i, \phi_r, \theta_r, \lambda) = \frac{\partial L}{\partial E} \quad [sr^{-1}] \tag{3.8}$$

Since it is expressed as a ratio of infinitesimal quantities, it cannot be directly measured, as real measurement must be made over a finite interval (United States National Bureau of Standards, 1977). However, many physical quantities have similar infinitesimal quantities in their definition, but are widely accepted as measurable quantities as they are closer estimates of the truth. Similarly, in our context, any references to BRDF in this document refers to an approximate estimate of the BRDF which cannot be measured. This is implicitly understood but stated explicitly here to avoid any misinterpretations.

Bidirectional reflectance factor (BRF) is the ratio of the reflected radiant flux from a surface area ∂A to the reflected radiant flux from an ideal and diffuse surface of the same area ∂A under identical illumination and view geometry. BRF is a unitless quantity as it is a ratio of two radiant fluxes.
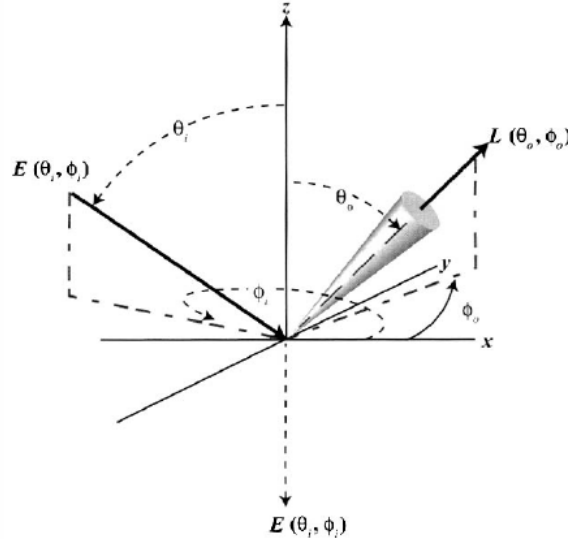
Figure 3.3: Bi-directional reflectance geometry (Schott, 2007).

$$r(\phi_i, \theta_i, \phi_r, \theta_r, \lambda) = \frac{\partial \Phi_r}{\partial \Phi_{r(ideal)}} \quad [unitless] \tag{3.9}$$

An ideal and lambertian surface reflects the same radiance in all directions and its BRDF is $\frac{1}{\pi}$. Thus, BRF can be simply expressed as

$$r(\phi_i, \theta_i, \phi_r, \theta_r, \lambda) = \pi \; f(\phi_i, \theta_i, \phi_r, \theta_r, \lambda) \tag{3.10}$$

Sometimes it is easier to visualize the directional reflectance characteristics of a surface as BRF since they can be easily related to the total reflectance factor. For example, a green leaf typically expressed as 0.5 (50%) in reflectance factor in the near infrared region when viewed in the laboratory can be expressed in BRF as 0.5 in a particular view and illumination angle (nadir, in this case). Thus, BRF is similar to reflectance factor but with the directional component of the source and viewer.

3.1.3.2 Hemispherical-Conical Reflectance Distribution Function (HCRDF)

BRDF and HCRDF differ only in the geometry, specifically the solid angle of the incoming irradiance and the outgoing radiance.

$$HCRDF(\phi_r, \theta_r, \omega_r, \lambda) = \frac{1}{\pi} \int_{2\pi} \int_{\omega_r} f(\phi_i, \theta_i, \phi_r, \theta_r, \lambda) \; \partial \Omega_i \; \partial \Omega_r \tag{3.11}$$

Typical field measurement or measurements from satellite have a detector that has an instantaneous field of view with a finite solid angle. The illumination of the ground by sun has both a direct and diffused sky light components resulting in a hemispherical irradiance on to the surface. This is what most of the instruments measure in the field. In the laboratory, conditions could be imposed on the illumination in which case it would be a bi-conical reflectance distribution function. BRDF simulations using DIRSIG, strictly speaking, are bi-conical reflectance measurements. However, the solid angles of the detector and that of the sun over a small region on the ground are extremely small such that these measurements are considered as bi-directional reflectance measurement in this document. In this document, the term bi-directional includes any spectral component of the bi-directional reflectance, unless otherwise stated.

3.2   BRDF Modeling

BRDF is a useful radiometric quantity that can be used to standardize reflectance observations with varying sun-view geometry to a common standard geometry as necessary in image mosaicking (Strugnell et al. (1998), Leroy and Roujean (1994)). It is highly imperative that reflectance observations from multiple datasets are standardized to be useful in monitoring and performing long and short-term trend and change detections. Since BRDF of land surface is directly related to the optical properties of materials (such as vegetation, forest, etc), they reflect a variety of natural and human influences which are of interest to global change research and climate studies (Lucht et al., 2000). BRDF can be measured in the laboratory (Feng et al., 1993) or in the field using goniometers (Deering (1988), Biliouris et al. (2007)). However, it is impractical and sometimes even impossible (particularly field measurements) to measure reflectance in all different sun-view geometries. Also, BRDFs are available for only a relatively restricted number of materials and land cover types as it is difficult to measure BRDF for many land cover types. Further, any redundant measurements for robustness is extremely rare (Schott, 2007). As a result, researchers have investigated the possibility of modeling the BRDF for different material types. Since BRDF is widely used across many disciplines, any criteria on the classification of BRDF models tends to vary. In this research, the focus is on the BRDF of the forest canopy, and hence classification of BRDF models pertaining to forest

canopy are discussed. The BRDF models can be broadly classified into three types, physical models, empirical models and semi-empirical models.

Physical BRDF models describe the interactions of the incident radiations on a specific target (forest canopy), usually in terms of the biophysical parameters that govern the radiative transfer within the canopy (leaf area index, leaf optical properties, angular distributions, chlorophyll content, etc). These models are often complex and involve a large number of parameters some of which may be correlated. They are generally non-linear in formulations and require numerical methods for inversion which may lead to instability depending on the number of observations. Due to their complexity in nature, these models are primarily used for validation of other models and measurements and are strictly used only in the research domain (Chopping, 1998).

Empirical BRDF models do not explain biophysical parameters or the process which govern BRDF, but rather try to use mathematical descriptions to fit the patterns in the bidirectional reflectance observations. They can range from simple second-order polynomial functions (Gutman, 1991) which takes variation in viewing zenith angle into consideration, to complex polynomials. The parameters in these functions may not relate to any physical quantities, however, they may be able to describe a wide range of observed BRDF shapes with reasonable accuracy and are easily invertible (Walthall et al., 1985). Using these parameters to model unobserved sun-view geometries can lead to predictions that can be completely unrealizable.

Semi-empirical BRDF models are hybrid models that incorporate anisotropic scattering behaviors using simplified kernel functions. These models use approximations to the physical interactions to formulate a linear or non-linear function that is a combination of both isotropic and anisotropic scattering. The semi-empirical models are typically a function of three terms, the isotropic term accounting for the uniform scattering, the geometric function accounting for the effects of mutual shadowing and geometric structure, and the volumetric function which is based on the radiative transfer in a turbid medium. There are several semi-empirical BRDF models such as Roujean model (Roujean et al., 1990), RossLi BRDF model (Wanner et al., 1995), RPV model (Rahman et al., 1993), etc. There are many advantages and dis-advantages of these semi-empirical models, particularly that of linear semi-empirical models such as RossLi model. The advantages for the linear semi-empirical models are: they retain some physical meaning in their parameters unlike fully empirical methods, have small number of parameters unlike physical models, obey the Helmholtz reciprocity, and also can be inverted analytically without

predicting unrealistic values. One of the major dis-advantages is that the models cannot provide biophysical parameters such as Leaf Area Index (LAI), surface roughness, canopy structure, etc. Another dis-advantage, that is of interest in this research, is that there is a high probability of inaccuracies in predicting BRDF in unobserved regions (Chopping, 1998). Section 4.3.2 explains how this inaccuracy can be mitigated by appropriate sampling and measurement of BRDF using DIRSIG. The advantages that these methods provide outweigh their disadvantages. Chapter 5 will discuss the goodness of the fit between the RossLi model and the measured data for the simulated forest using DIRSIG. In this research, the RossLi BRDF model is used extensively and is briefly described in the next section.

3.2.1 RossLi BRDF Model

The RossLi BRDF model is a linear semi-empirical BRDF model formulated based on volumetric scattering kernel functions by Ross (Ross, 1981, pp. 392) and geometric scattering kernel function by Li (Li and Strahler, 1992). Wanner et al. (1995) has provided the necessary approximations and formulations for the RossLi BRDF model. An adaptation to this model for the reciprocal condition is found in Lucht et al. (2000).

Typical semi-empirical models either contain only geometric terms without any physical parameters, such as Ross canopy model, or it contains one or very few physical parameters as in the case of Li model. In the RossLi BRDF model, both of these characteristics are considered to derive a linear model.

3.2.1.1 Ross Model

The Ross-thick kernel has been derived by Roujean et al. (1992) while a modified version to account for smaller LAI was done by Wanner et al. (1995). The Ross model gives the bidirectional reflectance above a horizontally homogeneous plant canopy based on the radiative transfer theory in a single scattering approximations. This is different from what happens in reality, where, a single photon scatters more than once on leaves, tree trunks, ground, etc. As will be discussed later, DIRSIG in its radiometric solution, uses multiple scattering phenomenon to estimate the canopy leaving radiance. In the Ross model, the leaves are located randomly above a flat horizontal surface with a Lambertian reflectance. Also, the leaves are assumed to

have Lambertain reflectance and Lambertian transmittance, such that their reflectance equals to their transmittance. This is approximately true in the NIR region of the spectrum, however, the transmittance of leaves are higher than the reflectance in the visible region. The type of Ross model (thick or thin) depends on the Leaf Area Index (LAI) of the canopy. Large LAI canopies ($LAI >> 1$) are modeled using Ross-thick kernel and small LAI canopies are modeled using Ross-thin kernel. The Ross-thick kernel function can be expressed as

$$R_{thick} = C_1 \ K_{thick} + C_2 \tag{3.12}$$

where, C_1 and C_2 are constants.

Further, the kernel K_{thick} is assumed to be zero when the illumination and viewing zenith angles are zero. i.e, when the sun and sensor is perfectly nadir. In such a case, the Ross reflectance equals the constant C_2.

$$K_{thick} = \frac{(\pi/2 - \xi)\cos\xi \ + \ \sin\xi}{\cos\theta_i \ + \ cos\theta_v} - \frac{\pi}{4} \tag{3.13}$$

$$C_1 = \frac{4s}{3\pi}(1 - e^{-LAI*B}) \tag{3.14}$$

$$C_2 = \frac{s}{3} + (\rho_0 - \frac{s}{3}) \ e^{-LAI*B} \tag{3.15}$$

where,

$$\cos\xi = \cos\theta_i \ \cos\theta_v + \sin\theta_i \ \sin\theta_v \cos\phi \tag{3.16}$$

$$B = \frac{1}{2(\sec\theta_i + \sec\theta_v)} \tag{3.17}$$

and

s is leaf reflectance

ξ is the phase angle of scattering

LAI is the leaf area index

ρ_0 is the surface reflectance

θ_i is the zenith angle of illumination

θ_v is the zenith angle of view

ϕ is the relative azimuth angle between illumination and view

For the Ross Thick model, constant C_1 will be the weight of the thick volume scattering kernel and C_2 is the isotropic scattering constant.

Ross Thin Kernel :

The Ross thin kernel derivation is similar to the thick kernel except that when the LAI is very small, a first order approximation for the exponential function is used. In this case, since the LAI is small, the photons scattered by the layer beneath the canopy (litter or soil) are assumed to be either absorbed or undergo multiple scattering, and thus becomes isotropic. The Ross thin kernel function can be expressed as ,

$$R_{thin} = C_1 \; K_{thin} + C_2 \tag{3.18}$$

where, C_1 and C_2 are constants.

Similar to thick kernel, it is assumed that the thin kernel is zero for nadir illumination and nadir viewing.

$$K_{thin} = \frac{(\pi/2 - \xi)\cos\xi \; + \; \sin\xi}{\cos\theta_i \; \cos\theta_v} - \frac{\pi}{2} \tag{3.19}$$

$$C_1 = \frac{2s \; LAI}{3\pi} \tag{3.20}$$

$$C_2 = \frac{s \; LAI}{3} + \rho_l \tag{3.21}$$

where

ρ_l is the average Lambertian reflectance of the layer beneath the governing thin canopy on top.

The phase angle of scattering is the same for both thin and thick kernel as shown in Equation 3.16. The constant C_1 will be the weight of the thin volume scattering kernel and C_2 is the isotropic scattering constant.

3.2.1.2   Li Model

The Li kernel approach was developed by Li and Strahler (1986) where the reflectance of the scene is given by the areal proportions K of sunlight crown (reflectance C), sunlight ground (reflectance G), shadowed crown (reflectance K), and shaded ground (reflectance Z).

$$R_{geo} = CK_c + GK_g + TK_t + ZK_z \tag{3.22}$$

In its derivation, the shadows were assumed to be perfectly black and hence T = Z = 0. Another assumption is that the sunlit crown and the sunlit ground have the same reflectance to reduce the complexity in its derivation. However, this is typically not the case across the VIS-NIR-SWIR spectrum, where the soil reflectance or litter reflectance tends to be slightly higher than the leaf reflectance. Thus the equation reduces to

$$R_{geo} = C(K_c + K_g) \tag{3.23}$$

Therefore, the entire model depends primarily on the proportion of the sunlit crown K_c and sunlit ground K_g.

As in the case of Ross model, the canopy can be sparse or dense, and hence two different formulation of Li kernel models were formulated accordingly. In the dense canopy, there is a larger contribution due to mutual shadowing and hence cannot be ignored as in the case of sparse canopy. More details on the derivation of the Li kernels for sparse and dense canopy can be found in Wanner et al. (1995). Li model formulations are linear, similar to Ross model. Li sparse model can be expressed as,

$$R_{sparse} = C_1 \ K_{sparse} + C_2 \tag{3.24}$$

$$K_{sparse} = O(\theta_i, \theta_v, \phi) - \sec\theta_i' - \sec\theta_v' + \ (1 + \cos\xi') \ \sec\theta_v' \tag{3.25}$$

$$C_1 = C\lambda\pi r^2 \tag{3.26}$$

$$C_2 = C \tag{3.27}$$

Li dense model is expressed as,

$$R_{dense} = C_1 \ K_{dense} + C_2 \tag{3.28}$$

$$K_{dense} = \frac{(1 + \cos\xi') \ \sec\theta_v'}{\sec\theta_v' \ + \ \sec\theta_i' - O(\theta_i, \theta_v, \phi)} \ - \ 2 \tag{3.29}$$

$$C_1 = \frac{C}{2} \tag{3.30}$$

$$C_2 = C \tag{3.31}$$

where,

$$O = \frac{1}{\pi}(t - \sin t \cos t)(\sec \theta'_i + \sec \theta'_v) \tag{3.32}$$

$$\cos t = \frac{h}{b} \frac{\sqrt{D^2 + (\tan \theta'_i \tan \theta'_v \sin \phi)^2}}{\sec \theta^i_i + \sec \theta'_v} \tag{3.33}$$

$$D = \sqrt{\tan^2 \theta'_i + \tan^2 \theta'_v - 2 \tan \theta'_i \tan \theta'_v \cos \phi} \tag{3.34}$$

$$\cos \xi' = \cos \theta'_i \cos \theta'_v + \sin \theta'_i \sin \theta'_v \cos \phi \tag{3.35}$$

$$\theta' = \tan^{-1}(\frac{b}{r} \tan \theta) \tag{3.36}$$

and

C is the crown reflectance

r is the radius of the sphere when vertical-scale transformation is performed to make spheroidal crown to a spherical crown

The constant C_1 denotes the weight of the sparse or dense scattering kernel in a complete kernel-driven model and constant C_2 is included to represent the isotropic scattering constant.

The Li kernels are non-linear since it has two parameters b/r and h/b describing crown shape and relative height. Thus, Li kernels can provide a family of kernels depending on h/b and b/r parameters. For MODIS processing, these values are set to 2 and 1 respectively (Schaaf et al., 2002).

The original form of the Li kernel is not reciprocal in illumination and viewing zenith angles. For homogeneous natural surfaces, reciprocity is expected at coarser resolution (Lucht et al., 2000). In the original derivation, the reflectance of the scene components are assumed to be constant. By assuming their variation to be a function of illumination zenith ($\frac{1}{\cos \theta'_i}$), the reciprocal kernel can be given as

$$K_{sparse} = O(\theta_i, \theta_v, \phi) - \sec \theta'_i - \sec \theta'_v + (1 + \cos \xi') \sec \theta'_v \sec \theta'_i \tag{3.37}$$

$$K_{dense} = \frac{(1 + \cos \xi') \sec \theta'_v \sec \theta'_i}{\sec \theta'_v + \sec \theta'_i - O(\theta_i, \theta_v, \phi)} - 2 \tag{3.38}$$

The overlap and phase angle of scattering has the same form for both reciprocal and non-reciprocal kernel as shown in Equation 3.32 and Equation 3.35 .

The Ross and Li kernels can be used together to represent the BRDF of the forest canopy (Schaaf et al., 2002). The combined Ross and Li BRDF model is a semi-empirical linear model and can be expressed as,

$$R(\theta_i, \theta_v, \phi, \lambda) = f_{iso}(\lambda) + f_{vol}(\lambda) \ K_{vol}(\theta_i, \theta_v, \phi) + f_{geo}(\lambda) \ K_{geo}(\theta_i, \theta_v, \phi) \qquad (3.39)$$

where,

$R(\theta_i, \theta_v, \phi, \lambda)$ is BRDF (or BRF) of the canopy

f_{iso} is the isotropic scattering

f_{vol} is the radiative transfer-type volumetric scattering as from horizontally homogeneous leaf canopies

f_{geo} is the geometric-optical surface scattering from scenes containing 3D objects that cast shadows and are mutually obscured from view

The volumetric scattering term expresses the effects caused by the interleaf gaps in the canopy and the geometric-optical term expresses the effects caused by the intercrown gaps with in the forest. In Equation 3.39, the Ross model is used for estimating the volumetric kernel and Li model is used for geometric kernel. Depending on the type of forest modeled, one may use thick or thin Ross kernels and sparse or dense Li kernels. In this research, reciprocal Li kernels are used for both sparse and dense canopy models.

3.2.2 Visualization of BRDF (or BRF)

The visualization of BRDF helps to understand the structural and optical properties of the surface or materials much more clearly than looking at a table of reflectance values. BRDF are functions of illumination angles (zenith, azimuth), view angles (zenith, azimuth) and wavelength and hence can be considered as 5 dimensional (5D) dataset. It is hard to show visualization of more than 2 dimensional (2D) dataset clearly on a document. Many techniques have been developed by the computer graphics community to show 3 dimensional data using volume rendering, isosurfaces techniques, etc. Typically, any higher dimensional data ($> 3D$) are represented by projections on to lower dimensions. For BRDF, we would use a similar approach. BRDF visualization should primarily focus on clear indication of interesting characteristics such as "bowl shaped" or "bell shaped" reflectance profiles observed along the solar principal plane

for certain materials. Three different methods of visualizing BRDF are used in this research for better understanding of the forest canopy or material optical properties, and are described in the following subsections.

3.2.2.1 Polar surface plot

Polar coordinate system is a 2D coordinate frame in which each point is represented by a distance from a reference point (origin) and an angle from a reference direction. Thus any point can be referenced in 2D by (r, θ). For polar BRDF plot, zenith angles are used as distance r (zenith is zero at origin) and azimuth angle measured in clockwise direction as angle theta from the reference direction (North in the local coordinate system). Typically, contour or isolines are drawn by connecting the same reflectance or reflectance factor. In some cases, symbols are used to represent the different reflectance for varying zenith and azimuth angles. In this document, reflectance for these azimuth and zenith angles are plotted with a suitable color based on a color-map. Keeping 3 dimensions fixed, such as wavelength, illumination zenith and illumination azimuth of BRDF, one can visualize the reflectance distribution across all view angles using this type of plot as shown in Figures 3.4 and 3.5.
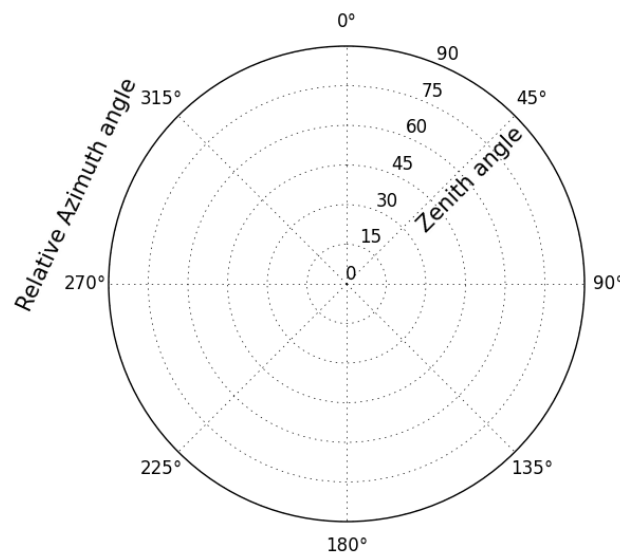
Figure 3.4: Plot showing the azimuth and zenith angles as two axes of polar plot coordinate system.
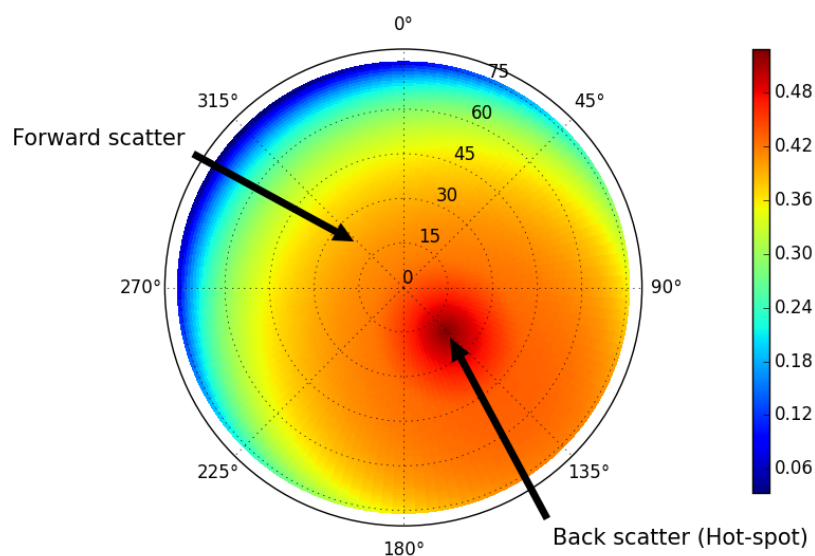
Figure 3.5: Plot showing BRF in polar plot with corresponding colormap. The BRF is for varying view angles when Sun Azimuth is 135° and Sun Zenith is 20°

3.2.2.2 3D Plot

There are many ways in which a surface can be represented in 3D. 3D rectangular plots are typically used to represent surface such as digital elevation model (DEM) as a mesh or wire-frame model.

The polar surface plot can be made visually more appealing by representing it as 3D plot. In this document, BRDF is shown in 3D format using spherical coordinate system rather than using the rectangular coordinate system. This is done by converting the azimuth angle, zenith angle, and reflectance to 3D rectangular coordinates as shown in Appendix A. A perfect Lambertian surface is represented by a hemisphere as shown in Figure 3.6. One of the advantages of this plot is that the differences in reflectance for different directions can be observed without a need for color-map. Even though this plot allows for better representation of BRDF on a computer using a visualization tool, on a 2D document, these plots can be viewed clearly only for very few angles at a time. Due to this limitation, these plots are hardly used in this document.

3.2.2.3 Principal plane and Cross plane plot

In general, natural surfaces show most variability in the solar principal plane. The principal plane is defined as a plane formed by the source of illumination, target and viewing sensor. In the

principal plane, the viewing azimuth angle is either the same azimuth angle as the illumination source or offset by 180 degrees from the source azimuth angle. The cross plane is defined as a plane perpendicular to the principal plane. The reflectance differences in the cross plane and principal plane can provide a good visualization of the BRDF in these important directions. Reflectance for varying zenith angle along the principal and cross planes can be plotted as 2D line plot (or scatter), as shown in Figure 3.7. This plot has an advantage of showing the backscatter and forward scattering phenomenon observed by many natural surfaces including forest canopies.
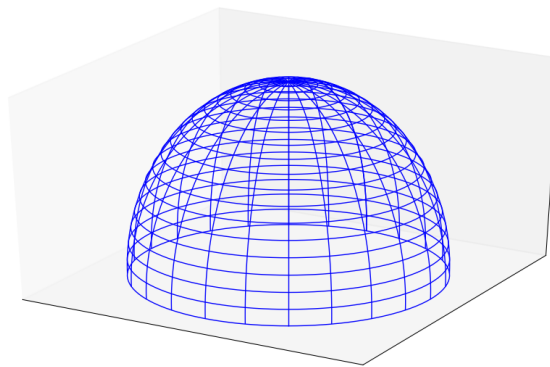
Figure 3.6: 3D Plot of BRF for Lambertian surface. Lambertian surface has equal reflectance across all directions as shown in 3D plot as hemisphere.
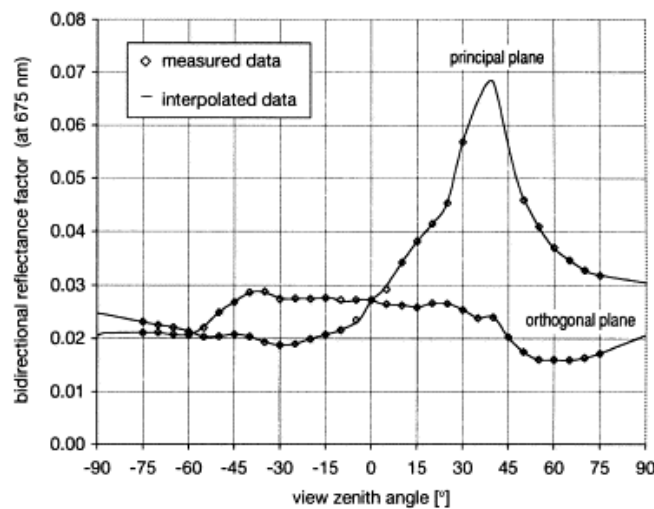
Figure 3.7: Measured versus interpolated reflectance data of the grass lawn in the principal and cross planes (Sandmeier et al., 1998)

For forest canopies, the hot-spot is a bright region where the reflectance (due to back scattering effect) is highest in comparison to any other viewing angles. Since this occurs along the principal plane, these plots help to understand the importance and the extent of the reflectance in the hot-spot region. However, the spread of the hot-spot region in the azimuth direction cannot be easily observed, and hence this type of plot is not used as widely as the polar surface plot.

In this document, the zenith angles are positive when the azimuth angle of the viewing sensor and illuminating source are the same unless explicitly stated.

3.3   DIRSIG

The Digital Imaging and Remote Sensing Image Generation (DIRSIG) model is a first principle based radiation propagation models actively developed by Rochester Institute of Technology (RIT) since the late 1990s (Brown and Goodenough, 2015) for synthetic image generation. DIRSIG can produce passive single-band, multi-spectral, hyper-spectral, and polarized imagery from the visible through the thermal infrared region and also has the capability to simulate active laser using Light Detection and Ranging (LIDAR) techniques, and active radio frequency (RADAR) image modalities. DIRSIG can be used to test image system designs, provide test imagery for exploiting algorithms and can help in system trade studies. In this research, DIRSIG will be used as the experimental engine for performing design of experiment analysis. The DIRSIG tool has been validated for its overall performance and accuracy over the past decades which are summarized in Brown and Schott (2010). The validation of DIRSIG for forest canopy studies is presented in Chapter 5. The main components of DIRSIG modeling are scene modeling, sensor modeling, platform motion and tasking, and atmosphere modeling. This is illustrated in Figure 3.8. The radiometric computations are performed in DIRSIG by a specialized set of tools known as Radiometry solvers.

3.3.1   Scene Modeling

The key features of the DIRSIG model is that all modalities are simulated from a common scene description. The scene is composed of 3D geometrical objects (polygons, mathematical objects) which are assigned a material description in the spectral domain of interest. The 3D geometric models can be imported from a variety of tools (AutoCAD, Blender3D, OnyxTree, etc) into

DIRSIG. The material description includes thermodynamic properties to enable temperature prediction and optical properties to drive radiometric prediction and propagation.
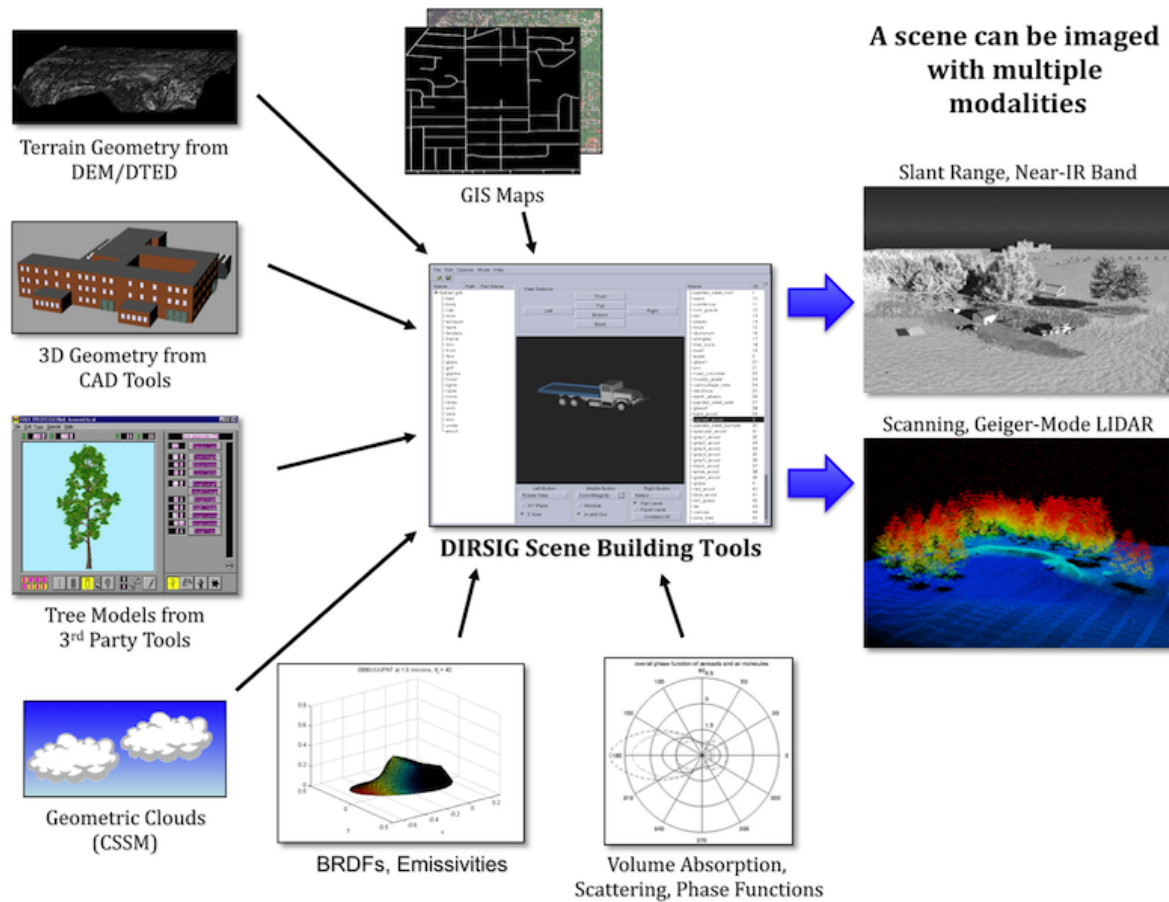
Figure 3.8: DIRSIG Scene modeling across all modalities employing a single scene database (Brown and Goodenough, 2015)

The collection of 3D objects can be placed as static or can be positioned dynamically in a coordinate system of choice. The choice of coordinate system is dependent on the type of scene being modeled and the size of it. For example, if the scene being modeled is small, then a flat earth assumption is valid and hence a local coordinate system such as Easting, Northing, Up (ENU) coordinate system can be used. For large area scene modeling, the curvature of the earth needs to be accounted for, and in such cases, the scene can be modeled either in a cartesian coordinate system such as Earth Centered Earth Fixed (ECEF) or in geogrpahic coordinates using latitude, longitude and altitude. Internally, irrespective of the scene coordinate system used, DIRSIG will automatically translate all the coordinates into the required coordinate system depending on the platform modeling and simulation.

3.3.2 Sensor Modeling

The DIRSIG instrument module was designed to provide a framework for many types of sensor. It could be passive imaging sensors with single pixel scanning or multi-pixel whiskbroom architecture, 2D imaging arrays, and pushbroom sensors with multispectral or hyperspectral configurations. DIRSIG can also model LIDAR and SAR sensors for active imaging. For the passive imaging sensors with one or more detectors, spatial detector layout can be modeled by providing pixel size, pixel pitch or by providing line of sight angles for each detector. Spectral sensitivity of each detector is modeled either by assigning detector specific RSR or using a mathematical expression. DIRSIG uses ray tracing techniques to estimate the incoming radiance in front of the aperture for each detector and also provides options to increase the fidelity of the measurement by increasing the number of ray samples per detector. This is illustrated in Figure 3.9, where a single detector randomly samples the radiance from nine different location within the Ground Instantaneous Field Of View (GIFOV). The nine samples are averaged to estimate the radiance reaching the detector's IFOV.
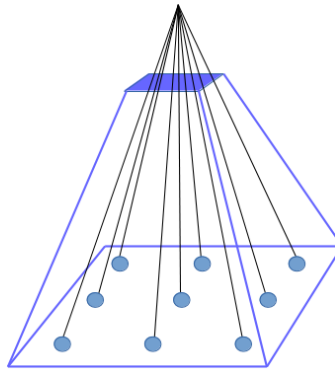
Figure 3.9: Detector sub-sampling strategy used in DIRSIG. Three sub-samples along x and three sub-samples along Y are used to estimate the average radiance reaching detector's IFOV.

3.3.3 Platform motion and tasking

The DIRSIG model features an advanced data acquisition model using a concept of platform which is analogous to the spacecraft bus. Multiple sensor payloads can be mounted on the platform. This modular design allows each sensor to be spatially translated and/or rotated from each other and can support more complex acquisition geometries. The motion of a platform

can be provided using a data driven or parameterized descriptions such as fixed trajectory using orbital parameters or using Two Line Element (TLE) descriptions. DIRSIG can also support a static acquisition by providing the position and orientation of the platform in a specific coordinate system. The platform position can be provided in a local coordinate system or in a global coordinate system, and the orientations can be provided either as Euler angles or quaternions. For orbiting satellites, it is best to specify the platform motion using TLE. The tasking aspect of platform modeling is analogous to instrument on and off times. The tasking provides the ability to start and stop acquiring the synthetic scene based on the platform motion and time.

3.3.4 Atmospheric modeling

DIRSIG has a suite of interface modules that leverage externally developed atmospheres like MODerate resolution atmospheric TRANsmission (MODTRAN) (Brown and Goodenough, 2015). DIRSIG has many options to define the atmospheric optical properties. In its simplest case, one can define the source irradiance without any atmosphere for the simulations. This option is useful particularly for testing the radiometric computations of the surface materials or for generating the BRDF of materials. The DIRSIG model can use MODTRAN input files to define the atmosphere as a function of altitude and wavelength.

3.3.5 Radiometry Solver

In DIRSIG, the radiometric calculations due to the interaction of different objects in the scene with the incident photons are performed by a set of tools which are called as Radiometry Solver (radsolver). The radsolver predicts or estimates the radiant flux from a specific radiational element for a specific set of conditions. The radiational element could be a surface such as a leaf, trunk of a car, etc. or a volume element such as clouds, water, etc. The radsolvers provide the capability to perform the radiative transfer to compute the flux contribution along the path and through the element (Brown and Goodenough, 2015). The radsolvers uses the material or optical properties of the surface as defined in the constructed scene for computing the radiation contributions. In DIRSIG, there are a few specialized radsolvers that are designed and optimized for specific models and radiational elements. In this research, the "Geodensity" radsolver is used extensively for modeling the scattering interactions in a forest canopy.

3.3.6 Geodensity Radiometry Solver

In a forest canopy, the primary scatterers are leaves of the trees which reflect as well as transmit the incoming photon. A single photon can interact with many scattering elements (leaf) before it is either absorbed by an absorbing element or leaves the canopy. The contributions due to multiple scattering in the forest canopy are very high, particularly in the NIR region where the leaves have very high reflectance and transmittance. Capturing the multiple scattering effects using a hierarchical ray tracing radsolver in DIRSIG is computationally not feasible as it is an exponential growth problem (Goodenough and Brown, 2015). A surface photon map based approach as explained by Goodenough et al. (2006) is suited for capturing the multiple scattering effects. It is based on Monte Carlo propagation of photon bundles to build an abstract map of flux density across surfaces and use the flux density map to approximate the local multiple-scattered contributions. The photon mapping approach cannot be directly applicable for tree modeling since the map will be very sparse and the search areas will have to be small since the leaves themselves are very small. This can impact the local density estimate. These limitations are overcome by associating photon bundles with the geometry components such as facets and primitives rather than an abstracted structure in the photon map. As a result, a very good approximation of the flux density can be estimated since the area of these facets are known. This radsolver is referred to as Geodensity radsolver.

The spatial fidelity of the estimate is tied to the fidelity or the detail in the constructed scene model which itself can impose a strict requirement and limitation that the geometry elements are fairly small and uniform for this radsolver. Fortunately, these requirements are easily satisfied in the case of leaves, which usually are small and are modeled geometrically with smaller facets.

One of the requirements for the canopy scattering model is to account for the scattering in the leaf volume due to photons arriving from outside the immediate area. For a continuous canopy, this effect is more pronounced at the boundaries. This is accomplished by using a cyclic propagation of photon bundles at the boundaries, i.e., any photon leaving from one of the boundaries does not really escape the canopy but acts as an incoming photon from the opposite boundary. This is illustrated in Figure 3.11. A photon leaving from one side of the canopy shown as an arrow from a location (x,y) and oriented at an angle θ, re-emerges from the other side from the same position and orientation. The photons leaving from the top of the canopy alone escapes, while the photon leaving from the canopy boundaries recirculate to interact with

the canopy. The cyclic propagation emulate the characteristics observed in a large homogeneous canopy surface where the boundaries are not considered while modeling the canopy surface.
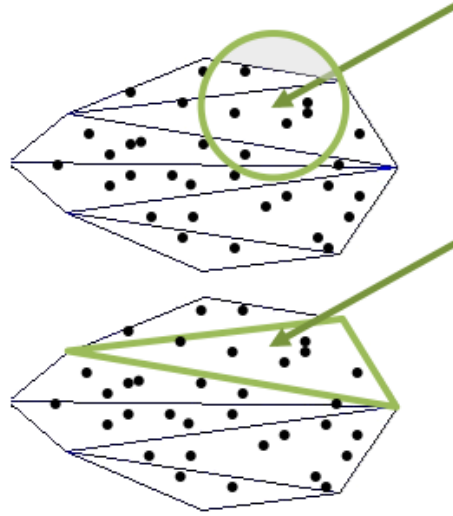
Figure 3.10: Difference between photon map (top) and Geodensity (bottom) radsolvers. The edge overlapping issues observed in photon map due to abstract search area is not present in the geodensity radsolver, since it uses the surface area of the geometry primitives (facets) for estimating the flux density (Goodenough and Brown, 2015).
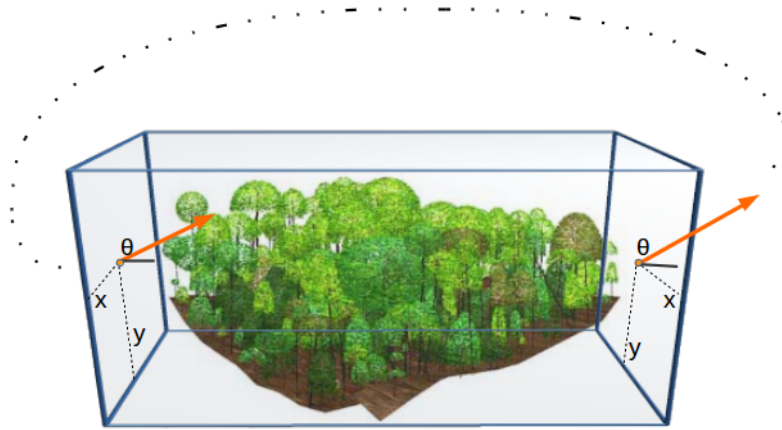
Figure 3.11: Cyclic propagation of photon bundles in the Geodensity radsolver. A photon leaving the canopy from one side of the boundary at a specific position and orientation (x,y,θ), re-emerges from the opposite side at the same position and orientation to emulate the characteristics observed in a large homogeneous canopy.

3.4 Sensors

Optical remote sensing sensors are primarily designed for land cover and land use monitoring and for near and long-term resource monitoring and trend analysis. Although these sensors work in the same optical domain, from the visible to shortwave infrared region (occasionally, thermal infrared region), their design parameters and their sensitivities are quite different. This section will provide a background on the types of remote sensing systems with an emphasis on sensors from the Landsat 8 and Sentinel 2 missions.

Remote sensing systems can be divided into three general categories: whiskbroom sensor, pushbroom sensor and frame sensor. Whiskbroom sensors uses a rotating mirror to scan the ground in the across track direction while the platform motion helps to image the ground in the along track direction. A whiskbroom sensor can have either a single detector (line scanners) or a linear array of detectors in the along track direction. Since a line scanner has a single detector, its design allows for simple optics. Also, radiometric calibration is easier since the entire ground is imaged using the same detector. However, the swath width coupled with platform motion in the along track reduces the dwell time for each detector's Instantaneous Field Of View (IFOV). It is for this reason that the line scanners are typically found in aerial sensors and are seldom used for observations from satellite platforms where the platform velocity at low earth orbit further reduces the dwell time and lowers the Signal to Noise Ratio (SNR).

Using an array of detectors (in the along track direction) can increase the dwell time for each detector and can improve the SNR compared to a line scanner without requiring a large optical design. However, the increase in dwell time is realized only by having the scan mirror rotate at a slower rate. Since the platform motion is continuous and the rotating mirror takes finite time to scan from one end to the other, typically, a scan line corrector mirror is used to account for the scan line gaps due to platform motion if the data are collected in both directions of the mirror rotation. Examples for these type of sensing system includes sensors such as Multispectral scanner (MSS), Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) in earlier Landsat missions (Landsat 1 – Landsat 7).

Pushbroom sensors have an array of detectors in the across track direction covering the swath while the image in the along track direction is acquired by the platform motion. Thus, each detector has a longer integration time resulting in improved SNR performance compared to line

scanners or whiskbroom sensors. Since the detectors need to cover the entire Field Of View (FOV), the focal plane for these sensors are large compared to whiskbroom sensors. The spectral information is collected by using spectral filters placed in the along track direction. Thus, each spectral band has varying view angle in the along track direction. In the across track direction, the FOV is covered by a row of detectors for each spectral band that are either continuous or segmented. For example, in Landsat 8, the FOV is about 15 degrees, and to cover this entire swath, about 6000 detectors are required (multi-spectral band). It is sometimes difficult to fabricate many detectors with good detector characteristics such as responsivity, quantum efficiency, etc. To improve the fabrication efficiency and performance of the detectors, the entire row of detectors covering the FOV is segmented into smaller sections which will be referred as Focal Plane Modules (FPMs). This is illustrated in Figure 3.12, where 14 FPMs are staggered in two rows in the along track direction. The spectral bands are placed in the along track direction resulting in band parallax.
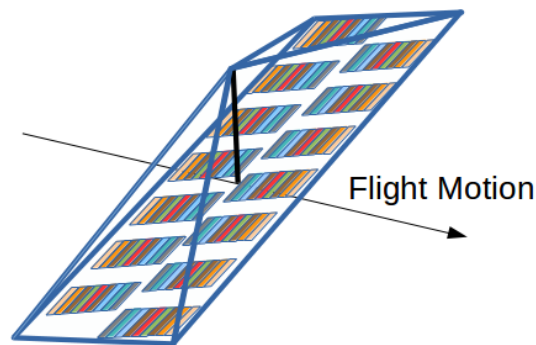
Figure 3.12: The arrangement of 14 FPMs in a focal plane of a pushbroom sensor. The FPMs are staggered in the along track direction resulting in varying band parallax for spectral bands.

In Landsat 8, there are 14 FPMs for the OLI (Landsat 8) sensor while MSI (Sentinel 2) has 12 FPMs to cover the entire swath width. The FPMs are staggered in the along track direction to reduce any cross-talk between the electronics and for easier read out and electronic integration. This imposes further variations in the view angle between the odd and even FPM. As a result, detectors in each spectral band have varying view angle in the across track direction and also have varying view angle in the along track direction depending on whether the detectors are from odd or even FPM.

3.4.1 Landsat 8 Operational Land Imager (OLI)

The Landsat mission provides a continuous collection and monitoring of the Earth's land surface for various scientific application in agriculture, land mapping, forest studies, land cover change detection, and more recently in observing change detection in fresh and coastal waters. Landsat 8 is the most recently launched and operational satellite in a series of Landsat missions dating back to 1972 when Landsat 1 (Earth Resources Technology Satellite 1) was launched. Landsat 8 has two sensors on-board for monitoring in the visible, near-infrared, short-wave infrared (VIS-NIR-SWIR) spectrum, and the thermal infrared spectrum of the EM radiation. The VIS-NIR-SWIR sensor is called the Operational Land Imager (OLI) and the thermal sensor is called the Thermal Infrared Sensor (TIRS). In this research, since VIS-NIR-SWIR spectrum alone is considered for canopy interactions, the thermal sensor TIRS is not discussed.

The OLI sensor collects a 190 KM wide image swath from a 705 KM orbital altitude. The telescope in OLI uses four anastigmat mirrors with an effective focal length of 886 mm at the center. The OLI instrument collects the image using fourteen individual overlapping focal plane modules. Each of these module covers a portion of the 15 degree cross-track field of view and about two degree along track FOV. The FPMs are aligned in a staggered line within the focal plane array so that the adjacent FPMs overlap each other, avoiding any gaps in the cross-track coverage. The alternate FPMs are rotated by 180 degrees to keep the active detector areas as close together as possible. The prelaunch geometric model was constructed for each detector's location in the FPM along with the nominal locations of FPM, and then projecting these locations through the OLI optics into object space based on the OLI telescope design parameters (Storey et al., 2014). A plot showing the line of sight angles for end points and center of each band on each FPM is shown in Figure 3.13. A slight rotation of the FPMs away from the center was introduced intentionally to compensate for optical distortions. These line of sight angles were corrected post-launch using ground control points on the ground, reference images, and using band to band comparisons.

Each FPM includes rows of detectors with interference filters mounted on the module to provide the spectral separation (Knight and Kvaran, 2014). OLI has nine spectral bands which are summarized in Table 3.1. Improvements in SNR for OLI in comparison to Landsat 7 ETM+ is shown in Table 3.2. The coastal/aerosol and cirrus bands are new for OLI. Except for the

panchromatic band, all spectral bands have a nominal ground sample distance of 30m for the products.
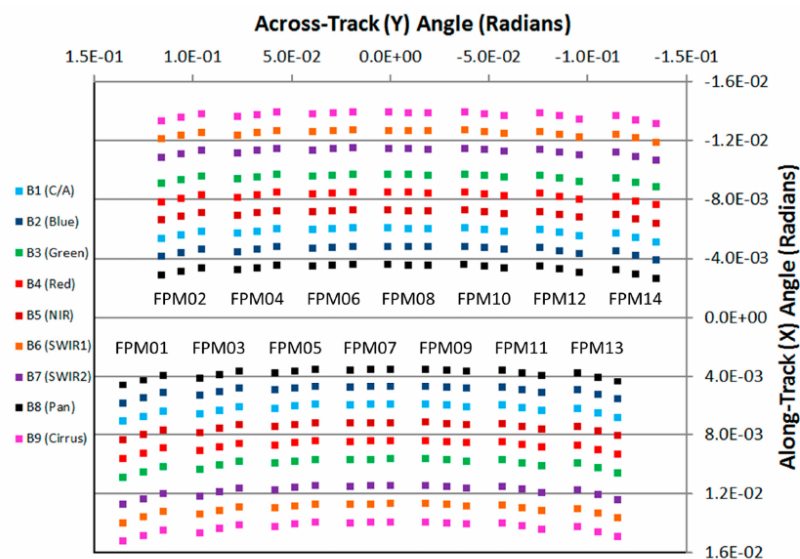
Figure 3.13: OLI line of sight pattern with the along-track scale exaggerated by a factor of ~ 8 to highlight the off-nadir FPM rotation to compensate for optical distortion (Storey et al., 2014).

Table 3.1: OLI spectral bands with spectral and spatial requirements. Spectral (center wavelength, bandwidth) and Spatial (GSD and SNR) (Knight and Kvaran, 2014).

Band Name	Center Wavelength (nm)	Bandwidth (nm)	GSD (m)	SNR Requirement
Coastal/Aerosol	443	20	30	130
Blue	482	65	30	130
Green	562	75	30	100
Red	655	50	30	90
NIR	865	40	30	90
SWIR 1	1610	100	30	100
SWIR 2	2200	200	30	100
Pan	590	180	15	80
Cirrus	1375	30	30	50

Each module consists of nine rows of detectors for spectral band separation and a tenth masked row for radiometric calibration. Figure 3.14 shows the OLI focal plane array and FPM. The detectors from silicon bands (1-5,8) and Mercury-Cadmium-Telluride bands (6,7,9) provides spectral sensitivity in the VIS-NIR and SWIR regions respectively. There are redundant detectors for each bands which improves on-orbit reliability and the physical offset it introduces

Table 3.2: Comparison of OLI SNR to ETM+ SNR at typical radiance levels (Morfitt et al., 2015).

OLI Band #	Band Name	L-typical $Wm^{-2}sr^{-1}\mu m^{-1}$	OLI SNR	ETM+ SNR
1	Coastal/Aerosol	40	237	-
2	Blue	40	367	39
3	Green	30	304	37
4	Red	22	227	26
5	NIR	14	201	34
6	SWIR 1	4	267	36
7	SWIR 2	1.7	327	27
8	Pan	23	148	16
9	Cirrus	6	160	-

are corrected during calibration and geolocation process. The spectral bands image a given ground target at slightly different time (collected by flight motion) and also the same bands in alternate FPM look at the ground at different view angles, resulting in band parallax. Typically, the band parallax is so small that the variation in the sensor observed radiance due to atmospheric path length differences is negligible. However, the band parallax can introduce appreciable variation in the observed radiance for non-Lambertian objects.
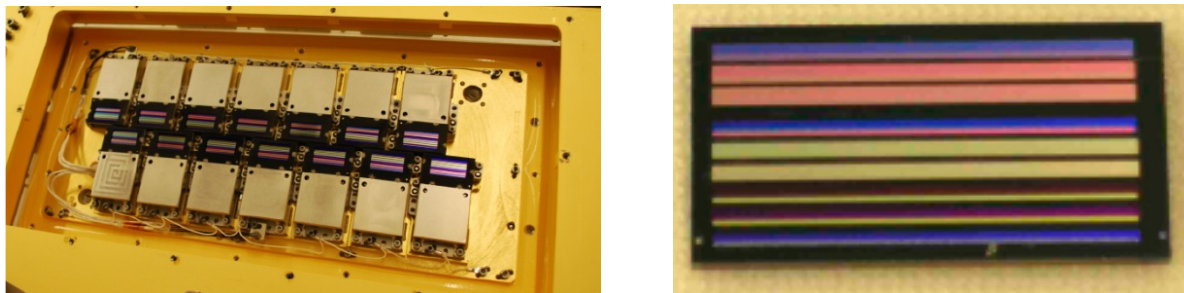
Figure 3.14: OLI full focal plane array (left), individual spectral filter array (right) (Knight and Kvaran, 2014).

The OLI sensor has variable integration times from 90 to 3600 microseconds, but for operational imaging of the earth, they are fixed to 3600 microseconds to account for the instrument velocity and required GSD. The focal plane electronics digitize the signal to 14 bits, but due to data rate limitations, only 12 bits are sent to the ground (Knight and Kvaran, 2014). The upper or lower 12 bits are sent depending on whether the sensor images the ground or imaging calibration of dark targets respectively. The RSR for different spectral bands were characterized prior to launch and tests were conducted to validate the spectral requirements. The test procedure and

the results can be found in Barsi et al. (2014). Figure 3.15 shows the measured RSR prior to launch and the summary of the band responses are shown in Table 3.3
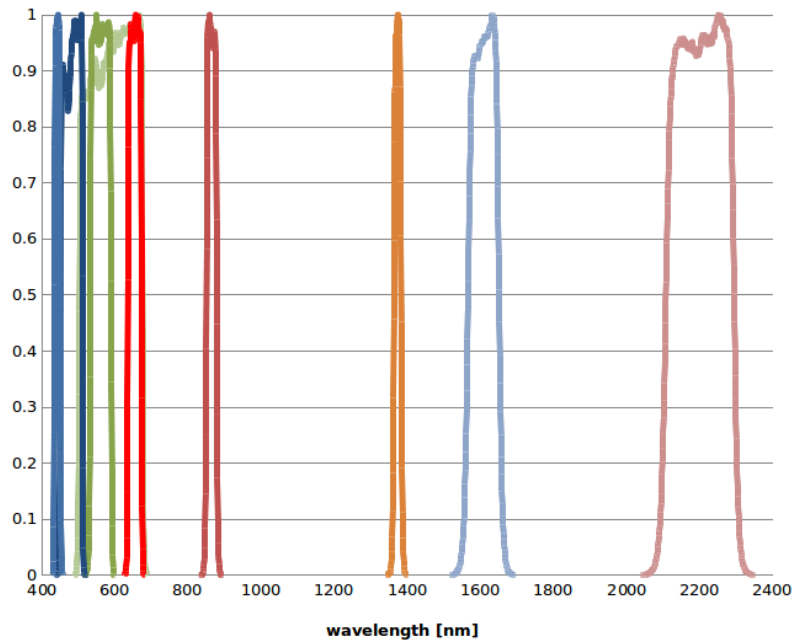
Figure 3.15: The band-average relative spectral radiance responses of the Operational Land Imager (Barsi et al., 2014).

Table 3.3: OLI Spectral band's bandwidth, Full-width Half Max wavelength and Center wavelength (Barsi et al., 2014).

Band Name	Bandwidth [nm]	FWHM wavelength [nm] lower	upper	Center Wavelength [nm]
Coastal/Aerosol	15.98	434.97	450.95	442.96
Blue	60.04	452.02	512.06	482.04
Green	57.33	532.74	590.07	561.41
Red	37.47	635.85	673.32	654.59
NIR	28.25	850.54	878.79	864.67
SWIR 1	84.72	1566.50	1651.22	1608.86
SWIR 2	186.66	2107.40	2294.06	2200.73
Pan	172.40	503.30	675.70	589.50
Cirrus	20.39	1363.24	1383.63	1373.43

The OLI sensor has gone through rigorous calibration prior to launch and also after launch. The calibration and characterization of the OLI sensor showed that the sensor continues to perform well and has exceeded the required specifications in spectral, spatial and geometric characteristics of the sensor (Morfitt et al., 2015, Storey et al., 2014).

3.4.2 Sentinel -2 MultiSpectral Instrument (MSI)

The Global Monitoring for Environment and Security (GMES) program is a joint initiative of the European Commission (EC) and the European Space Agency (ESA) to establish a European capacity for the provision and use of monitoring information for environmental and security applications (Drusch et al., 2012). As part of this initiative, the Sentinel 2 mission will ensure continuity of services that rely on multispectral high resolution optical observations over global terrestrial surfaces. The main objectives of the Sentinel 2 mission is to provide high resolution multispectral imagery with a high revisit frequency, to support and complement the continuity of SPOT and Landsat missions, and to support the next generation of operational products such as land cover maps, land use change detection and geophysical parameters.

The need for a high revisit period imposed a requirement to have two identical Sentinel 2 satellites that would operate simultaneously in the same Sun-synchronous orbit at an altitude of 786 KM (14 +3/10 revolutions per day) with 10:30 as the mean local time of the descending node (LTDN). The two satellites will be phased at 180 degrees on opposite sides of the orbit as shown in Figure 3.16. The LTDN of 10:30 was selected as the best compromise between the need for minimal cloud cover and to ensure suitable solar illumination. Further, it is also close to the LTDN for Landsat and SPOT (10:00), allowing the possibility to combine the historical data sets from Landsat and SPOT with Sentinel 2 for long term time series applications.
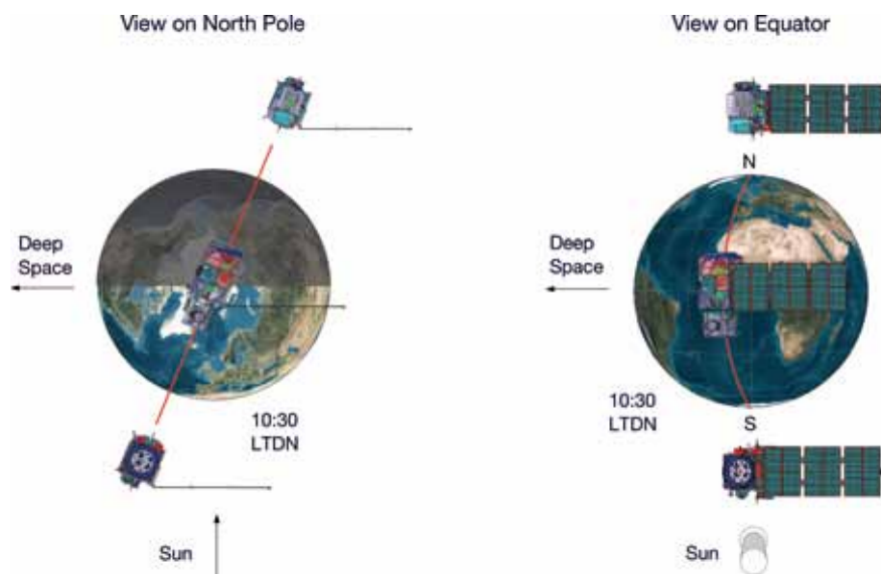
Figure 3.16: Sentinel 2 Satellite orbital configuration showing two sentinel satellites with a phase of 180 degrees to increase the revisit period to 5 days (Drusch et al., 2012).

These requirements have driven the Sentinel 2 design to use a MultiSpectral Instrument (MSI) with 13 spectral bands ranging from the visible and near infrared to the short-wave infrared spectrum. The nominal spatial resolution varies from 10 m to 60 m, depending on the spectral band with a 290 KM field of view. The four bands that are at 10 m nominal resolution include, classical blue (490 nm), green (560 nm), red (665 nm) and near infrared (842 nm). The six bands that are at 20 m nominal resolution includes four narrow bands in the vegetation red-edge spectrum (705 nm, 740 nm, 783 nm, 865 nm) and two large SWIR bands (1610 nm and 2190 nm). The remaining three bands at 60 m nominal resolution are mainly dedicated to atmospheric compensation and cloud screening (443 nm for aerosol retrieval, 945 nm for water vapor retrieval and 1375 nm for cirrus cloud band detection). The spectral band coverage and corresponding ground sample distance are shown in Table 3.4 and in Figure 3.17.

The MSI sensor is a pushbroom sensor with a three mirror anastigmatic telescope and a pupil diameter of 150 mm. There are 12 FPMs that are staggered similar to OLI to image the ground swath of 290 KM. This translates to about 20.6 degrees in instrument FOV (+/- 10.3 degrees) and approximately 3.5 degrees in the Along track direction. The cross track angles are higher in MSI than OLI which is about 15 degree FOV. Since there are fewer bands in OLI, the along track FOV is also smaller in OLI than MSI. The focal length is approximately 600 mm. The VNIR focal plane is based on monolithic Complementary Metal Oxide Semiconductor (CMOS) detectors while the SWIR focal plane is based on Mercury Cadmium Telluride detectors hybridized on CMOS readout circuit.
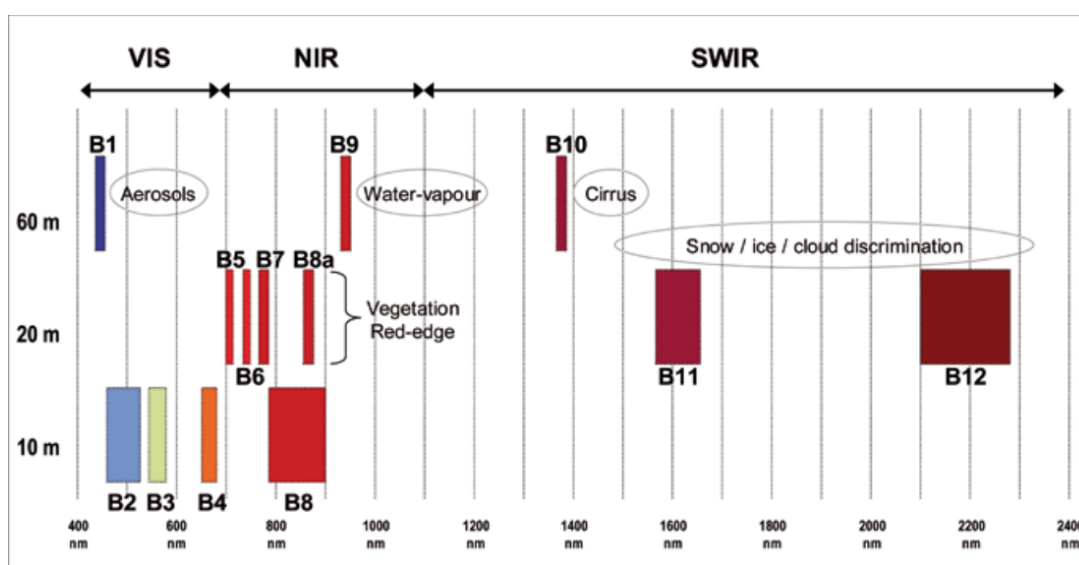
Figure 3.17: Nominal resolution for different spectral bands (Drusch et al., 2012).

Table 3.4: Table showing different bands in the MSI sensor with their corresponding center wavelength, bandwidth, nominal resolution, and Ltyp radiance and SNR (Drusch et al., 2012).

Band #	Central Wavelength [nm]	Bandwidth [nm]	Spatial Resolution [m]	L_{ref} $[Wm^{-2}sr^{-1}\mu m^{-1}]$	SNR @ L_{ref}
1	443	20	60	129	129
2	490	65	10	128	154
3	560	35	10	128	168
4	665	30	10	108	142
5	705	15	20	74.5	117
6	740	15	20	68	89
7	783	20	20	67	105
8	842	115	10	103	174
8a	865	20	20	52.5	72
9	945	20	60	9	114
10	1380	30	60	6	50
11	1610	90	20	4	100
12	2190	180	20	1.5	100

For the SWIR bands, the FPM is made of three rows of detectors for Band 10 and four rows of detectors for Band 11 and Band 12. These three bands work in the Time Delay Integration (TDI) mode by optimal selection of required pixels for TDI in bands 10 , 11 and 12. A dichroic beam splitter provides spectral separation of the VNIR and SWIR focal plane. As a result of beam splitting, there is a possibility in reducing the incoming signal on to these detectors, which affects the SNR performance. Another difference between OLI and Sentinel 2 is that the pixel size of Sentinel 2 (nominal resolution of 20 m from 785 KM altitude) is smaller than that of OLI, thus further reducing the incoming signal onto the detector. Based on the pre-launch characterization, the SNR of the OLI sensor is better by a factor of 2 to 4, for most of the bands when compared at the same radiance levels (see Table 3.2 and Table 3.4). However, if convolved to common resolution, the SNRs of OLI and MSI sensors are similar for most of the bands. The MSI sensor collects the data in a 12 bit quantization and these observed data are downlinked to the ground station by using a lossy compression based on wavelet transform to reduce the data rate and data volume. However, the compression ratio is fine tuned for the spectral bands independently to ensure that there is no observable impact on the image quality due to compression (Drusch et al., 2012).

The RSR for different spectral bands were characterized prior to launch and tests were conducted

to validate the spectral requirements. Figure 3.18 and 3.19 shows the measured RSR prior to launch. MSI sensor has completed pre-launch calibration for spatial, radiometric and geometric characteristics of the sensor and have met all the requirements. Comparison between Landsat 8 and Sentinel 2 sensors indicated that the two sensors are comparable both in geometric and radiometric performance (Storey and Haque, 2016).
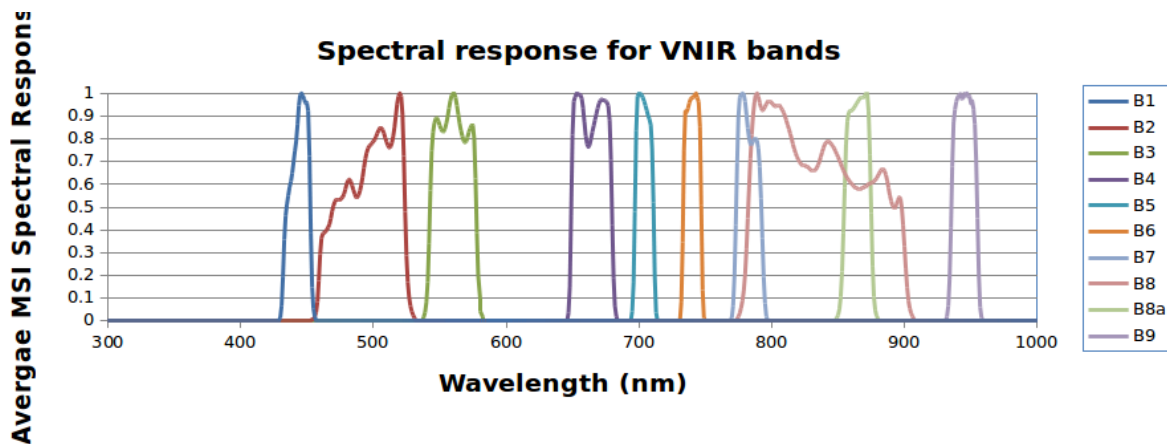
Figure 3.18: Measured RSR (prior to launch) of Sentinel 2 for VIS-NIR spectral bands (Drusch et al., 2012).
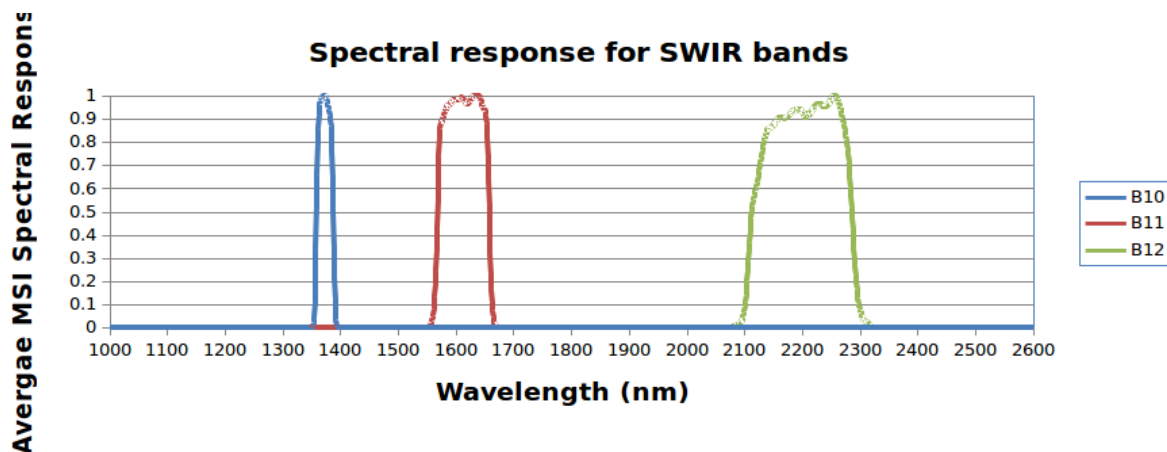
Figure 3.19: Measured RSR (prior to launch) of Sentinel 2 for SWIR spectral bands (Drusch et al., 2012).

3.5 Design of Experiments (DOE)

Experimentation is an important part of any scientific research. An experiment in its formal definition can be defined as a test or series of tests in which useful changes are made to input variables of the process or system to observe and identify the reasons for the changes

in the output response variables (Montgomery, 2012). It is important to have well-designed experiments as the results and conclusions drawn from those experiments are highly dependent on the manner in which the data are collected. The approach to planning and conducting the experiment is dependent on the application that an experimenter is interested. The experimentation is performed while seeking to maximize the efficiency and be able to analyze statistically to derive valid and objective results and also to infer meaningful conclusions. Thus, there are two aspects to any experimental problem: the design of the experiments and the statistical analysis of the data. These two are closely related since the method of analysis depends on the type of experiment conducted.

Montgomery (2012) has outlined seven steps as guidelines for design of experiments and analysis.

1. Recognition of and statement of the problem.

2. Selection of the response variable.

3. Choice of factors, levels and range.

4. Choice of experimental design.

5. Performing the experiment.

6. Statistical analysis of the data.

7. Conclusions and recommendations.

The statement of the problem is the foremost step in the experimental design process. This drives the requirements for the response variable, type of design, number of levels, type of statistical analysis, etc. The selection of the response variable is very much dependent on the type of problem that is needed to be solved. The choice of factors is dependent on the problem as an experienced experimenter would know which factors are to be considered as design factors or nuisance factors. The nuisance factors may have large effects yet they may not be of interest for the particular problem. These nuisance factors could be controllable or uncontrollable. The controllable nuisance factors are compensated as part of appropriate design technique and uncontrollable nuisance factor can be accounted by statistical analysis. Once the design factors are selected, then an appropriate number of levels needs to be decided for each of the factors, which influences the type of design. Sometimes the experiments are

conducted to screen for important factors among many factors. In such cases, the number of factor levels are generally low (about 2 levels). The choice of experimental design is influenced by the number of factors, levels, and the type of problem being solved. While performing the experimental design, it is important to consider the three basic principles of experimental design: randomization, replication and blocking.

Randomization typically means that both the allocation of the experiment and the order of the experimental runs are randomly determined. Also, statistical methods require that the observations or errors be independently distributed random variables. Replication is the process of independent repeat of each factor combination. In a natural experimental set up, observations will be different from one replicate to another. Blocking is a technique used to reduce or eliminate the variability due to nuisance factors.

In this research, DIRSIG is used as an experimental engine in performing the experiments. The Geodensity radsolvers in DIRSIG uses Monte Carlo based ray tracing approach, which can introduce variability in the observations. However, when a BRDF model is used for describing the scene, DIRSIG simulations produce the exact same results consistently for any number of runs. This affects the requirement for randomization and replication. The modifications to account for randomness using the DIRSIG tool will be discussed later.

Statistical methods are used with the experimental results so that the conclusions are objective and valid. Since many of the questions can be cast into a hypothesis-testing framework, it is necessary to use hypothesis testing and confidence interval estimation procedures for a designed set of experiments.

In a broad sense, there are two types of strategy in performing experiments.

1. One factor-at-a-time approach.

2. Factorial experimentation approach.

The One factor-at-a-time method consists of selecting first a set of levels for all input variables or factors, followed by varying each factor over its levels, while keeping all the other factors at a fixed level. This method of varying one factor at a time and evaluating its response for each factor across their levels is typically done by plotting a series of graphs for each factor. The inference from this method is straightforward, as one can see directly a relationship between

the response and each factor independently. This is illustrated in Figure 3.20. The SNR value increases with increase in time, changes from illumination level A to B, and also with increase in pixel size from 10 microns to 20 microns. This was observed by changing one factor at a time while keeping all the other factor levels constant.
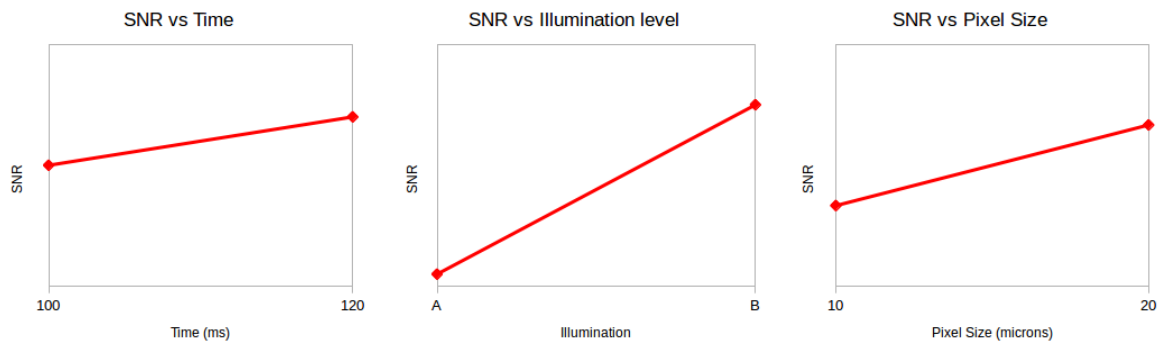
Figure 3.20: Experimental study showing one factor at a time analysis

However, one of the major disadvantages of this strategy is that it fails to consider any interaction between the factors. It is common to expect a particular factor to produce a different effect when the levels of some other factor changed from one level to another. In remote sensing, these interaction effects may be common. For example, a sensor system can show an increasing SNR as a function of increasing integration time for a particular illumination level. However, when the same system is used for measuring SNR at a different illumination level, then the SNR increase may not be appreciable enough at a different integration time as seen in Figure 3.21.

In this case, the SNR increase was almost negligible at the illumination level B for increasing integration time, while the increase in SNR was very high at illumination level A. This indicates that increasing integration time is very useful only when operating at illumination level A. These kind of interaction effects cannot be studied using a one factor-at-a-time strategy.

A better strategy of experimentation would be to use factorial experiments where the experiment is conducted by varying the levels of many or all the factors and analyzed for main and interaction effects.

The design techniques for factorial experiments are useful to study multiple factor interactions efficiently. The analysis may use the Analysis of Variance (ANOVA) model to assess the effects of the factors and their relative significance. The general data model and the hypothesis testing

procedure used in the ANOVA model can be described using a simple design: a design with a single factor with many levels.
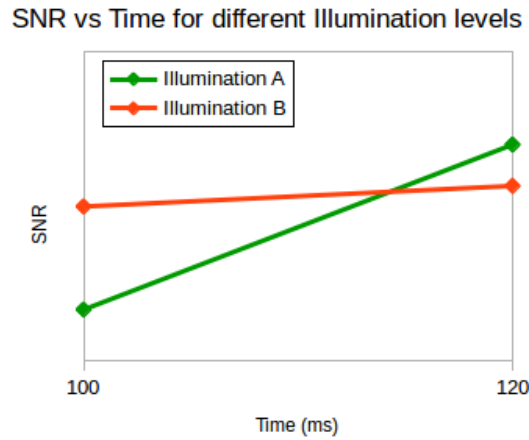
Figure 3.21: Experimental study showing interaction effects

Single factor effects using ANOVA model

For the case of one factor with 'a' levels (treatments), the observed response from each of the levels is a random variable. If there are n observations for each levels, the model for the data can be expressed as shown in Equation 3.40.

$$y_{ij} = \mu_i + \epsilon_{ij} \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., n \end{cases} \tag{3.40}$$

The equation can be rewritten as,

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., n \end{cases} \tag{3.41}$$

where,

y_{ij} is the ij th observation

μ_i is the mean of the i th level

μ is the overall mean for all observations

τ_i is a i th treatment effect

ϵ_{ij} is the random error that incorporates other sources of variation including the nuisance factors.

Equation 3.40 is called the "means model" and Equation 3.41 is called the "effects model" and they both are linear functions of the model parameters. The effects model is widely used and intuitively understood, as the parameter τ_i represents the deviations from the overall mean when a specific type of treatment is used. The main objective of mean model is to test the hypothesis about the treatment means and to estimate them to determine their contribution.

The null and alternate hypothesis for the mean model and the effects model are given in Equation 3.42. They represent the same model, but their hypothesis is based on different parameter.

$$
\begin{aligned}
\text{Mean Model Hypothesis} \quad & H_0 \ : \ \mu_1 \ = \ \mu_2 \ = \ ... \ = \ \mu_a \\
& H_A \ : \ \mu_i \ \neq \ \mu_j \text{ for at least one } \quad i \ \neq \ j
\end{aligned}
$$

$$(3.42)$$

$$
\begin{aligned}
\text{Effects Model Hypothesis} \quad & H_0 \ : \ \tau_1 \ = \ \tau_2 \ = \ ... \ = \ \tau_a \ = \ 0 \\
& H_A \ : \text{ at least one } \quad \tau_i \ \neq \ 0
\end{aligned}
$$

While the null hypothesis in Equation 3.42 indicates that the testing is performed for the equality of the treatment means for the means model, the null hypothesis for the effects model indicates that the treatment effects are zero. The rejection of null hypothesis indicates that at least one of the treatment effect is non-zero, i.e, at least one of the treatments introduce significant effect in the response.

ANOVA is a technique used to test this hypothesis. It is a method to assess the equality of the means based on the comparison of the variation between the sample to within the sample. The total variance can be partitioned into two terms; between variance and within variance. These variances are represented in Equation 3.43.

The $SS_{treatment}$ and SS_E represents the between variance and within variance respectively. Equation 3.43 is referred as the ANOVA identity. The test of hypothesis for no difference in treatment means can be performed by comparing $MS_{treatment}$ and MS_E. The hypothesis testing requires that the model errors are assumed to be normally and independently distributed with zero mean and σ^2 variance, and the variance is assumed to be constant for all levels. If the null hypothesis is true, the ratio of $MS_{treatment}$ and MS_E is distributed as an F distribution with $a-1$ and $an-a$ degrees of freedom. Equation 3.44 is the test statistic for the hypothesis shown in Equation 3.42.

$$
\begin{aligned}
SS_T &= SS_{treatment} + SS_E \\
SS_T &= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 \\
SS_{treatment} &= n\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 \\
SS_E &= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2 \\
MS_{treatment} &= \frac{SS_{treatment}}{a-1} \\
MS_E &= \frac{SS_E}{an - a}
\end{aligned}
\tag{3.43}
$$

where,

SS_T is the total corrected sum of squares

$SS_{treatment}$ is the sum of squares of difference between treatment average and total average

SS_E is the sum of squares of difference between observations from the treatment averages

$$
\begin{aligned}
F_{stat} &= \frac{MS_{treatment}}{MS_E} \\
\text{Reject } H_0 \text{ ; if } & F_{stat} > F_{1-\alpha,a-1,an-a}
\end{aligned}
\tag{3.44}
$$

where, α is the significance level

Typically, for an ANOVA model, the p-value is used to reject the hypothesis. The p-value of a test is the probability that the test statistics will take on a value as extreme as the observed value when the null hypothesis is true. This is shown in Equation 3.45. The ANOVA table for a single factor and three factors are shown in Appendix B.

$$
\begin{aligned}
\text{p-value} &= P(F \geq F_{stat}|H_0 \text{ is true}) \\
\text{Reject } H_0 \text{ ; if p-value} &\leq \alpha
\end{aligned}
\tag{3.45}
$$

The ANOVA analysis typically requires that the experimental runs are performed in random order to ensure that the experimental design is a completely randomized design. A plot of residuals against another variable is used as one of the diagnostic tools to verify the consistency of the data and validate the assumptions on the model. The residuals plot (with fitted values

or factor levels) are very useful to verify the equal variance assumption. Figure 3.22 shows an example for the non-constant and constant variance cases.

(a)

(b)
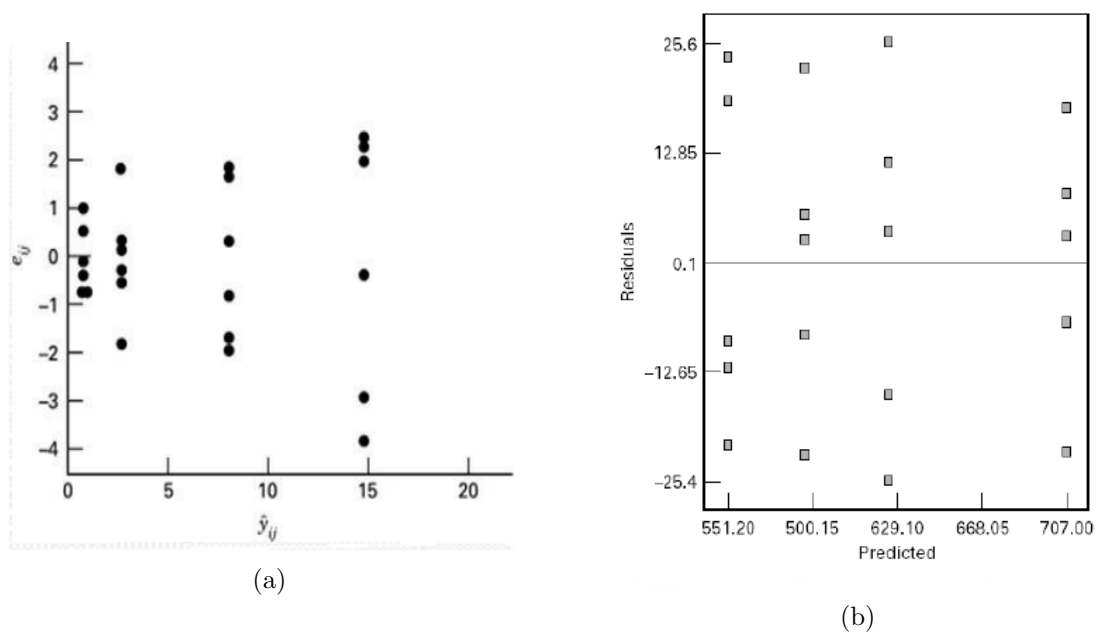
Figure 3.22: Residual plots used to verify the model. The plot on the left shows that the residuals increase with large estimates and these type of pattern indicates that the model is not adequate. The right plot does not show any apparent pattern and the residuals are approximately the same. (Montgomery, 2012).

Figure 3.22a shows an obvious pattern, where the residuals increase with large estimates. This indicates that the model fails the constant variance assumption. There are no apparent patterns observed in Figure 3.22b, so the equal variance assumption is valid. The residual plots can be useful to detect outliers by observing unusually large or small residuals.

The independence assumption is verified by checking the plot of residuals with the order in which the experiment was run. The residuals should be randomly scattered about zero with no distinct patterns to justify the independence assumption. This is illustrated in Figure 3.23. The obvious pattern of decreasing residuals for increasing observation order indicates that nuisance factors are affecting the response. In such cases, either the cause is found and eliminated if possible or the model may need to be adjusted for such factors. In general, this plot is checked first as the other plots could be misleading if this assumption does not hold true.

The normality assumption is verified using a normal probability plot, which is a plot of residuals against their normal scores. The plot will resemble a straight line if the error distribution is normal. In general, moderate departures from normality are of little concern in the fixed analysis

of variance (Montgomery, 2012). Figure 3.24 shows the normal probability plot. Since all the residuals lie along the straight line, the normality assumption is valid for this model.
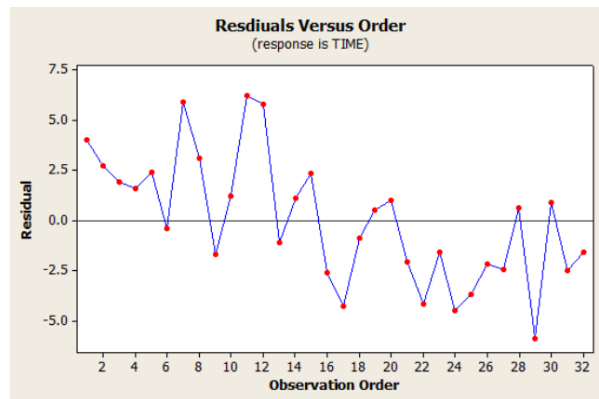
Figure 3.23: The plot of observation order vs residuals. The response variable has a unit of time. As the observation order increases, the residuals decreases. The obvious pattern in this plot indicates that some kind of noise from nuisance factor is affecting the response (Montgomery, 2012).
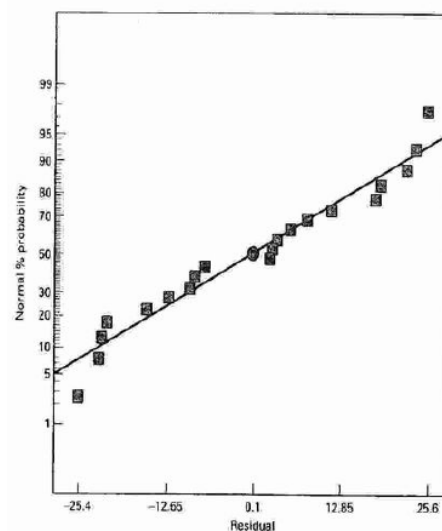
Figure 3.24: The plot of normal scores vs. residuals. The normality assumption is valid for this model as the residuals lie along the straight line (Montgomery, 2012).

The model validation techniques for factorial experiment are similar to the single factor case, and the procedure for analyzing the factorial experiment is explained in the next section.

3.5.1 Factorial experiment

In a factorial experiment, the experiment is performed varying all the factors at the same time to provide information about the process variability. There are two possible effects that can be observed from these experiments: main effects and interaction effects. The main effect is defined to be the change in response produced by a change in the level of the factor when averaged over the levels of all of the other factors. In experiments where there are more than one factor, interaction effects are possible. The interaction effect indicates that the difference in response of one factor varies based on the levels of one or more other factors. This was illustrated in Figure 3.21. In such cases where an interaction effect is present, examining main effects separately may not provide useful information about the process.

The ANOVA model shown in the previous section can be extended easily to more than one factor. For a three factor case, the effects model is shown in Equation 3.46 and the ANOVA model for A through K factors are shown in Equation 3.47. The ANOVA table for a three factor, fixed effect case is shown in Appendix B.

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \epsilon_{ijkl} \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., b \\ k = 1, 2, ..., c \\ l = 1, 2, ..., n \end{cases} \quad (3.46)$$

where,

y_{ijkl} is the $ijkl$ th observation

μ is the overall mean for all observations

$\tau_i, \beta_j, \gamma_k$ are the main effects of ith , jth, kth level of Factor A, B, C with 'a', 'b', 'c' levels respectively

$(\tau\beta)_{ij}, (\tau\gamma)_{ik}, (\beta\gamma)_{jk}$ are the effect of the two factor interaction

$(\tau\beta\gamma)_{ijk}$ is the effect of three factor interaction

ϵ_{ijkl} is the random error component

$$SS_T = SS_A + SS_B + ... + SS_{AB} + SS_{AC} + ... +$$
$$SS_{ABC} + ... + SS_{AB..K} + SS_{ERR} \quad (3.47)$$

where,

$SS_A, SS_B, ..., SS_K$ represent the sum of squares for the main effects for factor A through K

$SS_{AB}, SS_{AC}, ..., SS_{AB...K}$ represent the sum of squares for the interaction effects between all combination of factors

SS_{ERR} represent the sum of squares for the error

The typical procedure used for analyzing a factorial design is as follows:

- Run the ANOVA model and check for significance

- Model reduction

- Analyze residuals (graphically)

- Interpretation of the results

The ANOVA table, as shown in Appendix B, is used for generating the ANOVA model results. This is followed by testing the model hypothesis to determine the significant factor effects. Similar to the single factor experiments, the p-value is used for checking the significance. The model reduction is the process of removing the insignificant effects from the model, and add those effects (sum of squares and degree of freedom) with the error term. The significant effects are used to refit the model. In the model reduction process, the hierarchy principle is used for internal consistency. The principle states that if a higher order term is significant, then all the terms of lower order containing factors involved in the higher order term should also be kept in the model. As explained in the previous section for the single factor experiments, the residual plots are used to check for model adequacy and the normality assumption is verified using a normal probability plot. Finally, the results are interpreted according to the objective of the problem. Although the procedure is similar for different factorial experiments, some designs may use different methods for model adequacy validation. A special case of factorial designs, namely, 2^k designs, are typically used for factor screening experiments.

3.5.2 2^k design

The 2^k factorial design is a special case of the general factorial design with k factors, each operating at two levels. The two levels for each factor dictates that the response is

approximately linear over the range of chosen factor levels. The two levels can be either qualitative or quantitative, and are typically written in Yates order with coded levels (+1 and -1). The Yates order for a 2^3 design is shown in Table 3.5. The treatment combination is determined based on the factor that is set to +1 level.

Table 3.5: Design Matrix for 2^3 design based on Yates order

Treatment Combination	A	B	C
(1)	-1	-1	-1
a	1	-1	-1
b	-1	1	-1
ab	1	1	-1
c	-1	-1	1
ac	1	-1	1
bc	-1	1	1
abc	1	1	1

The coded factor units (+1,-1) can make the interpretation easier. Also, the coded factors help to determine the relative size of the factor effects. In this research, the qualitative and quantitative levels for each factor is converted to coded (+1,-1) level for the factor screening experiment.

In the 2^2 design, there are two main effects and one two-factor interaction effect. For a 2^k factorial design, there are k main effects, $\binom{k}{2}$ combinations of two-factor interactions, $\binom{k}{3}$ combinations of three-factor interactions, and it continues ending with one k-factor interaction.

The analysis procedure and techniques for the 2^k design is the same as the factorial experiment procedure discussed earlier. The ANOVA model can be used, however, a different hypothesis test based on a t-distribution could also be employed to test the hypothesis on the size of the effect. The corresponding hypotheses and the t-statistic used to test the hypothesis is shown in Equation 3.48.

$$t_{stat} \;=\; \frac{\hat{A} - A}{se(\hat{A})}$$

$$H_0 \;:\; A \;=\; 0 \quad \text{for any effect } A$$

$$H_A \;:\; A \;\neq\; 0 \tag{3.48}$$

$$\text{Reject } H_0 \text{ ; if } \text{p-value} \;<\; \alpha$$

where,

$se(\hat{A})$ is the standard error of estimate of effect A

The significance test in Equation 3.48 for 2^k design provides the same result as the ANOVA model, but allows us to see the effect size and check for practical significance, unlike the signficance test using F-distribution. This makes the interpretation of the results simple. For these designs, both the magnitude and direction of the factor effects are typically used to determine the importance of the factors.

In some cases, it is not possible to run all the 2^k combinations in a factorial experiment, particularly when the number of factors is large, and in such cases, fractional factorial designs such as 2^{k-p} are useful. In this research, the 2^{k-p} fractional factorial design is used for factor screening experiment.

3.5.3 2^{k-p} **design**

The 2^{k-p} design is a special case of 2^k design with a smaller number of experimental runs than a 2^k design. Table 3.6 shows the number of experimental runs required in a 2^k design for different k.

For large k, the number of experimental runs is very high, but the percentage of main and two factor interaction effects with respect to all the effects decreases. Typically, higher order effects are seldom necessary in a process as they are usually dominated by the lower order effects and so the higher order effects can be aliased with the lower order effects to reduce the number of experimental runs. The 2^{k-p} design is very useful in the early stages of experiments, when many factors are to be investigated and screened for their significance.

In the 2^{k-p} design, there are k factors with two levels each, but the number of experimental runs is dependent on the amount of fractionation, denoted by the parameter 'p'. For example,

Table 3.6: Number of experimental runs required for different factors in a 2^k design. The percentage of the main and two factor interaction effects with respect to all the effects are shown in the last column.

| k | 2^k | Number of Effects | | | $\% \frac{(main\ +\ 2f.i)}{(\ all\ effects\ )}$ |
		Main	2 f.i	higher order	
3	8	3	3	1	85 %
4	16	4	6	5	67 %
5	32	5	10	16	48 %
6	64	6	15	42	33 %
7	128	7	21	79	22 %
...					
15	32768	15	105	32647	2 %

a fractional factorial design with 4 factors for a 2^k design requires 16 experimental runs, while a 2^{k-p} design requires only 8 experimental runs, if $p = 1$. The design matrix for this example is shown in Table 3.7. In this case, the main effect D is aliased with the three-factor interaction term (ABC). Hence, the effect estimated for D is the sum of the effects of D and the effect of ABC. The aliasing of D with ABC causes ambiguity in determining the significance of the effects as the effect of D cannot be separated from the effect of ABC. This type of aliasing is not uncommon as the effect of a three factor interaction is generally assumed to be insignificant.

Table 3.7: Design matrix for 2^{4-1} fractional factorial design. The main effect for D is confounded with a three factor interaction (ABC).

Standard Order	A	B	C	D=ABC
1	-1	-1	-1	-1
2	1	-1	-1	1
3	-1	1	-1	1
4	1	1	-1	-1
5	-1	-1	1	1
6	1	-1	1	-1
7	-1	1	1	-1
8	1	1	1	1

Similar to the aliasing of D with ABC interaction effect, few other effects are also aliased. The defining relation provides a relation among all the main and interaction effects. The defining relation for this case is given as $I = ABCD$. The alias structure helps to identify the weakness in the design by inspecting the aliased effects.

The resolution of the design indicates the amount of aliased effects, and is given by the minimum

Table 3.8: The aliasing structure for 2^{4-1} design.

Effect	Confound
A	BCD
B	ACD
C	ABD
AB	CD
AC	BD
BC	AD
ABC	D

word length in the defining relation excluding I. The resolution for the 2^{4-1} design shown in Table 3.7 is IV since the defining relation is $I = ABCD$, whose word length equals four. It indicates the weakest part of the design, and in this case, it is aliasing of a two-factor interaction with another two-factor interaction. For screening the factors for significance, the experiments are conducted usually with resolution III or IV designs.

In general, the type of design chosen is dependent on the problem that is to be solved. For example, in some processes, it is necessary to determine the best factor levels that provide the maximum or minimum response. In such cases, response surface methodologies are useful. In this research, one of the primary tasks is to determine the significant factors that affect the sensor response. Therefore 2^{k-p} design is used for screening the factors. The other objective is to estimate the relative importance of these factors, which is accomplished using multi-variate regression analysis.

3.5.4 Regression analysis

A single dependent variable that depends on several independent variables (regressors) are typically characterized by a mathematical model called the regression model. In some cases, the functional relationship is well defined between the response and the variables, but in most-cases, it is difficult to establish a functional relationship, and so polynomial models are used as approximating functions. Models that have interaction terms can also be analyzed using the regression model. In this case, the interaction terms are considered as new regressors. Equation 3.49 shows an example of a regression model with 2 independent variables with and without the interaction terms. The interaction term $x_1 x_2$ can be replaced by a new regressor term x_3. The model can be generalized to k regressor variables using multiple linear regression model as

shown in Equation 3.49. The parameter β_i represents the expected change in response y per unit change in x_i, when all the other regressors are held constant. Similarly, a function with higher order polynomials can be easily analyzed using the multiple linear regression models. Note that the model is linear with respect to the parameters.

$$
\begin{aligned}
\text{Linear regression without interaction} : y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon \\
\text{Linear regression with interaction} : y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon \\
&:= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon \\
\text{Multiple Linear regression} : y &= \beta_0 + \beta_1 x_1 + \beta_1 x_2 + ... + \beta_k x_k + \epsilon
\end{aligned}
\tag{3.49}
$$

The coefficients in a multiple linear regression model are estimated using the least squares method. The matrix notation and the estimator of β is shown in Equation 3.50.

$$
\begin{aligned}
y &= X\beta + \epsilon \\
\hat{\beta} &= (X'X)^{-1}X'y \\
\hat{y} &= X\hat{\beta} \\
e &= y - \hat{y} \\
SS_E &= e^T e \\
\hat{\sigma^2} &= \frac{SS_E}{n-p}
\end{aligned}
\tag{3.50}
$$

where,

SS_E is the sum of squares of the residuals

$\hat{\sigma^2}$ is the unbiased estimator of variance

The off-diagonal elements in $X'X$ are the sums of cross products of the columns in X which are equal to zero in the orthogonal designs. The 2^k factorial design is an orthogonal design for fitting the multiple linear regression model, however not all factorial designs are orthogonal. For example, fractional factorial design with factors at multiple levels (unequal intervals) may not be an orthogonal design. Similar to other factorial designs, the error in the regression model is assumed to be normally and independently distributed with mean zero and variance σ^2. As a

result of this assumption, the observations y_i are normally and independently distributed with mean $\beta_0 + \sum_{j=1}^{k} \beta_j x_{ij}$ and variance σ^2 (Montgomery, 2012).

3.6 Summary

This chapter is an overview of the concepts necessary to understand the approach that is discussed in the following chapter. The importance of defining certain radiometric terms in this chapter is essential to avoid any confusion, particularly when dealing with BRF or BRDF of the forest canopy. Using DIRSIG as a virtual Gonionmeter, one can measure a simulated forest's BRDF as will be shown in the next chapter. In this research, due to the limitation of time and system resources, the entire BRDF was not measured in DIRSIG. Instead, the measurements were used to accurately fit to a forest canopy BRDF model. Even though many models are available in the literature, RossLi BRDF model was used in this research owing to its simplicity and effectiveness in modeling forest canopy BRDF. Koukal and Schneider (2010) compared the accuracy of different BRDF models and concluded that RossLi BRDF model for different biomes is as accurate as leading non-linear models such as the RPV model. Using modeled parameters in DIRSIG for BRDF also helps to run the simulations efficiently. In DIRSIG, the sensor and environmental factors can be modeled to simulate the radiance reaching the sensor. In that context, the Landsat 8 and the Sentinel 2 sensor parameters, as discussed in this chapter, are used for modeling. By using established DOE practices, the interaction between factors and their relative significance can be studied to accomplish the stated objectives.

Chapter 4

Methods and Approach

This chapter delves into the details necessary to implement the goals and objectives addressed in Chapter 2. In Section 4.1, the validation strategy used in comparing the DIRSIG to other radiative transfer models are discussed. This is important since all the simulations are performed in DIRSIG. Section 4.2 discusses the methodology involved in modeling the forest trees using OnyxTree software, and also briefly explains the tree placement strategy to create a forest canopy in DIRSIG. Section 4.3 elaborates on the methodology necessary to measure BRDF using DIRSIG, specifically about the sampling strategies for the view angles, sun angles and wavelength. This section also discusses the sensitivity study performed to determine the optimal number of photons and ray tracers for modeling complex structures such as a forest canopy in DIRSIG. This is critical, as at some level increasing the number of photons or rays affect the processing time without any measurable improvement in the results. Section 4.4 describes the approach used to fit the BRDF measurements to the RossLi BRDF model and its associated sensitivity to the measurements. This section also discusses the approach used to generate different BRDF models such as forest on sloped terrain, forest senescence, etc. Section 4.5 discusses the approach used to model and characterize the defoliation of forest for different remote sensing products. The characterization of forest defoliation helps to analyze and estimate the effect of sensor and environmental factors using one-factor-at-a-time strategy as discussed in Section 4.6. The approach used in the validation of simulated data by comparing it with the data measured by the real sensors, is explained in Section 4.7. Section 4.8 describes the methodology used to screen the important factors that affect the sensor's response based on

the factorial design and analysis technique. Lastly, the chapter discusses the method used to determine the relative significance of the factors using regression based analysis.

4.1 DIRSIG Validation

As alluded to in Chapter 3, DIRSIG is used as the experimental engine for performing all the experiments and simulations. This requires that all the algorithms used by DIRSIG are validated accurately. DIRSIG has been extensively used in the remote sensing community and has evolved both in terms of accuracy and performance over the past two decades, and its validations are summarized in Brown and Schott (2010). The algorithms related to sensor and platform geometry have been validated for its accuracy to represent the real world sensors in the simulated environment. Solar and emissive spectral illumination for point and non-point sources have also been validated for its consistent and accurate representation of reality. DIRSIG uses MODTRAN for all its radiative transfer computation for atmospheric constituents. MODTRAN solves the radiative transfer equation to characterize molecular and particular absorption, emission, and scattering, as well as reflections, emissions and transmissions. MODTRAN has been well documented and investigated for its accuracy in representing the radiative transfer for atmosphere from the ground to the height of the sensor (Anderson et al., 1999, Kotchenova et al., 2008).

Many published articles as outlined in Brown and Goodenough (2015) have shown validation of radiometric solvers and other BRDF models such as Phong, Shell, Priest-Germer, Ward, etc. However, most of these validation uses simple geometrical surfaces with single scattering or dual scattering at the most. In forest canopies, most of the energy is reflected due to multiple scattering. In particular, the reflected energy due to multiple scattering can be as high as 50% of the total reflection in the NIR spectral region. In a typical forest canopy, a single photon can bounce as many as 50 times before finally leaving the canopy or being absorbed by the trees or ground. It is extremely difficult and computationally inefficient for generic radiometric solver algorithms to keep track of all photon bounces. The Geodensity radsolver, as explained in Section 3.3.6, reduces the computational time required in tracking the multiple photon bounces. Although this algorithm is similar to the photon mapping algorithm developed and validated by Goodenough et al. (2006), it has not been validated for its accuracy in representing the canopy

interactions. In this research, one of the important tasks is to validate the Geodensity radsolver algorithms.

In general, it is extremely difficult to validate any radiative transfer algorithm due to the limitations with the physical observations. The validation of the Geodensity radsolver in DIRSIG against a published radiative transfer algorithm presents additional challenges. Firstly, the radiative transfer models may not require the complex 3D geometries that are needed in DIRSIG. Also, some models are dependent on parameters such as LAI, leaf orientation, leaf thickness, etc., which cannot be directly used as input parameters in DIRSIG. Secondly, the unique workflow of DIRSIG as explained in Section 3.3, makes it difficult to implement other radiative transfer algorithms efficiently in DIRSIG. Thirdly, the results or output parameters that these models generate may not be consistent with the DIRSIG results. For example, some radiative transfer algorithms provide the BRDF of the canopy directly in a functional form or generate results in reflectance units, but DIRSIG images are in radiance units and needs to be converted to reflectance for appropriate comparisons. Due to these limitations, the validation of DIRSIG is performed using a mix of qualitative and quantitative methods as summarized below.

4.1.1 Scene Reflectance

The scene generated using DIRSIG for a complex 3D geometry such as a forest canopy can be validated using qualitative methods, where, the image radiance or reflectance is compared to the expected results based on intuitive knowledge of the light interaction within the canopy. For example, multiple scattering from leaves can increase the incoming irradiance on to other leaves, resulting in an increased radiance for a specific view and illumination angle. Such non-quantitative techniques can be used to verify DIRSIG for its true reproduction of light interaction within the canopy.

4.1.1.1 Multiple Scattering Evaluation

In DIRSIG, multiple scattering can be validated in two ways; verifying the response on shadow and non-shadow regions.

Evaluation based on shadow is a simple qualitative method to verify that the results produced by DIRSIG are closer to what is expected based on first principle approximations. In this method, an image of a tree is simulated in DIRSIG with given illumination zenith and azimuth angles and compared subjectively against the expected result over the shadow region.

In the real world, a tree viewed at nadir from a non-nadir illumination will produce its shadow on the ground. Figure 4.1 shows an example of a tree with shadow cast over the ground. This can be simulated in DIRSIG using a nadir viewing sensor with illumination source (sun) at a specific illumination angle. DIRSIG is primarily intended for natural illumination condition with a MODTRAN defined atmosphere. In order to validate the multiple scattering phenomenon for leaves, it is necessary to remove any scattering due to the diffused skylight. This can be accomplished in DIRSIG by simulating with a dark (black) sky with no scattering from the atmosphere. The outgoing radiance from the shadow pixels in the simulated image is due to the incoming irradiance from many scattered (reflected and transmitted) photons from the surrounding regions (trees and ground). The radiance expected from a shadow pixel is dependent on the ground reflectance, ground terrain, LAI, wavelength, reflectance and transmittance properties of the tree species, and the radsolver parameters. High LAI indicates more leaf elements in the virtual scene, which reduces the transmitted energy to the shadowed region, resulting in dark shadows. Also, low reflectance and transmittance for tree elements (and ground) can reduce the measured radiance over shadow regions. For example, tree reflectance in the visible region of the spectrum is much lower than the NIR spectral region. Further, reducing the number of photon bundles (light rays) illuminating the scene or the number of rays captured by a detector reduces the observed radiance.

As discussed in Section 3.3, the Geodensity radsolver associates the photon bundles to primitive geometries such as facets. As a result, their flux density estimates are valid only when geometry elements are fairly small and uniform as seen in a typical leaf geometry. This constraint limits the use of Geodensity radsolver for larger geometries like ground surfaces, which use the generic radsolver for their radiometric computations. Taking these limitations into consideration, a simulation can be performed with trees illuminated by more than 1 million photon bundles, detector's ray sampled at very fine resolution in the NIR spectral region, where the reflectance and transmittance are approximately equal with very low absorption.

The scene with a single tree will casts its shadow on the ground, but the ground uses a different
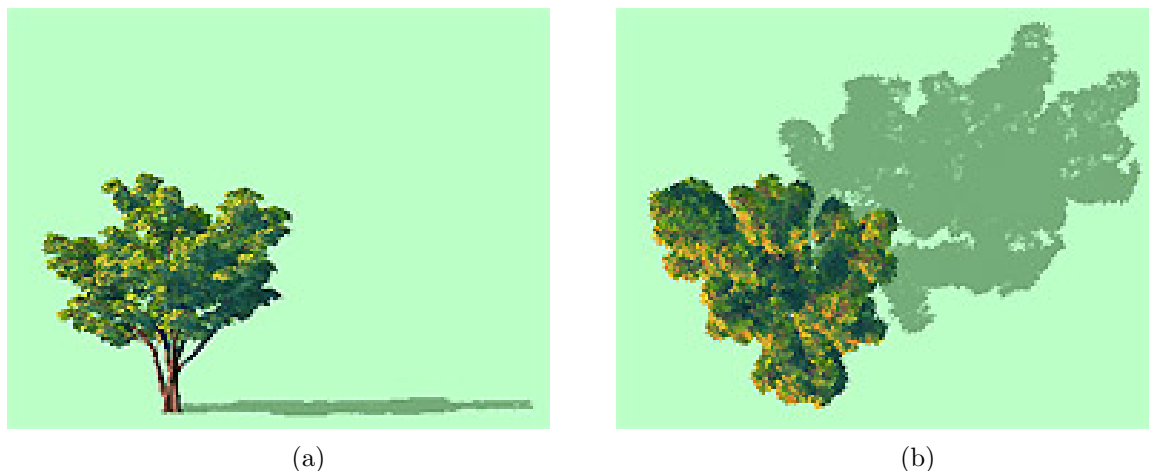
Figure 4.1: Modeled tree showing shadows for non-nadir illumination for (a) front view and (b) top view (Onyx Computing, 2015).

radsolver. Hence two identical trees are used, such that one tree casts its shadow on the other. The shadowed part of the tree will be expected to reflect lower radiance to the sensor than the tree directly illuminated by sun and this is verified in the simulated image. Also, the shadowed part of the tree reflects much higher radiance than a near zero reflector (dark shadow), due to multiple scattering of light from adjacent elements.

The qualitative evaluations for a non-shadow region can be performed using the same strategy and simulated image as described above. The scene consists of two identical trees and one tree casts its shadow on the other. The tree under shadow will have both shadow and non-shadow region of the crown when viewed from nadir. Leaves, having higher transmission and reflection in the NIR, allow more photons to be transmitted and reflected, that in turn can increase the incoming radiation onto the other leaves or trees. This will result in an increased radiance at the non-shadow part of the tree. This effect is expected in the simulated image and is used to verify the validity of the multiple scattering interactions in DIRSIG. Chapter 5 discusses the qualitative validation results for multiple scattering in DIRSIG.

4.1.1.2 Total reflectance evaluation

The qualitative analysis approach used to verify the multiple scattering effect was discussed in the previous section. In this section, the method used to verify the total reflectance from the ground and the trees are explored by comparing them to its typical values for different wavelength.

In general, the forest canopy's reflectance is more influenced by the structural variability (leaf position and orientation) than due to the BRDF of the individual elements. Hence, the facets of leaves, trunks and branches can be modeled as Lambertian reflectors. Further, the closed canopy in a typical deciduous forest reduces the interaction of direct illumination with the ground. Thus, the ground can also be assumed as a Lambertian reflector. The transmittance of the leaves are very high in the NIR region, and therefore, the leaf facets are modeled assuming as a Lambertian transmitter. These BRDF assumptions simplify the construction and the radiometric computations within the DIRSIG model. A virtual forest scene, thus constructed, can be used to estimate the total reflectance of the canopy. Since DIRSIG provides its simulated image in radiance units, they need to be converted to reflectance.

The radiance image from DIRSIG can be converted to reflectance either using simulated panels based on the Empirical Line Method (ELM) (Schott, 2007, p. 280) or by varying the irradiance of the source. The latter technique is used in this study to get the measurements in reflectance units for modeling BRDF. In this research, the ELM method is also used as a simple technique for compensating atmospheric attenuations for the simulated data. The governing conditions to convert the radiance to reflectance image by varying the irradiance of the source is shown below.

$$
\begin{aligned}
L(\theta_i, \theta_v, \phi_i, \phi_v, \lambda) &= E(\theta_i, \phi_i, \lambda) \ \cos(\theta_i) \ BRDF(\theta_i, \theta_v, \phi_i, \phi_v, \lambda) \\
BRF(\theta_i, \theta_v, \phi_i, \phi_v, \lambda) &= \pi \ BRDF(\theta_i, \theta_v, \phi_i, \phi_v, \lambda)
\end{aligned}
\tag{4.1}
$$

From Equation 4.1,

$$
BRF(\theta_i, \theta_v, \phi_i, \phi_v, \lambda) \ = \ \frac{\pi \ L(\theta_i, \theta_v, \phi_i, \phi_v, \lambda)}{E(\theta_i, \phi_i, \lambda) \ \cos(\theta_i)}
\tag{4.2}
$$

Changing the magnitude of irradiance to a convenient value,

$$
E(\theta_i, \phi_i, \lambda) = \frac{\pi}{\cos(\theta_i)}
\tag{4.3}
$$

then,

$$
\begin{aligned}
BRF(\theta_i, \theta_v, \phi_i, \phi_v, \lambda) &= \frac{\pi \ L(\theta_i, \theta_v, \phi_i, \phi_v, \lambda)}{(\frac{\pi}{\cos(\theta_i)}) \ \cos(\theta_i)} \\
BRF(\theta_i, \theta_v, \phi_i, \phi_v, \lambda) &= L(\theta_i, \theta_v, \phi_i, \phi_v, \lambda)
\end{aligned}
\tag{4.4}
$$

As seen from Equation 4.4, signals captured by DIRSIG are, in essence, in reflectance factor units when appropriate irradiance values are chosen. The images simulated using DIRSIG in the NIR spectral bands can be used to qualitatively verify the reflectance measured over the tree crown and ground.

The evaluation of radsolver algorithms using the total reflectance and multiple scattering methods are useful to preclude any gross errors with the radsolver. While methods mentioned in this section are simple, they do not provide any quantitative comparisons with other radsolver algorithms. The next section explores the quantitative verification of DIRSIG by comparing it with published canopy radiative transfer algorithms. This is important because these models were specifically developed to model tree canopies, whereas, DIRSIG is a general scene simulation tool whose utilities for forest canopy modeling has not been rigorously validated.

4.1.2 Validation of DIRSIG with RAMI

RAMI proposes a mechanism to benchmark models designed to simulate the transfer of radiation at or near the Earth's terrestrial surface (RAMI, 2015). RAMI is an on-going activity that operates in successive phases with each one aiming at re-assessing the capability, performance and agreement of the latest generation of Radiation Transfer (RT) models. It is expected that these benchmark tests will lead to model enhancements and developments that will benefit the RT modeling community. Note that so far, the entire focus has been on models designed to model forest canopy just above the canopy.

The first phase of RAMI (RAMI I) was launched in 1999 with its prime objective being to document the variability that existed between canopy reflectance models when run under well controlled experimental conditions (Pinty et al., 2001). This was extended with the launch of second phase (RAMI II) in 2002 to focus on performance of models dealing with structurally complex 3D plant environments. The third phase of RAMI (RAMI III) in 2005 further increased the number of test cases and evaluated the relative performance of 18 different RT models. RAMI III comparisons showed a general convergence of the submitted RT simulation models that agreed to better than 1% between six of the participating 3D Monte Carlo based RT models.

The fourth phase of RAMI (RAMI IV) introduced a completely new set of architectural scenarios subdivided into "abstract" and "actual" canopies. The latter are based on detailed inventories

of existing forest and plantation sites. DIRSIG validation for actual canopies would be very useful, but could not be attempted due to the unavailability of RAMI IV results at the time of validation. Hence, the validation is focused primarily on abstract scenes from RAMI III.

RAMI III test cases can be broadly subdivided in to two groups: Homogeneous experiments and Heterogeneous experiments. In homogeneous experiments, the properties of a given canopy environment is independent of geographic location (spatial coordinates). The environment consists of simple and uniform canopy structures. In homogeneous experiments, scenes with scatterers having optical properties with typical values encountered in the solar domain (red and NIR spectral regions) are referred to as "Solar domain" experiments. Scenes with optical properties that do not correspond to any realistic scenarios but are required to assess the energy conservation are referred to as "purist" experiments. For DIRSIG validation, only the solar domain experiments are validated, as they exhibit the realistic scenarios within the forest canopy.

In heterogeneous experiments, the properties of the environment are dependent on spatial coordinates. While four different heterogeneous environments have been provided, only one of them, "Real Zoom-In" scenario, is used for validation.

4.1.2.1 Homogeneous Experiments

Homogeneous environmental scenes are composed of a large number of non-overlapping disc-shaped objects representing the leaves, located over a horizontal plane standing for the underlying soil surface. These objects were randomly distributed finite size scatters with specific optical properties (transmittance, reflectance), and the orientation of the normals follow either a planophile or erectophile distribution function. The optical properties of the underlying soil follows Lambertian BRDF. The scene geometry and the associated optical properties are found in Table 4.1. The graphical representation of the homogeneous scene is shown in Figure 4.2.

Table 4.1: RAMI III geometry and optical properties for planophile and erectophile distribution. The disc scatterer has Lambertian BRDF for transmittance and reflectance, and soil has Lambertian BRDF for reflectance (RAMI, 2015).

	Case 1 RED-20	Case 2 NIR-20	Case 3 RED-50	Case 4 NIR-50
Solar zenith angle [degrees]	20	20	50	50
Solar azimuth angle [degrees]	0	0	0	0
Scatterer radius (planophile) [m]	0.1	0.1	0.1	0.1
Scatterer radius (erectophile) [m]	0.05	0.05	0.05	0.05
LAI	3	3	3	3
Height of canopy [m]	2	2	2	2
Leaf reflectance	0.0546	0.4957	0.0546	0.4957
Leaf transmittance	0.0149	0.4409	0.0149	0.4409
Soil reflectance	0.127	0.159	0.127	0.159
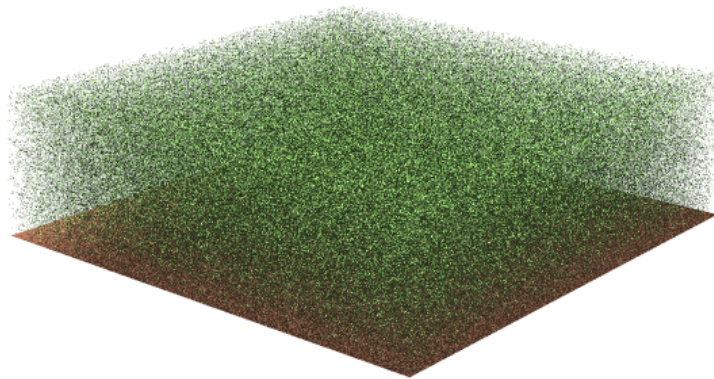
Figure 4.2: Geometrical representation of the scene used for homogeneous experiments (RAMI, 2015).

A file with the exact coordinates of every leaf and its normal is provided, which are converted to DIRSIG scene format. The DIRSIG simulation provides results in BRF units when appropriate irradiance is used as shown in Equation 4.3. RAMI (2015) published BRF results for principal and cross plane directions along with many intermediate results from single scattering and multiple scattering. However, in this research, validations are performed only for BRF along the principal and cross plane for total canopy reflectance since agreement in these two planes validates all the intermediate steps. The principal plane BRF is generated by performing DIRSIG simulations for different view zenith angles along the principal plane. The DIRSIG BRF is then compared against the published RT models to validate its consistency.

4.1.2.2 Heterogeneous Experiments

The "Real Zoom-In" experiments of heterogeneous environmental scenes are carried out at the top of the canopy in the NIR wavelength at different spatial resolutions (270 m, 90 m, 30 m). The scene is composed of a large number of disc-shaped scatterers that are contained within a series of non-overlapping spherical and cylindrical volumes. The spherical and cylindrical volumes represent plant crowns and they are fixed in their sizes and positioned at pre-determined locations above a horizontal plane (soil surface). The disc shaped scatterers are randomly distributed within the spherical and cylindrical volumes and are characterized by their optical properties as well as the orientation of their normals. As in the homogeneous case, the soil is assumed to have Lambertian BRDF for reflectance, while the leaves (disc-scatterers) are assumed to have Lambertian BRDF for reflectance and transmittance. The scene geometry and the associated optical properties are found in Table 4.2. The graphical representation of the heterogeneous scene is shown in Figure 4.3.

Table 4.2: RAMI III geometry and optical properties for "Real Zoom-In" test case of Heterogeneous experiments. The disc scatterer has Lambertian BRDF for transmittance and reflectance, and soil has Lambertian BRDF for reflectance (RAMI, 2015).

	Real Zoom-In
Scene dimensions (LxWxH) [m]	270 x 270 x 15
Scatterer radius [m]	0.05
Scatterer normal's distribution	uniform
Sphere radius [m]	4
Cylinder radius [m]	3
Cylinder height [m]	12
Top of canopy height [m]	15
LAI	5
Number of spheres	205
Number of cylinder	409
Scene coverage	30 %
Solar zenith angle [degrees]	20
Solar azimuth angle [degrees]	0
Leaf reflectance (sphere)	0.49
Leaf reflectance (cylinder)	0.45
Leaf transmittance (sphere)	0.41
Leaf transmittance (cylinder)	0.30
Soil reflectance	0.15
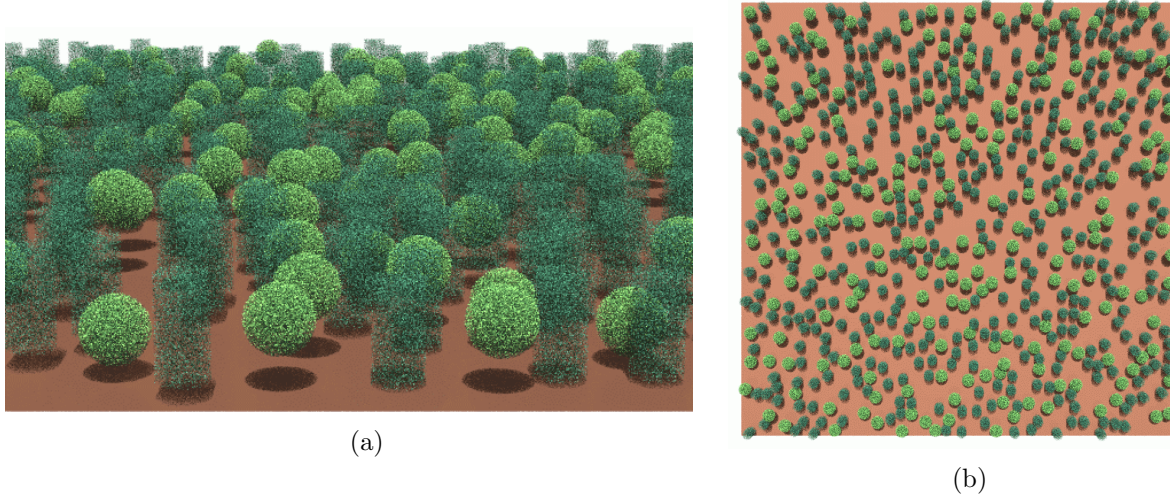
Figure 4.3: RAMI III Heterogeneous scene for "Real Zoom-In" experiment from (a) side view and (b) top view (right) (RAMI, 2015).

The different spatial resolutions are subdivided into 19 test cases which are graphically represented in Figure 4.4.
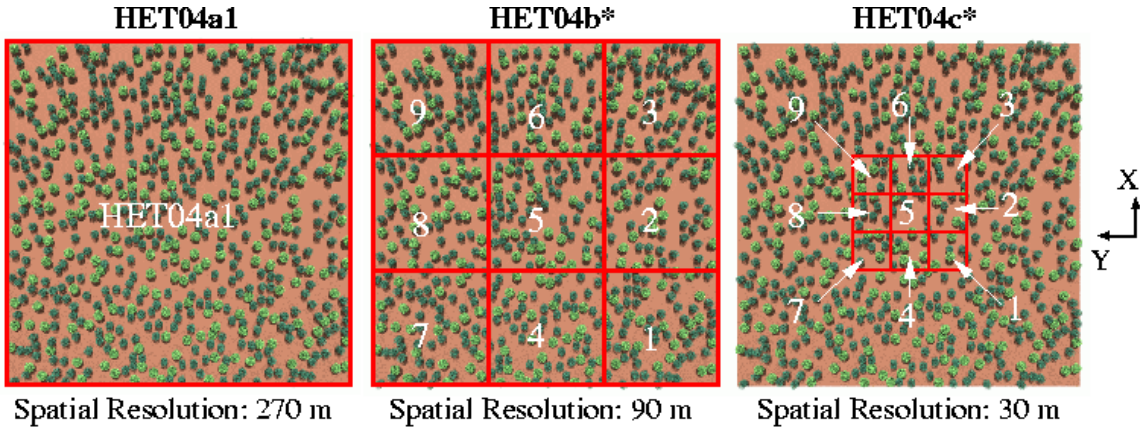
Figure 4.4: Different test cases used to validate DIRSIG. HET04a1 provides BRF for 270 m x 270 m scene extent, while 9 test cases in HET04b[1-9] has a spatial dimension of 90 m x 90 m. HET04c[1-9] covers a 30 m x 30 m scene extent with 9 test cases whose average is expected to be consistent with HET04b5. Similarly, the average BRF of HET04b[1-9] experiments are expected to be consistent with HET04a1 (RAMI, 2015).

In all these measurements, the entire scene is illuminated irrespective of the different field of view measurements. The average of all the HET04c* experiments should be comparable to the values of the HET04b5 test case, and average of all the HET04b* experiments should be consistent to the HET04a1 experiment. The fraction of absorption and transmission through the canopy are computed using the incident radiation that illuminates from the top of the canopy. In DIRSIG, this can be easily simulated by using direct illumination from the source

for the specific spatial extent of the scene. The BRF measurement for varying zenith angle imposes a change in the solid angle for a sensor with a constant field of view. While this can be compensated by appropriate cropping of the output results, the reduction in samples used for representing the scene are inevitable as we move away from nadir. This issue is mitigated by allowing the samples to be generated at the measured area directly, while maintaining the same sample density for all different view angles (Goodenough and Brown, 2015).

As with the homogeneous experiments, a file with the exact coordinates of all the scatterers and their normal's orientation are provided. These are converted into DIRSIG scene format. The BRF results published by RAMI (2015) are then compared with the DIRSIG simulated BRF results along the principal plane direction. The validation results for all the RAMI III experiments are provided in Chapter 5.

4.2 Modeling of Forest Canopy

The successful validation of DIRSIG to produce consistent and accurate results in comparison to other RT models, discussed in Chapter 5, elevates the confidence in using DIRSIG for modeling the real world forest. In the real world, forest canopies consist of similar types of trees with varying height on an undulating surface covered with litter from twigs and leaves. It is extremely difficult if not impossible to recreate an exact tree geometry with their associated optical properties. Hence, appropriate assumptions are needed to model the tree and canopy geometry. This section describes the process involved in modeling a forest canopy with an introduction to the Harvard forest site that is used for modeling the deciduous forest canopy in DIRSIG.

4.2.1 Harvard forest

Research teams from University of Massachusetts, Boston (UMB) and Boston University (BU) are monitoring certain ground sites in Harvard Forest located in Massachusetts. The team maintains and collects forest inventory data over hardwood tree sites annually. These sites are also used for validation of data products from sensors like MODIS (Shabanov et al., 2003). The Hardwood tree site is subdivided into 9 geographical sections (East, West, North, South, North East, North West, South East, South West and Center). Modern surveying techniques are used

to measure the distance and bearing of each tree within its section relative to the center of the section. In this research, the central section of the site is used to model the Hardwood forest (covers approximately 50 m x 50 m). The forest inventory provides information on the tree number, its relative position and orientation, Diameter at Breast Height (DBH), species, status of the tree such as dead leaf, broken trunk, etc., and crown dominance. Figure 4.5 shows an example of four types of crown dominance observed in forest stands. Some of the trees have additional information such as tree height, tree diameter, crown extent, and height to living crown. In this site, there are about 100 trees of which about 40% of the trees are dominant or co-dominant trees and the three most common tree species include birch, red maple and red oak trees. Figure 4.6 shows an approximate location of the central section of the Hardwood tree site.
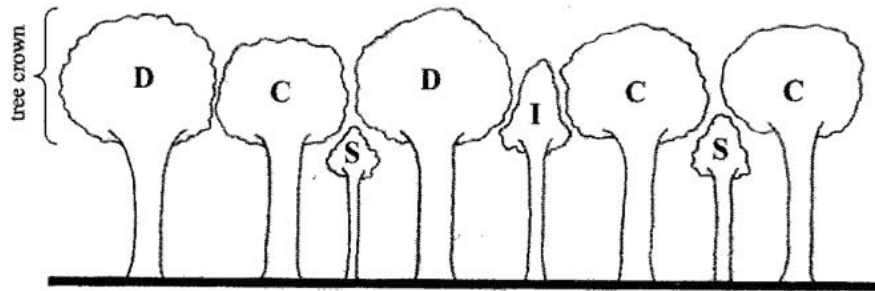
Figure 4.5: Tree crown positions as it relates to dominance in forest stand (D = Dominant, C = Codominant, I = Intermediate, S = Suppressed) (Smith, 1986).

(a)   (b)

Figure 4.6: Approximate location of Hardwood forest site during (a) leaf on condition acquired on Sept, 2010 and (b) leaf off condition acquired on May, 2015 (Google Earth, 2015).

The team also conducts ground campaigns to collect spectral signatures of leaves, trunk and ground litter using field spectroradiometers in the VIS-NIR-SWIR region. The spectral information collected are limited to a few tree species. The species with the missing spectral information are approximated with one of the existing species.

4.2.2 Forest geometry

The forest inventory is quite useful for modeling trees, but, building a virtual tree is a difficult and time consuming process. The LIDAR community has shown promising research in using ground based LIDAR to model high fidelity tree models (Disney et al., 2013, Reitberger et al., 2009, Rosell et al., 2009). However, lack of such a dataset in this case, has lead to modeling trees using commercially available tree building software called OnyxTree (Onyx Computing, 2015). OnyxTree is capable of generating geometries such as broadleaf, conifers, grasses, palms, flowers and bamboo. The OnyxTree environment allows for trees to be built using parameters as shown in Figure 4.7. However, the parameters required by OnyxTree cannot be directly measured or obtained from field inventories. Also, the parameters are highly correlated, and changing one parameter can influence the other parameters related to field measured data. A broadleaf tree object consists of trunks, boughs, branches (level 1, 2, 3), twigs, leaf stems, leaves, leaf plates and envelope. Depending on the complexity required, some or all of these objects can be exported to a facetized geometry format.
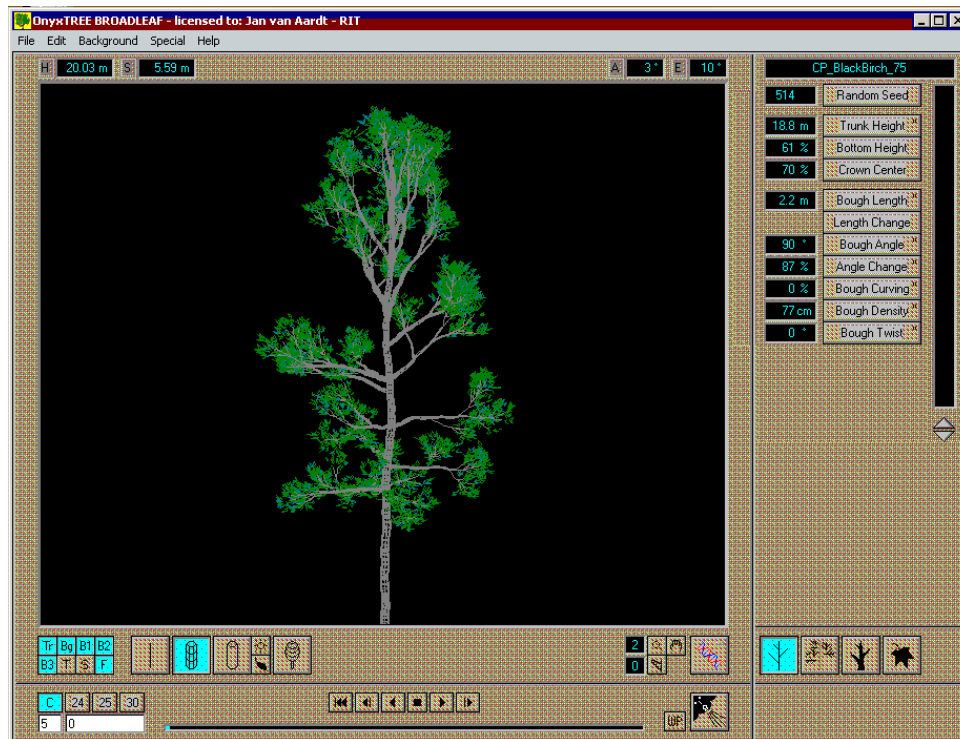
Figure 4.7: An example of a virtual broadleaf tree built using OnyxTree (Onyx Computing, 2015). Some tree parameters used to modify the shape of the tree are shown on the right.

The generation of a virtual tree in OnyxTree is more of an art than science. Further, creating small or large trees takes about the same time and have similar number of objects. From the four classes of crown dominance in forest stands, the suppressed and intermediate trees are relatively small compared to dominant trees. In a dense forest such as hardwood, the dominant and co-dominant trees covers the canopy crown predominantly and the other two classes are invisible from the top of the canopy. Hence, in this research, only dominant and co-dominant tree species and their inventory information are used to build the synthetic trees. Table 4.3 shows the distribution of dominant and co-dominant trees from forest inventory data. A typical tree from OnyxTree looks more like a park tree than a forest-like tree. In a forest, leaves (or branches) of the tree have to compete for sunlight due to the thick crown from adjacent trees. As a result, the branches tend to grow more vertically unlike park trees, where the branches tend to grow horizontally. While building the virtual trees, the OnyxTree parameters are adjusted to emulate a forest-like trees.

The tree generation process is best summarized in a flowchart as shown in Figure 4.8. In OnyxTree, a default tree model is selected for each tree species. The DBH and the height of the tree if available are used as input parameters to adjust the default tree. In such cases

Table 4.3: Distibution of dominant and co-dominant tree species in Hardwood forest site.

Tree Species	# Trees	Relative Percent
Black Birch	3	8 %
Paper Birch	4	11 %
Red Maple	14	38 %
Red Oak	15	40 %
Yellow Birch	1	3 %

where the height of the tree is unavailable, the height distribution of the species in the site is used to estimate the height such that the DBH and height of the tree is nominal for that species. The forest inventory data also provides information on the LAI for some of the trees. A forest-like tree is generated by adjusting the parameters such as pruning and the number of branches, to match closely with the LAI for that species in the forest inventory. More emphasis is given to match the LAI due to the difficulty in generating an accurate tree model in a virtual environment. Additionally, LAI provides a reasonable approximation to the structure in the canopy and it influences the bidirectional reflectance property more than any other tree parameters for medium-resolution sensors such as Landsat (Hasegawa et al., 2006, Shabanov et al., 2003).

Once a virtual tree species is generated in OnyxTree, a DIRSIG simulation of the virtual tree is performed to verify its LAI. This process is iterated until the LAI of the tree matches approximately to the LAI from the inventory. Once all the tree species are modeled in OnyxTree and verified, the next task is to determine their spatial positions relative to each other.

The forest inventory data provides geographical location for the trees in the site. Initial attempt at manual placement of dominant and co-dominant trees in their corresponding coordinates resulted in large gaps within the canopy since they cover only about 40% of the trees in the site. The gaps reduced but only moderately when intermediate and suppressed tree coordinates are also used to place dominant and co-dominant trees. These gaps are attributed to the inaccurate representation of the virtual tree, especially, its crown diameter. The drawbacks posed by manual placement to get thick canopy cover is mitigated by automatic tree placements using Poisson disc sampling method.
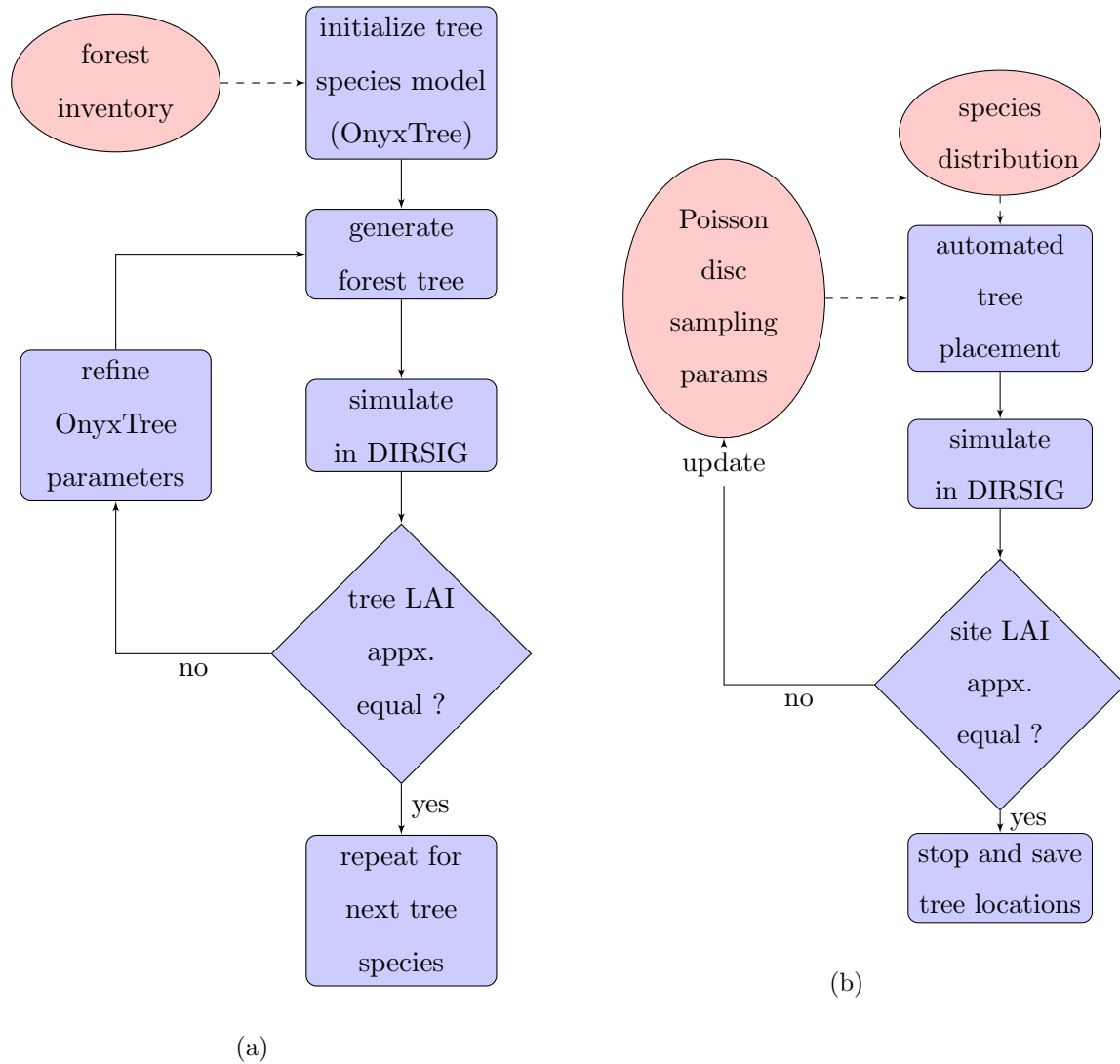
Figure 4.8: Workflow showing the (a) tree generation process in OnyxTree, and (b) the tree placement process using Poisson disc sampling method .

The Poisson disc sampling method provides an automated way of placing trees that approximates many natural spacings as observed in homogeneous forests (Cook, 1986, Mitchell, 1987). This is a method in which random samples are generated such that no two points are close to each other. A radius parameter defines how close any two points are sampled and changing the radius parameter changes the point density. The basic idea is to generate random points around existing points and check whether they can be added without affecting the minimum distance criteria. This results in an even but random distribution of samples. There are several algorithms available for producing a Poisson disc sample set. In this research, the code developed by Herman Tullekem in Python language (Herman Tullekem, 2015) is used for determining the samples. Figure 4.9 shows an example comparing three different random

sampling methods. Although the point density is approximately the same in all the cases, the Poisson disc sampling method provides a cluster-free uniform sampling, simulating the real conditions without any increase in the tree density.

(a)   (b)   (c)
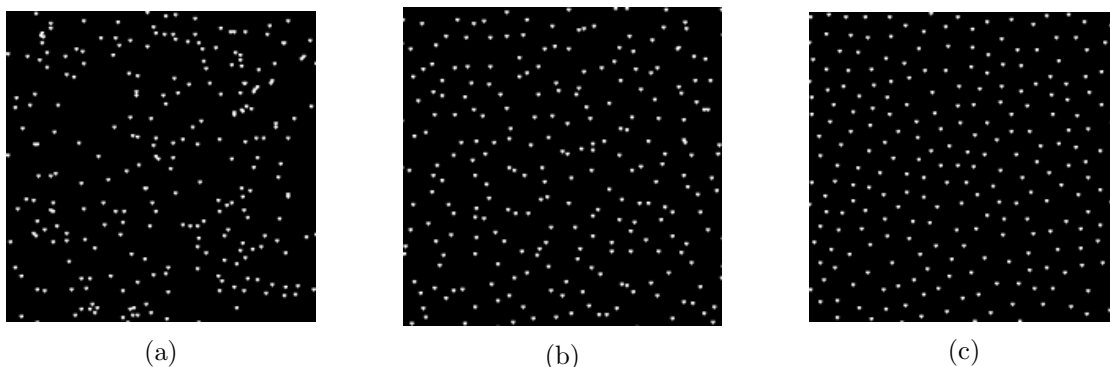
Figure 4.9: An example showing three different random sampling methods. The point density is approximately the same in all the three methods. (a) Uniform random points. The X and Y coordinates are randomly selected based on image extent. (b) Jittered grid. The image is divided into grids, and one point is randomly selected from every cell in the grid. (c) Poisson disc sample points. The Poisson disc sampling method clearly provides a uniform sampling compared to grid or uniform random sampling methods (Herman Tullekem, 2015).

The process of tree placement starts with determining horizontal coordinates (X and Y axes) using Poisson disc sampling method. The sample coordinates are assigned to different tree species such that the distribution of tree species in the site, as shown in Table 4.3, is maintained. The DIRSIG simulation of the virtual forest is used to check for large gaps in the scene, and to estimate its LAI. If the LAI of the simulated forest does not match closely to the LAI from forest inventory, then the process is repeated by adjusting the radius parameter in Poisson disc sampling method and/or by changing the random number seed generator. This approach ensures that the site has similar tree species distribution as the Hardwood forest site and the site and virtual forest's LAI matches closely with each other.

The scene geometry for the forest site consists of placement of trees on the ground, modeled as a plane. As discussed in Section 3.3, DIRSIG provides an option to create primitive shapes such as a plane for modeling a scene. Although it is inaccurate to render the ground as a plane, it is a good approximation for small spatial extent (50 m x 50 m), where the undulations in terrain is very small. The ground plane can also be rotated to account for slope in the ground as will be shown in Section 4.4.2.1. The trees are placed vertically (Z axis) over the ground plane using the option in DIRSIG to place trees on the terrain (to anchor trees over the underlying terrain).

This ensures that the trees grow vertically upwards, with their base connected to the ground, irrespective of the slope of the ground surface.

4.2.3 Forest optical properties

In 2013, research teams at UMB and BU collected the spectra for green leaves, trunks and ground using Analytical Spectral Devices(ASD) field spectroradiometers. They also collected the spectra in 2014 (October) when the leaves started to change color (brown). The brown spectra can be used to simulate the seasonal variation in forest. The two spectral collection campaigns did not collect leaf spectra for all the tree species in the site. Also, spectra for smaller tree objects like stems and twigs were not collected. The thick canopy cover on the hardwood site reduces the photon interaction with trunk and branches compared to leaves. Further, trunk and branches are opaque, causing any scattered photon from leaves to be reflected or absorbed. Due to the unavailability of trunk spectra for all but one species of interest, all the tree species are assigned the same trunk spectra. Similarly, black birch and yellow birch trees share the same leaf spectra. The yellow birch trees cover only about 3% of the forest site, and hence this does not introduce large error in using black birch spectra. The teams collected reflectance spectra for leaves, but a leaf is highly transmissive in the NIR and SWIR spectral region. Therefore, the spectral transmittance for leaves are estimated using the PROSPECT inversion model.

PROSPECT is a leaf optical model for estimating leaf-level reflectance and transmittance (Feret et al., 2008, Jacquemoud and Baret, 1990, Jacquemoud et al., 1996). The model is based on the representation of the leaf as one or more absorbing plates with rough surfaces. The model assumes the leaf is a stack of N identical elementary layers separated by N-1 air spaces. The input variables to the PROSPECT model are chlorophyll concentration (Cab), carotenoid concentration (Ccx), equivalent water thickness (Cw), dry matter content (Cdm), brown pigment (Cbp) and the leaf structure parameter (N). The model provides the leaf directional-hemispherical reflectance and transmittance as output. Feret et al. (2008) showed that the reconstruction of reflectance and transmittance from leaf biophysical parameters in the 400 nm - 2450 nm wavelength domain using PROSPECT-5 is accurate with low to negligible biases. It is also possible to invert the PROSPECT model and provide biophysical parameters from leaf reflectance spectra. The inverse model uses non-linear optimizations to minimize the root mean square difference between the measured reflectance (and transmittance, if available)

spectra and PROSPECT derived spectra. The workflow of model inversion process is shown in Appendix E Figure E.1. The Matlab code provided by Frederic Baret (2015) is used to perform PROSPECT model inversion to get leaf transmittance and reflectance.

The transmittance for field collected leaf species is derived through a two step process. First, the inverse model is used to estimate the biophysical parameters from the field-measured leaf spectra for each species. This is followed by forward PROSPECT model to estimate leaf reflectance and transmittance from biophysical parameters derived from inversion. Figure 4.10 shows a comparison between measured leaf's reflectance and transmittance to the derived data using the inverse-forward approach.
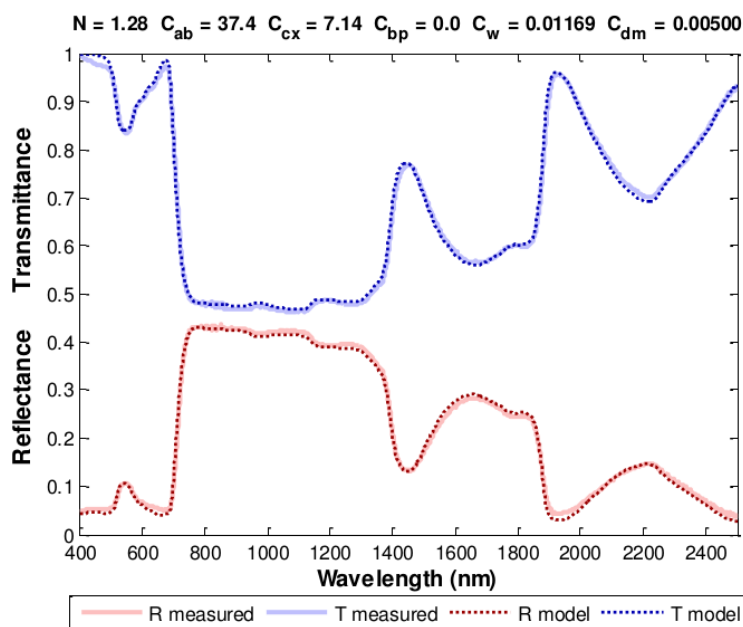
Figure 4.10: Comparison of measured reflectance spectra to the derived spectra using inverse-forward PROSPECT model for Poplar (Liriodendron tulipifera). The measured spectra matches very close to the estimated spectra in the VIS-NIR-SWIR spectral regions (Frederic Baret, 2015).

Similar to geometry, the optical properties are assigned to the trees and ground. The ground litter reflectance spectra collected by UMB and BU teams are used to define the ground optical properties and to provide texture to the ground. The results for the virtual forest geometry and the associated optical properties for the hardwood site are discussed in Chapter 5.

4.3 Measuring BRDF using DIRSIG

The following section describes the approach used in measuring the reflectance for varying illumination, view angles and wavelengths using DIRSIG as a virtual goniometer. The radiometric computations in DIRSIG were validated in Section 4.1, however, the parameters used in the radsolvers have to be optimized for accuracy and processing time. A sensitivity study is performed to determine the optimal parameter settings for the DIRSIG simulations.

4.3.1 DIRSIG sensitivity study

The Geodensity radsolver is used for all canopy radiative transfer computations in DIRSIG. It uses many parameters, some of which are used to optimize the simulation's processing time and accuracy in estimating the results correctly. The radsolver uses Monte-Carlo propagation of photon bundles from source to scene to build the flux density map across the surfaces as discussed in Section 3.3.6. In this radsolver, a bundle of rays are propagated randomly from the source and each bundle's interaction on different materials are stored and indexed for estimating the sensor reaching radiance.

Monte Carlo simulation relies on a large number of repeated measurements to estimate the population statistics from sample statistics. In the case of light interactions in the canopy, a large number of photon bundles are required to illuminate the scene to approximate reality. Increasing the number of photon bundles can provide more accurate estimate of the flux density, but will also increase the computational time linearly. Hence, it is necessary to estimate the optimal number of photon bundles for a canopy scene without impacting the fidelity or processing time.

In DIRSIG, the canopy interactions are captured by a sensor pointing at the scene. DIRSIG uses ray tracing technique to estimate the sensor reaching radiance for the detector's IFOV. A number of ray samples (sub-samples) are sent to the scene for every detector's IFOV. Each sample estimates the radiance along the direction of the ray. The detector's response is estimated by averaging the radiance from all the ray samples within an IFOV. A better estimate of each detector's response can be achieved by increasing the number of sub-samples in each dimension (length, width) of a detector, which increases the computational time quadratically for a rectangular detector array. Hence, an optimal number of sub-samples are necessary to reduce any impact on the processing time or accuracy of the results.

A sensitivity analysis can be useful to find an optimal value for the number of photon bundles and the detector's sub-samples by varying one factor at a time. The Geodensity radsolver generates a flux density map, which is dependent on radsolver parameters (ex: number of bundles, scene extent, etc), wavelength, incident angle (azimuth, zenith), and scene content. The incident angle, wavelength, and scene content are held constant for all the simulations while varying the number of bundles used to illuminate the scene, resulting in multiple flux density maps. For each flux density map, a number of simulations are performed with varying detector sub-samples, eliminating any uncertainty due to the random sampling of Monte-Carlo propagation.

The different configurations for the number of photon bundles and detector sub-samples used in this research are listed in Tables 5.4 and 5.5. The photon bundles illuminate the entire volume of the scene, but for simplicity, the height of the scene is not provided in the table. The best result can be achieved by simulating with the highest number of photon bundles and detector sub-samples that can be used within practical processing considerations. This simulation configuration is considered as the reference to which all the other simulation results are compared. A criterion is needed for determining the optimal configuration and in this case, ΔL (difference in radiance) is compared to the expected noise from a remote sensing system. The SNR of a remote sensing system is dependent on radiance (signal) from a target at a specific wavelength and the design of an instrument. For systems such as Landsat 8 and Sentinel 2, the SNR in the NIR spectral region is about 200. The high SNR is primarily due to high reflectance and transmittance expected from vegetations in NIR spectral region. Knowing SNR and the signal in radiance units, the noise in radiance units can be estimated from their ratio as shown in Equation 4.5.

For the sensitivity analysis, the simulations are performed using a simple frame camera imaging over a forest scene. The image generated from DIRSIG are modulated by the RSR of Landsat 8 (Band 5) to compute the band averaged radiance. The difference between the band average radiance and the reference is compared against the noise estimated at that radiance as shown in Equation 4.7 for all the configurations.

$$N \ [radiance] = \frac{MeanSignal \ [radiance]}{SNR \ [at \ signal \ radiance]} \qquad (4.5)$$

$$L_{eff} = \frac{\int_{\lambda_1}^{\lambda_2} L_\lambda \ R'_\lambda \ d\lambda}{\int_{\lambda_1}^{\lambda_2} R'_\lambda \ d\lambda} \tag{4.6}$$

Since the radiance of the image is measured at discrete wavelength,

$$L_{eff} = \frac{\sum_{\lambda_1}^{\lambda_2} L_\lambda \ R'_\lambda \ \Delta\lambda}{\sum_{\lambda_1}^{\lambda_2} R'_\lambda \ \Delta\lambda}$$

$$\Delta L = | \ L_{eff} \ - \ L_{eff} \ (ref) \ | \tag{4.7}$$

where

$\lambda_1 \ and \ \lambda_2$ are the bandwidth extent for OLI Band 5

R' is the RSR for OLI Band 5

L_λ is the mean radiance of the simulated scene for a specific wavelength, photon bundle and sub-samples

L_{eff} is the band average radiance for a specific photon bundle and sub-sample configuration

$L_{eff} \ (ref)$ is the band average radiance for the highest possible photon bundle and sub-sample configuration

ΔL is the absolute difference in radiance for a specific photon bundle and sub-sample configuration

A plot of absolute difference in radiance with respect to the different photon bundles or the detector sample configurations are expected to be an exponential function as shown in Figure 4.11. It is impossible to discern any differences in radiance when ΔL is below the system level noise (N). The smallest number of sub-samples and photon bundles that satisfies the condition $\Delta L < N$ are then used to determine the optimal parameters.
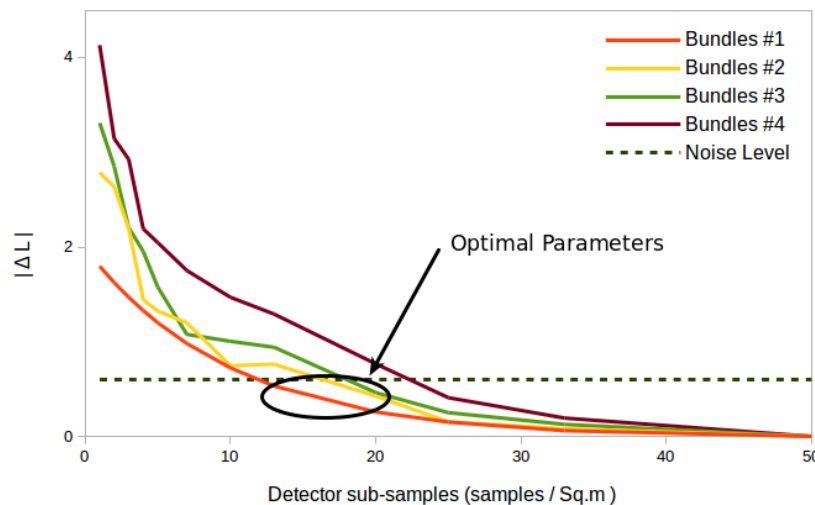
Figure 4.11: Difference in radiance as a function of number of photon bundles and detector sub-samples. The plot indicates a reduction in $|\Delta L|$ when the number of samples are increased (exponential function). Also, increase in photon bundles reduces the difference for a specific number of sub-samples as observed in multiple colors.

4.3.2 BRDF measurements

4.3.2.1 BRDF sampling

BRDF is a function of wavelength, view and illumination angles, and is measured at discrete samples. The number of samples increases geometrically based on the resolution for the view, illumination and wavelength parameters. For example, a low resolution of 10 degrees in zenith and 30 degrees in azimuth direction requires 120 (10 x 12) simulations to represent the hemispherical reflectance surface for a specific wavelength and illumination angle. If the same resolution is used for illumination angles, the number of simulation increases to 14400 (120x120) for each wavelength. As shown in the example above, it is exhaustive and impractical to measure even a low resolution spectral BRDF in the VIS-NIR-SWIR spectral regions. Further, a low resolution spectral BRDF could completely miss important canopy characteristics such as "Hot-Spot" region. Thus, it is necessary to intelligently sample for wavelength, view and illumination angles. The sampling strategy employed in this research is based on the geographical location of the scene, sensor's field of view, and its spectral bands.

Landsat 8 has an equatorial crossing time of 10 AM while Sentinel 2 crosses the equator at 10:30 AM (mean solar time). The relative position of the sun to the forest site, for any day, can be computed from the geographic location of the Harvard site. The sun's approximate zenith

and azimuth angles for the two local crossing times at the Harvard site for a year, are shown in Figure 4.12, and the equations used to estimate the angles are found in Appendix A.

The relative position of the sun over any geographic location follows a similar pattern for a specific time. This pattern is referred as an "infinity curve" and the position and the orientation of the curve is dependent on the geographic location and time of the day. This curve is used to constrain the sampling for the illumination angles. The samples are chosen for the illumination angles such that they are close to the curve for both the sensors. The canopy BRDF is influenced by the leaf-on and leaf-off condition of the forest. In general, the remote sensing community is interested in the leaf-on condition of the forest, which is dependent on the latitude. For the Northern Hemisphere, the leaf-on condition typically varies from April to September. The samples are chosen near that part of the curve between April and September for the Harvard forest since it is in the Northern Hemisphere.
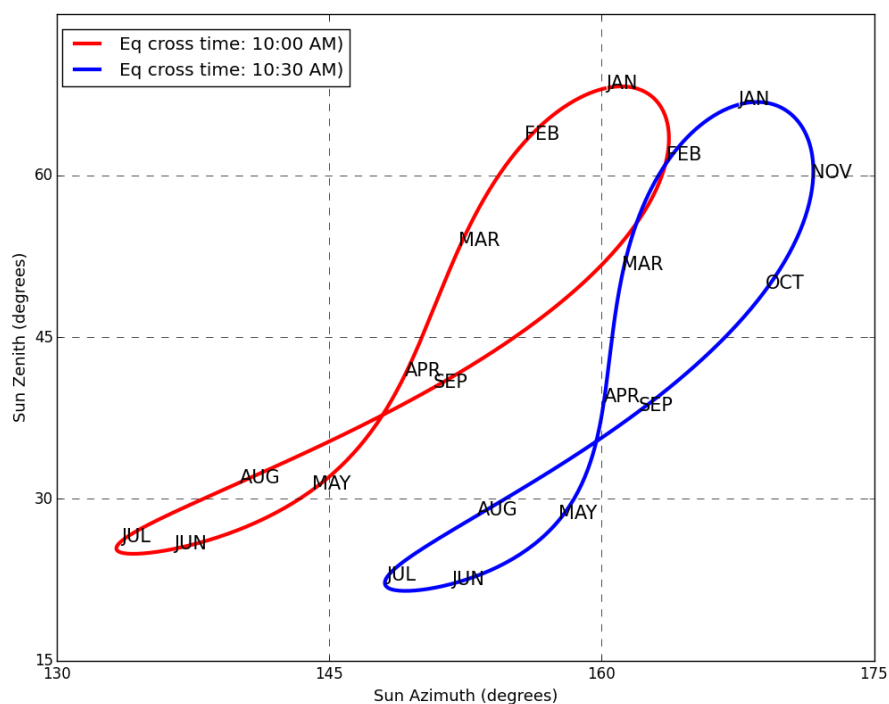
Figure 4.12: Position of sun (azimuth and zenith angles) in the local coordinate system at Harvard forest site [*Lat* : 43.531° / *Lon* : −72.182°] for an entire year.

In general, the atmosphere is used as part of a DIRSIG simulation and is modeled using MODTRAN. The presence of atmosphere leads to downwelled irradiance (diffused skylight) on the scene. The downwelled irradiance is sampled over the entire hemisphere in DIRSIG to estimate the total sensor reaching radiance. Full hemispherical BRDF is required as part of this

computation to estimate the outgoing radiance from the scene. The fitting of the measurements to the BRDF model provides the capability to estimate the reflectance for any illumination angle even when the measurements are limited to a few samples. However, the reliability of the fit coefficients to predict the reflectance correctly is dependent on the range of angles used for modeling the BRDF. This requirement is satisfied by choosing a few samples with large solar zenith angles.

The BRDF is a function of relative azimuth angle instead of sun and view azimuth angle, when the scene is azimuthally symmetric. In general, many biomes in nature including the forest canopy are assumed to be azimuthally symmetric. It is thus important to know the relative azimuth which is computed by subtracting the sun azimuth angle from the view azimuth angle. Both Landsat 8 and Sentinel 2 view the Harvard site from azimuth angles which are dependent on the orbital parameters of the satellite, orbital velocity of the earth, and the geographic position of the site. The azimuth angles for the two satellites are computed and subtracted from the chosen samples to get the relative azimuth angle, which are shown in Table 5.6 and in Figure 5.25.

Similar to the illumination angles, samples for the view angles are constrained based on the sensor's field of view. In the across track direction, the field of view (FOV) of the OLI and MSI sensors are $\pm 7.5°$ and $\pm 10.5°$ respectively. The FOV in the along track direction is less than $\pm 2°$ for the two sensors. This provides a limitation on the number of view angles that need to be sampled for measuring the BRDF. The strategy to choose samples near the expected view angle is appropriate, as the sensors are unlikely to image the ground beyond these view angles. The fitting of sampled measurements to a model provides a consistent and accurate BRDF with high precision, similar ot a high resolution BRDF measurement. If this study is extended to other sensors having wider FOV, view samples may have to be extended according to the sensor's FOV to improve the reliability of the fit coefficients and to get an accurate BRDF. The samples used for view angles in this research are shown in Table 5.6 and Figure 5.25.

The sampling strategy for wavelength is based on the spectral bands of the sensors. The OLI sensor has 9 spectral bands (Table 3.1) and the MSI sensor has 12 spectral bands (Table 3.4). Both sensors have two similar spectral bands that are not useful for vegetation applications (coastal / aerosol band and cirrus band). Avoiding these two bands, the sampling is performed for the remaining spectral bands in the VIS-NIR-SWIR region. For each spectral band, the

samples are chosen such that they can be used to approximate the shape of the RSR for either sensor. This careful selection reduces any error in the BRDF interpolation across the bandwidth of a spectral band due to the shape of sensor's RSR. The wavelength samples used to get the spectral BRDF are shown in Table 5.7 and Figure 5.26.

4.3.2.2   DIRSIG simulation

The decision on optimal parameters for the radiometry solver and the sampling strategy provides the necessary input information to simulate scenes using DIRSIG. The main components of DIRSIG modeling are scene, sensor, platform motion and tasking, and atmosphere as mentioned in Section 3.3.

The forest scene is modeled as discussed in Section 4.2. The same scene is used across all different view angles, sun angles and wavelength to generate spectral BRDF. The sun is used as the only source of illumination and its azimuth and zenith angle defines the location in the local coordinate system. A convenient value is chosen for the magnitude of sun's irradiance, as discussed in Equations 4.3 and 4.4, to convert the DIRSIG radiance image to a reflectance image. The atmospheric component of DIRSIG uses preselected MODTRAN atmosphere as default; however, options are available to simulate the scene without any atmospheric attenuation. Similar to the RAMI validation studies in Section 4.1.2, no atmosphere is used in the simulations. This ensures that the contribution to sensor reaching radiance (or reflectance) comes only from the interactions between the photons and the scene elements. In the DIRSIG simulation, the entire scene is illuminated by the sun using an optimal number of photon bundles.

The sensor used to image the forest is a simple frame camera whose parameters are listed in Appendix C. The camera's detector has a rectangular RSR with less than 1 nm bandwidth to avoid any error in the measurements due to the width of the spectral bands or the shape of the RSR. The camera's FOV at nadir is limited to 90% of the scene extent to reduce any edge effects at the scene boundaries. The image is acquired using an instant exposure of the scene. The position and orientation of the camera, for a specific view angle are determined as shown in Appendix A, Equation A.3.

The measurements are considered as directional (and not conical) since the solid angle subtended by the camera is very small and can be ignored. For example, the solid angle of the camera at an altitude of 100 KM over a scene of 50 m x 50 m is $\sim 0.000012°$ with a IFOV of $\sim 0.017°$.

The reflectance measurement for BRDF is estimated by averaging all image pixels that cover a region of no more than 30 m x 30 m on the ground and the corresponding solid angle measures less than 0.000004°. Warner et al. (2009) suggests that field spectroradiometer instruments with large field of view (in the order of 1° to 8°) should be compensated for their solid angle, and any such measurement should be considered as HCRF. The IFOV of airborne sensors are usually small ($0.02° - 0.2°$) while that of space platform such as MODIS is $\sim 0.1°$. In a strict sense, the measurements from airborne and space sensors are Hemispherical Conical Reflectance Factor (HCRF), but numerically approach to Hemispherical Directional Reflectance Factor (HDRF) values. Since the observations from airborne sensors and MODIS are considered as directional measurements by the remote sensing community, the DIRSIG measurements with solid angle much lower than the airborne sensors are also considered as directional measurements.

It takes one simulation run for each view angle, sun angle and wavelength combination. The total number of simulations performed for measuring BRDF is given by the product of the number of samples for the view, sun and wavelength. An automated script is used to modify the parameters for each of the DIRSIG components to generate all the simulation files for measuring BRDF.

4.3.2.3 3D Region of Interest (ROI)

The constant FOV of the camera causes its projected area on the ground to change as a function of view angle. As a result, the image covers a different extent on the ground when viewed at varying view angles, impacting the way the reflectance is measured. The reflectance of the forest scene, for a particular view angle, sun angle and wavelength is estimated based on the average of image pixels. In an ideal condition, the forest is large and homogeneous, and the reflectance over a region is shift invariant in its spatial position except at the forest boundaries. However, the synthetic forest scene built in DIRSIG is small and inhomogeneous and causes a change in reflectance when measured over different regions. The virtual forest canopy is also limited in size, unlike the real forest canopies where the trees are continuous and the FOV is limited to the crown part of the tree. The main trunk of the trees is invisible in any view angle unless when viewed at the edge of the forest. But, the edges of the forest are typically not modeled in any forest canopy models, owing to the edge effects caused by scattering mechanism between adjacent biomes. It is extremely difficult to simulate the same behavior of a continuous canopy

in the virtual environment, unless the forest is relatively large in size. A large scene requires more photon bundles to illuminate the scene, which significantly impacts the processing time in DIRSIG. The problems with the varying projected area and the crown-limited view of the canopy are solved by the 3D ROI method.

The 3D ROI is a technique that projects the 3D ground coordinates onto a 2D camera plane. The camera's position and Euler rotation angles in the local (ground plane) coordinate system is determined using Equation A.3. The corner points that define the 3D region of interest on the forest scene are determined, based on the height of the living crown and the horizontal extent on the ground. For example, if the height of the living crown is approximately 15 m and the horizontal extent required is 30 m x 30 m, then the 3D regions are selected such that they enclose a box with a height of 15 m to the top of the canopy and a base of 30 m x 30 m, centered within the forest scene at a distance of 15 m from the ground. The horizontal centering of the box is an input parameter and typically, they are centered at the center of the virtual scene. The 3D ground coordinates for the corner points of the box are projected on to the image coordinates (camera plane) using Equation 4.8.

$$
R_{XYZ} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(r_x) & \sin(r_x) \\ 0 & -\sin(r_x) & \cos(r_x) \end{bmatrix} \begin{bmatrix} \cos(r_y) & 0 & -\sin(r_y) \\ 0 & 1 & 0 \\ \sin(r_y) & 0 & \cos(r_y) \end{bmatrix} \begin{bmatrix} \cos(r_z) & \sin(r_z) & 0 \\ -\sin(r_z) & \cos(r_z) & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

$$
\begin{bmatrix} U \\ V \\ W \end{bmatrix} = [R_{xyz}] \left(\begin{bmatrix} X_g \\ Y_g \\ Z_g \end{bmatrix} - \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \right)
$$

$$
\begin{bmatrix} x \\ y \end{bmatrix} = -f \begin{bmatrix} \frac{1}{\Delta x} & 0 \\ 0 & \frac{1}{\Delta y} \end{bmatrix} \begin{bmatrix} \frac{U}{W} \\ \frac{V}{W} \end{bmatrix} + \begin{bmatrix} x_p \\ y_p \end{bmatrix}
$$

(4.8)

where

r_x, r_y, r_z are the orientation angles of the camera

X_g, Y_g, Z_g are the ground coordinates of a corner point in local coordinate system

X_c, Y_c, Z_c are the camera coordinates in local coordinate system

$\Delta x, \Delta y$ are the detector size in X and Y

f is the focal length of the camera

x_p, y_p are the principal point in the camera coordinate system

A polygon representing the projection of the 3D box onto a 2D plane is generated by a convex hull operation on the image coordinates. A mask built from the polygon limits the pixels that are used to estimate the reflectance for a specific view angle. This is illustrated in Figure 4.13. The projection of a 3D cuboid on the camera results in a 2D polygon. The shape of the polygon is dependent on the size of the 3D cuboid, and the view angle of the camera. The 3D ROI ensures that the same ground region is being used for measuring reflectance across varying view angles and provides a consistent measurement. Further, providing the height of the living crown in 3D ROI construction restricts the view only to the crown of the forest canopy.
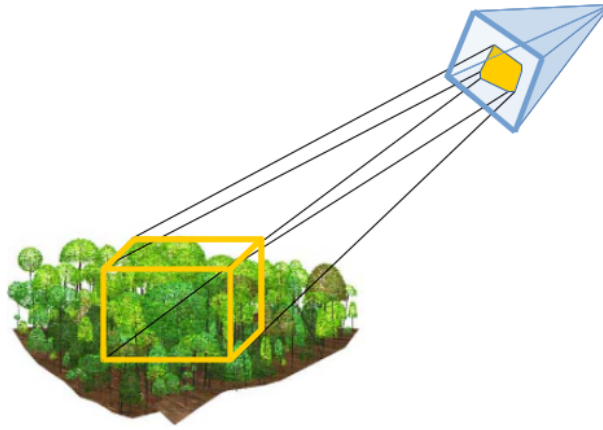
Figure 4.13: Projection of a 3D cuboid on to the camera results in a 2D polygon. The shape of the polygon is dependent on the view angle of the camera and the size of the 3D cuboid.

Although the 3D ROI technique works well, it cannot solve the limitations due to geometry and may introduce very small error. For example, at large zenith angles, the 2D projected ROI encompasses the trees that are not part of the box, as the trees inside the box could be hidden by the trees at the boundaries. At smaller zenith angles, few samples outside the box could be included in the mask. These artifacts are due to the limitation with the projection to resolve along one dimension. This is not an issue, since most of the sensors have smaller field of view ($\pm 20°$) and the effects introduced in such cases are very small and are statistically insignificant. The other limitation from geometry is due to an increase in ground extent covered by images for large zenith angles. This is a limitation that cannot be resolved using a camera with fixed

FOV. The number of pixels that covers the same region on the ground change, and is highest at nadir and reduces with an increase in zenith angle. This limitation is reduced by increasing the number of samples, such that the difference in average statistics is insignificant over a large numbers of samples. For example, a camera with a GIFOV of 1 m at nadir and 30 m x 30 m ground extent averages reflectance over 900 samples. The reduction in the number of pixels at small zenith angles ($\pm 20°$) is no greater than 100 samples leading to a small, if not negligible error. Any error in the measurement due to these limitations are considered unimportant since the measurements are used to fit to a BRDF model and are not used directly. The fitting of measurements to the RossLi BRDF model is explained in the next section.

4.4   BRDF Modeling

BRDF is a function of view and illumination angles defined over a hemispherical dome. Measuring complete BRDF even with a moderate azimuth and zenith resolution is difficult as mentioned earlier. To characterize a scene in the real-world requires complete BRDF due to skylight contribution and can, at best, be described only by models. The semi-empirical RossLi BRDF model is defined by very few parameters and can be modeled easily in DIRSIG. The RossLi BRDF model is discussed in Section 3.2.1, which in its simplest form can be expressed as in Equation 4.9.

$$\rho(\theta_i, \theta_v, \phi, \lambda) \; = \; f_{iso}(\lambda) \; + \; f_{geo}(\lambda) \; K_{geo}(\theta_i, \theta_v, \phi) \; + \; f_{vol}(\lambda) \; K_{vol}(\theta_i, \theta_v, \phi) \qquad (4.9)$$

where

$f_{iso}, f_{geo}, f_{vol}$ are the isotropic, geometric and volumetric coefficient

K_{geo}, K_{vol} are the geometric and volumetric kernel

The geometric kernel is dependent on two parameters, HB and BR, which defines the shape and relative height of the tree crown. These values are estimated for each tree model from OnyxTree using the height and radius of the crown in horizontal and vertical direction, as shown in Figure 4.14. The number of trees distributed within a section of interest for each modeled tree is used as its weight and a weighted mean is computed to represent the HB and BR for the scene as shown in Equation 4.10.

$$
\begin{aligned}
HB &= \sum_{i=1}^{n} w_i \left(\frac{h}{b}\right)_i \\
BR &= \sum_{i=1}^{n} w_i \left(\frac{b}{r}\right)_i
\end{aligned}
\tag{4.10}
$$

where

h, b, r are the height to the center, vertical radius and horizontal radius of the crown
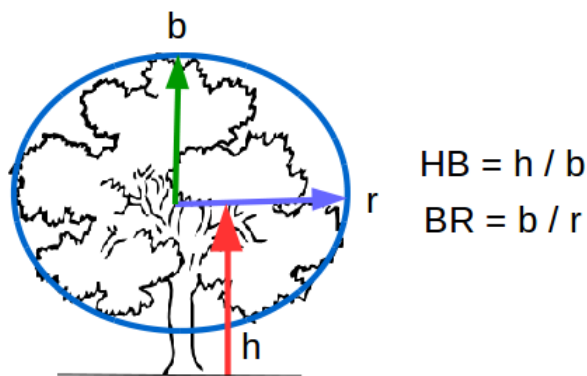
Figure 4.14: The two parameters (HB,BR) used in geometric kernel of RossLi model

The Ross-Thick and Li-Sparse model is used to model the virtual forest since its LAI is greater than 1 ($LAI >> 1$) and has nominal tree distribution. The measured reflectance data is used to estimate the model coefficients using a least square model fit. The formulations used to perform the least square solution are shown in Equation 4.11.

$$
\begin{bmatrix} \rho_1 \\ \vdots \\ \rho_n \end{bmatrix}_{(nx1)} =
\begin{bmatrix} 1 & K_{geo}(1) & K_{vol}(1) \\ \vdots & \vdots & \vdots \\ 1 & K_{geo}(n) & K_{vol}(n) \end{bmatrix}_{(nx3)}
\begin{bmatrix} f_{iso} \\ f_{geo} \\ f_{vol} \end{bmatrix}_{(3x1)}
$$

This is in the same form as $A x = b$, whose solution is given as

$$
x = (A^T A)^{-1} A^T b
$$

$$
\begin{bmatrix} f_{iso} \\ f_{geo} \\ f_{vol} \end{bmatrix} = \left(\begin{bmatrix} 1 & K_{geo}(1) & K_{vol}(1) \\ \vdots & \vdots & \vdots \\ 1 & K_{geo}(n) & K_{vol}(n) \end{bmatrix}^T \begin{bmatrix} 1 & K_{geo}(1) & K_{vol}(1) \\ \vdots & \vdots & \vdots \\ 1 & K_{geo}(n) & K_{vol}(n) \end{bmatrix} \right)^{-1}
$$
$$
\begin{bmatrix} 1 & K_{geo}(1) & K_{vol}(1) \\ \vdots & \vdots & \vdots \\ 1 & K_{geo}(n) & K_{vol}(n) \end{bmatrix}^T \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_n \end{bmatrix}
$$

<div align="right">(4.11)</div>

The solution to the least square fit provides 3 coefficients that describes the RossLi BRDF model. The RossLi coefficients as a function of wavelength are generated by performing an independent least square fit on the observations for each wavelength. DIRSIG has the capability to interpolate the coefficients between any two wavelength using a linear interpolation. As shown in Table 5.7, the sampling distance for wavelength is $\approx 5\ nm$ for the OLI and the MSI spectral bands, which is adequate for averaging over their RSR, as the RSR typically span ten to hundreds of nanometers.

The RossLi model approximates the canopy BRDF correctly for low solar zenith (SZN) angles but not as accurate for high SZN angles. This is not an issue as the remote sensing of the forest canopy is typically performed when the solar zenith angles are low. However, there is a need for reflectance at high solar zenith angles to account for the skylight interaction with the scene. This requirement is satisfied by fitting RossLi coefficients based on the solar zenith angle. In this approach, the solar zenith angle of $50°$ is used as the cutoff to delineate between high and low angles. Therefore, all the measurements for a specific wavelength with high solar zenith angles are used to generate independent RossLi coefficients. Similarly, all the measurements with low solar zenith angle are used to generate another set of RossLi coefficients. Thus, for a specific wavelength, two different sets of independent RossLi coefficients are generated based on the solar zenith angle. While this approach reduces the error in the fit coefficients when the solar zenith angle is low, at high solar zenith angles, the reduced number of samples (Table 5.7) used for fitting the measurement can introduce large errors in fit coefficients. The results are shown in Figures 5.28 and 5.29. The contribution of this error to the sensor reaching radiance is small and is discussed in the next section.

4.4.1 Measurement and model sensitivity

4.4.1.1 BRDF measurement sensitivity

Sensitivity of the RossLi fit coefficients to measurements can be studied using Leave One Out Count (LOOC) validation method. LOOC is a technique typically used to determine outliers in the data, but in this case, is used to study the sensitivity of the fit.

In the LOOC validation approach, the RossLi fit coefficients are determined initially using all the measurements for a specific wavelength. The RossLi coefficients, thus determined, are considered as reference coefficients for that wavelength. One of the view angles is randomly selected and removed from the measurements. This would reduce the number of measurements by n, where n is the number of sun angles. The RossLi fit coefficients are recomputed and the relative difference between the reference and the new fit coefficients are determined. This method is repeated by removing a different view angle and the corresponding relative error with respect to the reference is computed. The low sensitivity of the model to view measurements can be inferred from the consistency of the relative error for excluded view angles. The same approach is repeated for illumination angles to verify its sensitivity in modeling the RossLi coefficients. Finally, the sensitivity to wavelength can be inferred by verifying the consistency of the relative error across the VIS-NIR-SWIR region.

The LOOC validation approach can also be used to estimate the effect on reflectance in addition to the RossLi coefficients. In this technique, the RossLi coefficients are used to estimate the BRF for different sun and view angle combinations. Relative difference in reflectance between the reference BRF (using the reference coefficients) and LOOC BRF ("one sample removed" coefficients) for each combination of sun and view angles is estimated. The summary statistics such as mean and STD provides a good indication of the sensitivity of the measurements to the model in the reflectance domain. As in the case with coefficients, the sensitivity to wavelength can be inferred by checking the consistency of the relative error across the VIS-NIR-SWIR region.

4.4.1.2   RossLi BRDF model sensitivity

The RossLi BRDF model is a semi-empirical BRDF model, that approximates the real radiative transfer in the canopy. The approximations used in this model formulation affects the overall results (expected in the model), to fit the measurements accurately. However, Koukal and Schneider (2010) have shown that the RossLi BRDF model is comparable to many of the empirical and semi-empirical models, and equally accurate, if not better, than the non-linear models such as the RPV model. Although the model is considered accurate, a simple study on the validation of the RossLi BRDF model for forest canopy will be useful to eliminate any uncertainties associated in using this model to represent the canopy BRDF.

The approach to validate the RossLi BRDF model is simple and is performed for a few wavelengths to show its effectiveness in modeling the BRDF measurements. DIRSIG is used to measure the reflectance for the entire hemispherical dome in view angles, by sampling at a high resolution in the view azimuth and view zenith directions. The measurement is repeated for a few sun angles. Similar to the canopy BRDF model coefficient estimation, the least squares solution technique is used to fit the DIRSIG measurements to the model. The RMSE of the fit residuals and coefficient of determination (R^2) are used to assess the accuracy, sensitivity, and the limitations of the RossLi BRDF model in representing the hemispherical BRDF of the canopy.

4.4.1.3   BRDF sensitivity for high zenith angles

The radiance reaching the sensor is primarily influenced by the direct and diffuse (downwelled) contributions from the scene. These contributions vary depending on the optical properties of the scene. Hence, any error in the modeling of BRDF can potentially introduce error in the direct and diffuse contributions. The BRDF model is shown to accurately fit the measurements for low solar zenith angles (Section 5.4) and its error in the direct contributions are negligible. However, the model fit residuals are large for high solar zenith angles, which can introduce measurable error in the diffuse contributions. Hence, it is necessary to estimate the potential error in the sensor reaching radiance due to the downwelled contribution for high solar zenith angles. This is achieved by comparing the sensor reaching radiance between the downwelled irradiance (diffuse) and total irradiance (direct and diffuse) contribution. The governing equation to estimate the

sensor reaching radiance is shown in Equation 4.12.

$$L(total) = \text{direct radiance} + \text{diffuse radiance} + \text{upwelled radiance}$$

$$L(\theta_v, \phi_v, \lambda) = E_{sun}(\theta_s, \phi_s, \lambda) \; \tau_1(\theta_s, \phi_s, \lambda) \; \tau_2(\theta_v, \phi_v, \lambda) \; \cos(\theta_s) \; \rho(\theta_v, \phi_v, \theta_s, \phi_s, \lambda) \; +$$

$$\int_{i=0}^{2\pi} \int_{j=0}^{\frac{\pi}{2}} L_d(\theta_j, \phi_i, \lambda) \; \tau_2(\theta_v, \phi_v, \lambda) \; \rho(\theta_v, \phi_v, \theta_j, \phi_i, \lambda) \; \cos(\theta_j) \; \sin(\theta_j) \; \delta\theta_j \; \delta\phi_i + L_u$$

Since the downwelled radiance is computed for discrete angles,

$$L(\theta_v, \phi_v, \lambda) = E_{sun}(\theta_s, \phi_s, \lambda) \; \tau_2(\theta_v, \phi_v, \lambda) \; \tau_1(\theta_s, \phi_s, \lambda) \; \cos(\theta_s) \; \rho(\theta_v, \phi_v, \theta_s, \phi_s, \lambda) \; +$$

$$\sum_{i=0}^{2\pi} \sum_{j=0}^{\frac{\pi}{2}} L_d(\theta_j, \phi_i, \lambda) \; \tau_2(\theta_v, \phi_v, \lambda) \; \rho(\theta_v, \phi_v, \theta_j, \phi_i, \lambda) \; \cos(\theta_j) \; \sin(\theta_j) \; \Delta\theta \; \Delta\phi + L_u \tag{4.12}$$

$$\text{Downwell Contribution (DC)} = \frac{\text{diffuse}}{\text{diffuse} + \text{direct} + \text{upwelled}} = \frac{\text{diffuse}}{\text{total}}$$

$$\text{DC} = \frac{\sum_{i=0}^{2\pi} \sum_{j=0}^{\frac{\pi}{2}} L_d(\theta_j, \phi_i, \lambda) \; \tau_2 \; \rho(\theta_v, \phi_v, \theta_j, \phi_i, \lambda) \; \cos(\theta_j) \; \sin(\theta_j) \; \Delta\theta \; \Delta\phi}{L(\theta_v, \phi_v, \lambda)} \tag{4.13}$$

where,

θ_s, ϕ_s are sun zenith and azimuth angles

E_{sun} is the direct solar irradiance

L_d is the downwelled radiance

L_u is the upwelled radiance

L is the sensor reaching radiance

τ_1, τ_2 are transmission along the path from sun to scene and scene to sensor respectively

ρ is the BRDF of the scene

The fit coefficients derived from the least square method, as explained in the previous section, are used to determine the BRDF. The input parameters such as geographic coordinates of the forest site (Harvard) and the time and day of the year (10:00 AM equatorial crossing time on summer solstice) are used to estimate the solar angles, exoatmospheric irradiance and earth sun distance. The scattering by the atmosphere is simulated using MODTRAN, for a mid-latitude summer atmospheric profile. The parameters, thus determined, are used in

DIRSIG to estimate the transmission through the atmosphere for sun-scene-sensor path, direct irradiance, and downwelled radiance for different azimuth and zenith angles. The downwelled contribution, with respect to the total contribution, can be computed using Equation 4.13.

The downwelled contribution as determined using the above equation includes solar irradiance from all angles. This can be further divided into summation of two parts based on low or high zenith angles, as shown in Equation 4.14. Since there is a possibility of large error in reflectance for high solar zenith angles, it is helpful to estimate its contribution to the sensor reaching radiance. This can be easily computed as shown in Equation 4.15.

$$
\begin{aligned}
\text{Diff Rad} = {} & \sum_{i=0}^{2\pi} \sum_{j=0}^{\frac{\pi}{3.6}} L_d(\theta_j, \phi_i, \lambda) \ \tau_2 \ \rho(\theta_v, \phi_v, \theta_j, \phi_i, \lambda) \ \cos(\theta_j) \ \sin(\theta_j) \ \Delta\theta \ \Delta\phi \\
& + \sum_{i=0}^{2\pi} \sum_{j=\frac{\pi}{3.6}}^{\frac{\pi}{2}} L_d(\theta_j, \phi_i, \lambda) \ \tau_2 \ \rho(\theta_v, \phi_v, \theta_j, \phi_i, \lambda) \ \cos(\theta_j) \ \sin(\theta_j) \ \Delta\theta \ \Delta\phi
\end{aligned}
\tag{4.14}
$$

$$
\text{DC} = \frac{\displaystyle\sum_{i=0}^{2\pi} \sum_{j=\frac{\pi}{3.6}}^{\frac{\pi}{2}} L_d(\theta_j, \phi_i, \lambda) \ \tau_2 \ \rho(\theta_v, \phi_v, \theta_j, \phi_i, \lambda) \ \cos(\theta_j) \ \sin(\theta_j) \ \Delta\theta \ \Delta\phi}{L(\theta_v, \phi_v, \lambda)}
\tag{4.15}
$$

where,

DC is the contribution due to downwelled radiance with high solar zenith angle

$\frac{\pi}{3.6}$ is equivalent to 50° zenith angle

The results for the downwelled contribution as a function of wavelength will be shown in Chapter 5. Thus, the above approach helps to estimate the potential error in sensor reaching radiance due to the less precise modeling of BRDF for high illumination angles.

4.4.2 Auxiliary BRDF models

4.4.2.1 Forest on a sloped terrain

The generation of RossLi model coefficients using DIRSIG BRDF measurements was discussed in the previous section. These coefficients were derived for a forest on a flat terrain. However, in reality, forests are often found in mountainous region with steep gradients. Although the slope

of the underlying terrain has its normal pointing away from the Z direction ("up" direction), the trees grow vertically up to reduce the impact of gravity. Hence, rotating the entire BRDF by the same angle as the slope of the surface is not the right option for modeling forest canopies. A better approach is to use the same technique as discussed in the previous section, and fit a model to the scene with trees placed on a terrain with slope.

This can be accomplished by modeling a new forest scene, with trees placed on the rotated ground as shown in Figure 4.15. In DIRSIG, the ground surface is modeled as a plane and can be easily rotated to the required gradient. The trees are anchored to the terrain such that the horizontal positions are retained and their vertical position is accurately determined and placed over the underlying and tilted ground surface without any gaps between the trees and the surface. A new scene, thus modeled, is used to generate the BRDF measurements in DIRSIG, and are fitted to the RossLi BRDF model to estimate its coefficients.
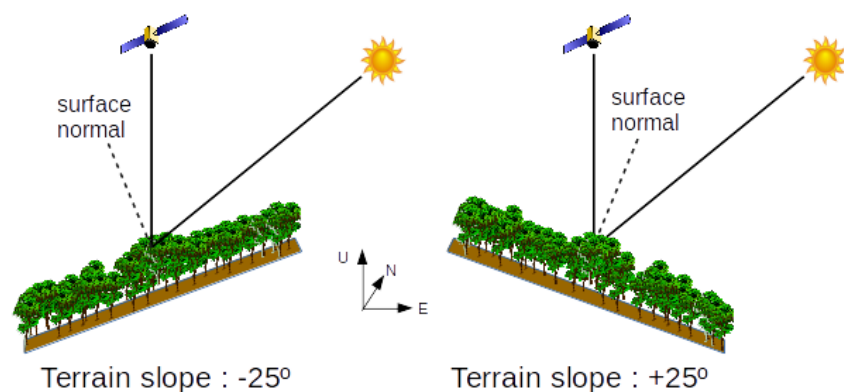
Figure 4.15: Two different orientations of the forest are shown to indicate the growth of the forest trees on a sloped terrain. The trees grow vertically upwards on a rotated ground.

4.4.2.2 Diversity in model coefficients

Forests can be broadly classified based on their geographical location and climatic condition, as tropical rainforest, deciduous, coniferous, and boreal forests. Each of these forest categories can be further sub-divided based on many factors such as, the type of species, tree structure, distribution, biomass production, location, etc. As seen in Section 4.2, it is extremely difficult to characterize and accurately represent a specific forest site in the virtual environment. Hence, to study the impact of sensor and environmental factors for all the different types of deciduous forest is beyond the scope of this research. However, it is important to recognize that the species distribution, height, DBH, and spatial distribution of the trees in the forest differ significantly

within a specific site. These factors affect the BRDF of the forest, especially its magnitude more than its shape, for a moderate resolution BRDF. These diversities in the BRDF for a particular forest type can be approximated by varying the RossLi model coefficients. There is no direct relationship between the coefficients and the variability within the forest. However, any variability in the forest is modeled by a different set of model coefficients. Typically, the model coefficients do not differ significantly for a small change in the forest. It is expected that the variation between the forest types (deciduous vs coniferous) induce larger differences in the RossLi coefficients than the variations within a forest type. This section discusses the approach used to vary the RossLi model coefficients, to account for the variation expected within the deciduous forests.

In this approach, the BRDF measurements using a 3D ROI, as discussed in Section 4.3.2.3, is centered at different locations within the image to generate multiple sections of the forest. This is illustrated in Figure 4.16. The locations for the center of the ROI are selected such that they are equally spaced within the image and the two adjacent sections overlap by about 80%.
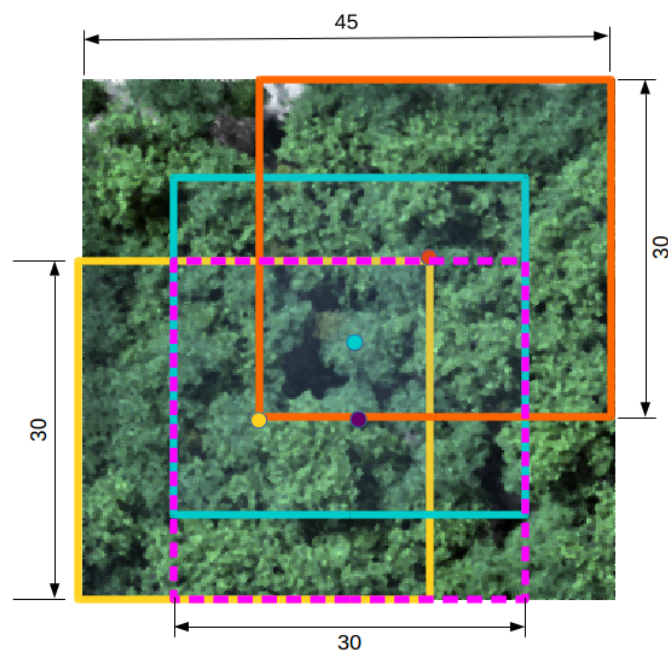
Figure 4.16: Nine sections can be selected based on the overlap criteria, but four sections (Yellow, Magenta, Cyan, Orange) with their center are shown here for clarity. Each of the section covers 30 m x 30 m extent and the adjacent sections (Yellow,Magenta) have 75% overlap, while the diagonal sections (Yellow,Cyan or Cyan,Orange) have little more than 50% overlap

.

Each of these sections provide BRDF measurements for different view, illumination and wavelength samples. These BRDF measurements are used to fit the RossLi model, as discussed in Section 4.4, to generate distinct RossLi model coefficients representing the possible variations within the forest site. For example, the camera captures an image covering 45 m x 45 m forest extent, while the 3D ROI selects 9 sections each about 30 m x 30 m (horizontally) to estimate the reflectance measurement for a specific wavelength, view and illumination angle. The overlap constraint ensures that the BRDF of any two sections differ only minimally in its magnitude and shape and thus satisfying the initial assumptions. As seen in the above example, this method provides a limited number of diverse RossLi coefficients. However, it is possible to generate a large number of distinct coefficients by using the normal random sampling, based on the mean and standard deviation from the limited set of RossLi coefficients.

Random samples for each of the coefficients $(f_{iso}, f_{geo}, f_{vol})$ can be selected independently from their normal distribution with corresponding mean and standard deviation, if the three coefficients are uncorrelated. It is evident from the RossLi model formulations (Section 3.2.1) that there exists a strong correlation between isotropic and Ross model coefficients and also between isotropic and Li model coefficients. The correlation between the three coefficients is illustrated in the Figure 4.17. The strong correlation between the coefficients negates the possibility of sampling them independently.

The three coefficients are treated as correlated variables and can be transformed into a set of uncorrelated variables using the Principal Component Analysis (PCA) technique. The set of observations for each coefficient is transformed into the principal component domain. The minimum and the maximum of the observations in the principal component domain can be used as the minimum and the maximum (a,b) for the uniform distribution $U(a,b)$. Random samples are chosen from this uniform distribution, for each of the three principal component basis independently. The independent selection is valid since the variables in the transformed space are uncorrelated. Alternatively, normal distribution can be used instead of uniform distribution to generate random samples from the transformed coefficients. The coefficients from the center forest and the deviations among the different sections of the forest can be used to represent the mean and standard deviation of the normal distribution respectively. A scaling parameter on the standard deviation can further increase or decrease the required variability in the generated samples. The chosen samples are then transformed back to the original basis using the inverse transform. An example for random sampling and inverse principal component transform are

shown in Figures 4.18 and 4.19.

(a)

(b)

(c)
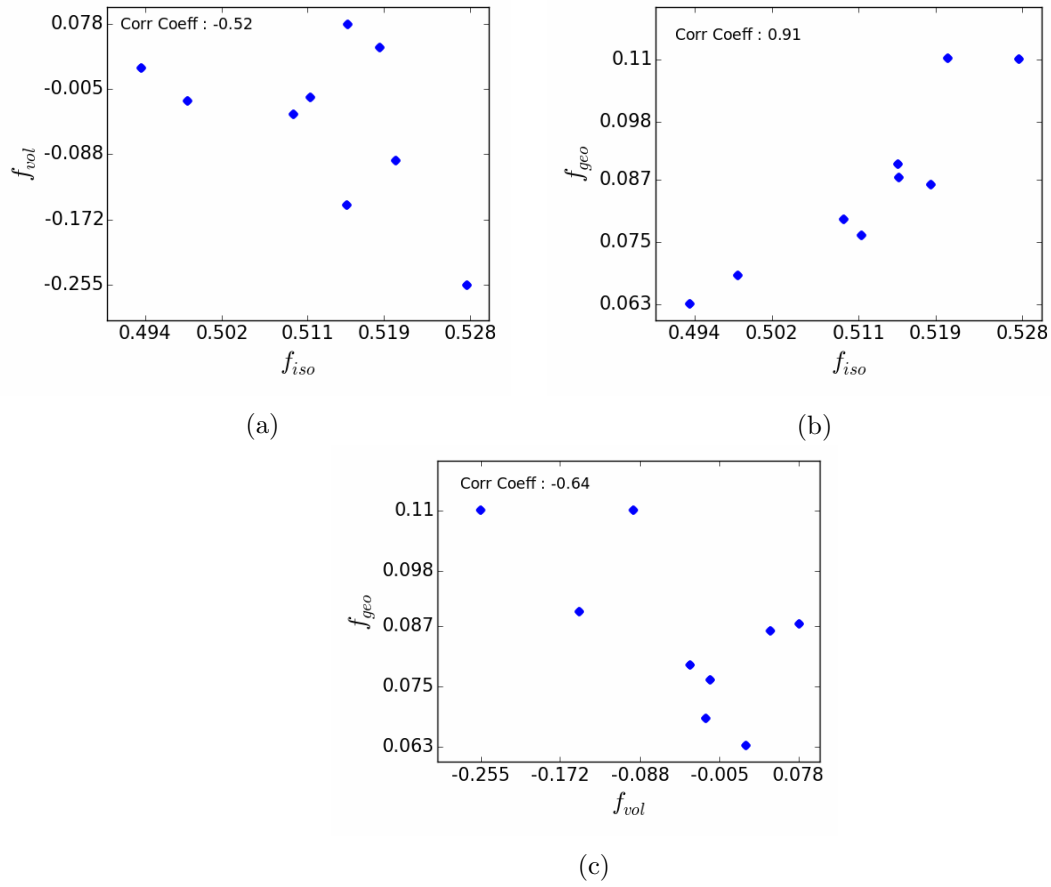
Figure 4.17: Correlation between RossLi model coefficients. (a), (b), (c) indicate the correlation between isotropic and volumetric, isotropic and geometric, and volumetric and geometric coefficients. The correlation coefficient between f_{iso} and f_{geo} is very high (0.91). The nine coefficients are estimated for the forest scene covering 30 m x 30 m.

The variability within the random BRDFs can be estimated by evaluating the relative variation of each random BRDFs to the BRDF of the central section of the forest based on the reflectance for various sun and view angle combinations, as shown in the equation 4.16.

$$\Delta\rho_{ij} = 100|\frac{\rho_{(CF)j} - \rho_{ij}}{\rho_{(CF)j}}| \quad \forall \quad i \in \{\text{random BRDFs }\} \ , \quad j \in \{1, 2, \dots, n\}$$

$$\mu_{\rho_i} = \frac{\sum\limits_{j=1}^{n} \Delta\rho_{ij}}{n} \tag{4.16}$$

where,

μ_{ρ_i} is the mean relative variation for a specific random BRDF

n is the number of sun and view angle combinations

CF is the BRDF for the central section of the forest
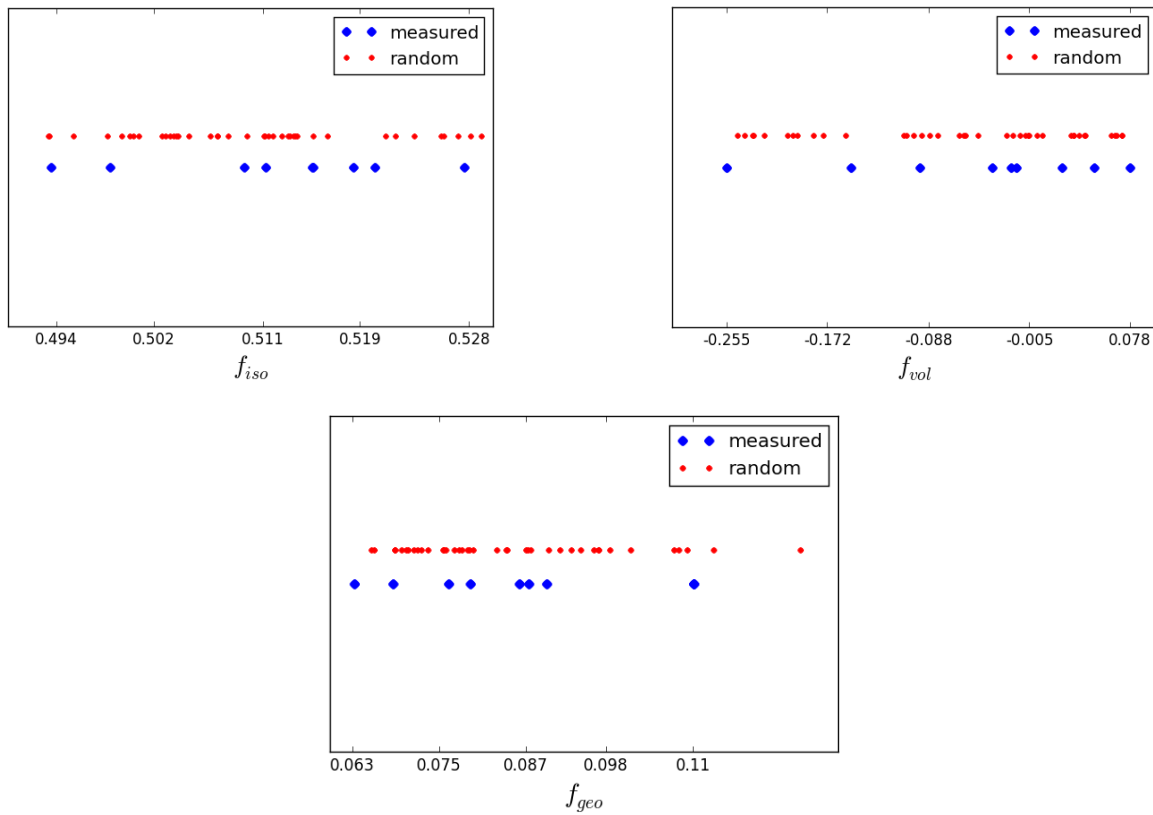
Figure 4.18: Distribution of nine measured RossLi coefficients, and 40 randomly sampled (from uniform distribution) coefficients transformed from the Principal Component space. The figure shows that the range for the random samples are approximately equal to the range for the measured coefficients (see X-axis). The measured and random samples are offset in the Y-axis to show their range clearly.

The approach explained in this section can generate large number of random and distinct canopy BRDFs that are very useful to represent the randomness exhibited in the real-world forests. The results for these auxiliary model coefficients are shown in Chapter 5.4.2.
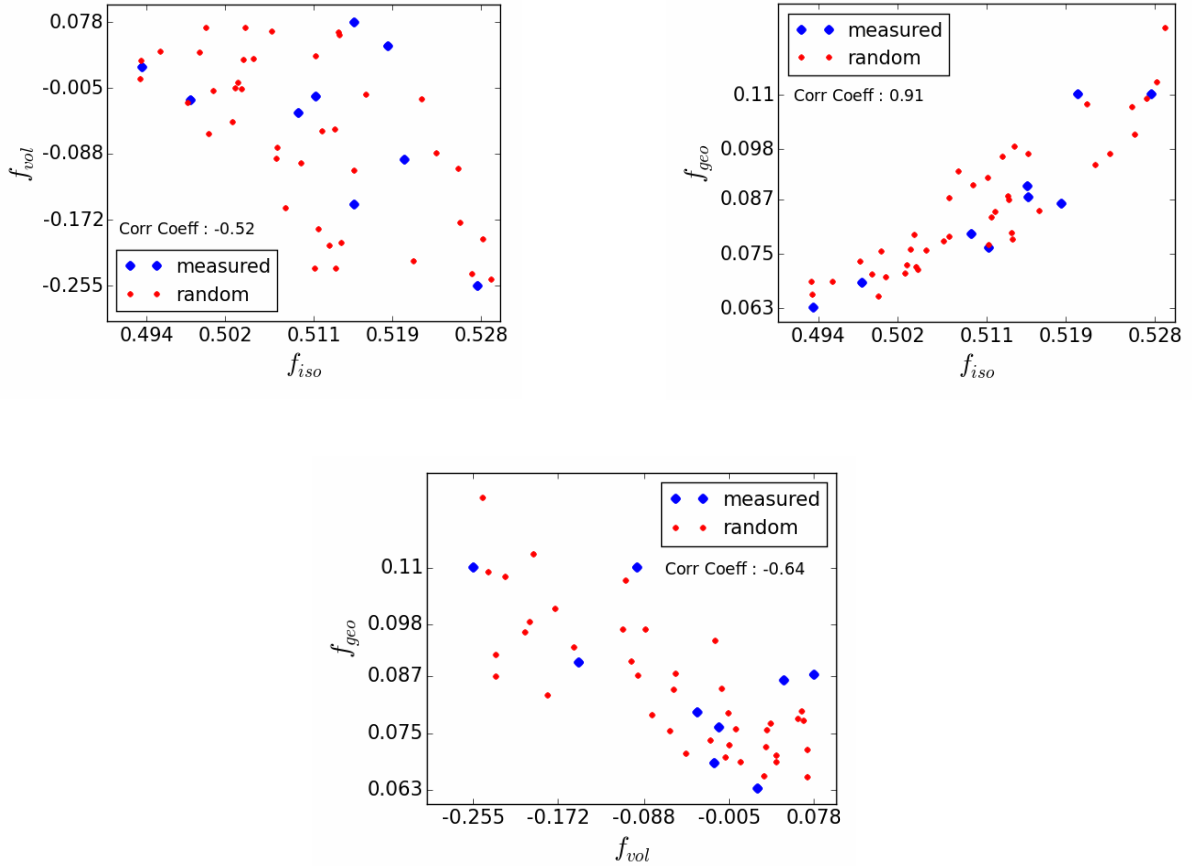
Figure 4.19: Correlation between RossLi coefficients for measured and randomly sampled data. The correlation coefficient of the 40 random samples is approximately equal to the correlation coefficient of the nine samples shown in Figure 4.17. The distribution of the random samples are very similar to the nine measured samples, as expected.

4.4.2.3 Temporal BRDF

Similar to the BRDF modeling for forest canopies on sloped surfaces, the seasonal changes of a forest can be modeled by changing the optical properties of the leaves. For example, in autumn season, the color of the leaves change from green to yellowish-brown, due to unmasking of carotenoids and antocyanins caused by significant reduction in the production of chlorophyll. The changes in the chemical structure within a leaf affects its optical properties. Typically, seasonal change affects both the shape and the color of a leaf. However, the changes in the optical properties of the leaves affects the spectral BRDF of the canopy much more than a small change in its shape for leaf-on conditions. The geometry of the trees in the scene can be retained while varying its optical properties. The leaf spectra collected by the research teams from UMB and BU, during autumn season, can be used to replace the reflectance properties for the modeled trees. The process of BRDF measurement and model fitting, as discussed earlier,

can be performed to generate a senescence BRDF. However, it is very difficult to quantify the senescence of a forest and cannot be mapped to a forest variable.

Hence, in this research, the temporal BRDF is modeled by varying the geometric properties of the forest. This is accomplished by modeling the forest canopy for varying levels of defoliation as described in Section 4.5. The variation in the BRDF can then be mapped to a forest variable such as LAI to better understand its effect.

4.5 Modeling forest canopy signal

Change detection techniques are widely used for mapping and monitoring the forest cover to detect the declining health and vigor of forests. These change detection techniques rely on two key aspects; the biophysical variation in a forest introduces corresponding variation in its reflectance, and secondly, the changes in the received signals can be detected by a calibrated sensor. The new generation sensors such as OLI and MSI are well calibrated and have very high SNR to detect the small changes on the ground. However, the effects due to the sensor and environmental factors can potentially affect the detection and estimation of the actual changes on the ground (or in a forest canopy). One of the key objectives in this research is to assess the impact of these factors, hence it is important to represent the effects in a meaningful way. Although the effect of these factors in reflectance or sensor reaching radiance can be measured and are useful to asses the factors' influence, it does not provide an intuitive understanding or a direct relationship to the actual changes on the ground.

Foresters typically are interested in parameters such as LAI, defoliation or biomass to assess the changes, but these parameters cannot be directly measured from moderate resolution remote sensing data (Landsat, Sentinel). However, within a simulation environment such as DIRSIG, a forest canopy can be modeled at a very high resolution with individual trees and leaves, and parameters such as defoliation can be easily modeled and/or estimated from the 3D geometry of the canopy.

Thus, the different levels of forest defoliation can be used as a measure of signal within a forest canopy and its relative variation can be used as a metric to evaluate the impact of the sensor and environmental factors.

4.5.1   Modeling the Harvard forest defoliation

Section 4.2 discussed the approach used in modeling the geometry and optical properties of the Harvard forest canopy in DIRSIG. The same modeling process is used to generate defoliated forest canopies. Defoliation refers to the loss of leaves from different branches of the trees. This can be due to a natural process of senescence or due to infection (e.g. gypsy moth). In the case of infection due to pests, the defoliation is rapid and affects the leaf area of the canopy with small to no changes to the leaf's optical properties, i.e, the leaves in the canopy are likely to show good health but are reduced in number due to pest infestation. For example, if less than 50 percent of the tree crown is defoliated, most hardwoods will experience only a slight reduction in radial growth and the trees, in general, can withstand one or two consecutive defoliations McManus et al. (1992). Hence, the defoliated canopy is modeled with an assumption that its optical properties does not change for varying levels of defoliated geometry. Note that in some cases, the infestation can cause spectral changes in the leaf in addition to loss of leaves. In such cases, the effect due to defoliation (in BRDF) will have larger impact than what is discussed here.

The forest geometry as modeled in Section 4.2.2 is considered as a canopy without any defoliation. For each of the tree models in that geometry, leaf facets are selected randomly and are then removed keeping the secondary level branches and twigs intact. The pests that feed on leaves (e.g. chewing insects) can eat an entire leaf, edges of the leaves, chew holes in the centers of leaves, skeletonize the leaves or eat only the upper or lower portion of the leaf (Tree diseases, 2016). In the majority of these cases, parts of the leaves are eaten, and hence the approach of removing leaf facets rather than an entire leaf is a valid approximation to the actual defoliation observed in forests. The number of leaf facets removed for each tree model is dependent on the levels of defoliation. For example, for a 20% defoliation, 20% of the facets from each tree model are randomly selected and removed. The different levels of defoliation and their corresponding number of leaf facets for each tree model are shown in Table 4.4.

The different levels of the defoliated forest canopies are constructed from the corresponding defoliated tree models with the same optical properties in DIRSIG, as discussed in Section 4.2.3. For these defoliated forest scenes, the methods suggested in Sections 4.3 and 4.4 are used to measure the BRDF and to fit these measurements to the RossLi BRDF model. Thus, every defoliated forest canopy is represented by an independent BRDF model. Since the LAI is

Table 4.4: Different levels of defoliation and the number of facets used to model the trees

Trees	Defoliation Levels					
	0%	10%	20%	25%	30%	40%
Black Birch	52328	47096	41863	39246	36630	31397
Paper Birch	51390	46251	41112	38543	35973	30834
Red Maple 1	80800	72720	64640	60600	56560	48480
Red Maple 2	106130	95517	84904	79598	74291	63678
Red Maple 3	24310	21879	19448	18233	17017	14586
Red Maple 4	99280	89352	79424	74460	69496	59568
Red Oak 1	66234	59611	52988	49676	46364	39741
Red Oak 2	215404	193864	172324	161553	150783	129243
Red Oak 2	368774	331897	295020	276581	258142	221265
Yellow Birch	190640	171576	152512	142980	133448	114384

directly related to the defoliation, the different levels of defoliation can be useful to establish a relationship between the variation in reflectance and LAI.

4.5.2 Modeling the signal

The defoliated BRDF provides the ability to compare the reflectance for varying degrees of defoliation. However, in most change detection applications, it is more important to measure the change from the reference data rather than estimating the actual changes in radiometric units. In this research, the change detection is estimated relative to the reference and thereby avoids any scaling or non-intuitive units for comparisons. The relative variation can be defined as shown in Equation 4.17, where the reference and the data are assumed to be of the same units. Since it is a fraction of two same units, the relative variation is unitless and it is scaled by 100 to get the variation in percentage. Thus, any changes from the reference can be easily interpreted as a relative change in percentage.

$$\text{Relative Variation} = \frac{100 * |\text{Data} - \text{Reference}|}{\text{Reference}} \tag{4.17}$$

As mentioned earlier, the different levels of defoliation can be interpreted as the signal levels, while the relative variation of the defoliation with respect to the reference can be interpreted as the manifestation of the signal. The level of defoliation can be easily measured in units such as LAI in a simulation environment, but these parameters cannot be directly measured from the

moderate resolution remote sensing data (Landsat, Sentinel). However, they can be inferred from calibrated data products such as reflectance data, at-sensor radiance data and also from higher level products such as the Normalized Difference Vegetation Index (NDVI). Since the defoliation is characterized as signal, the relative variations are estimated for different types of data products and are summarized below.

4.5.2.1 Signal for at-sensor radiance data

The signal for radiance data can be computed by estimating the radiance under different sensor and environmental conditions. Table 4.5 shows the visibility conditions, view and sun angles, and the two RSRs used to estimate the at-sensor or Top-of-Atmospheric radiance(TOA) using the DIRSIG tool.

The DIRSIG simulation consists of a scene to represent the ground, a sensor to image the scene, and the parameters to model the atmosphere. In this case, the sensor is assumed to fly at Landsat-8 altitude and has a single detector with an IFOV equivalent to 30m on the ground. The visibility conditions are provided to the MODTRAN tape5 file to simulate the atmospheric attenuations for mid-latitude summer atmosphere with rural aerosol. The scene is modeled as a spheroid with the radius of the Earth and its surface property is modeled as BRDF (according to the level of defoliation). The detector uses both the RSRs (OLI ,MSI) as two bands but with the same line of sight angles (across track view angles) shown in Table 4.5. The view angles are represented only in the across-track direction, since the OLI and MSI have very narrow field of view in the along-track direction ($< 2°$). The different simulations for each defoliation are shown pictorially in Figure 4.20.

Table 4.5: Sun angles, defoliation levels, X-Track angles, Visibility conditions, RSRs, and the spectral bands used in the sensitivity analysis

Sun angles		Defoliation levels	X-Track angles (deg)	Visibility (KM)	RSR	Spectral bands
Zenith	Azimuth					
35	145	0%	-12	10	OLI	RED
30	137	10%	-7.5	15	MSI	NIR
33	157	20%	0	20		
25	150	25%	7.5	25		
27	135	30%	12			
35	165	40%				
20	150					
23	132					
40	150					
38	153					
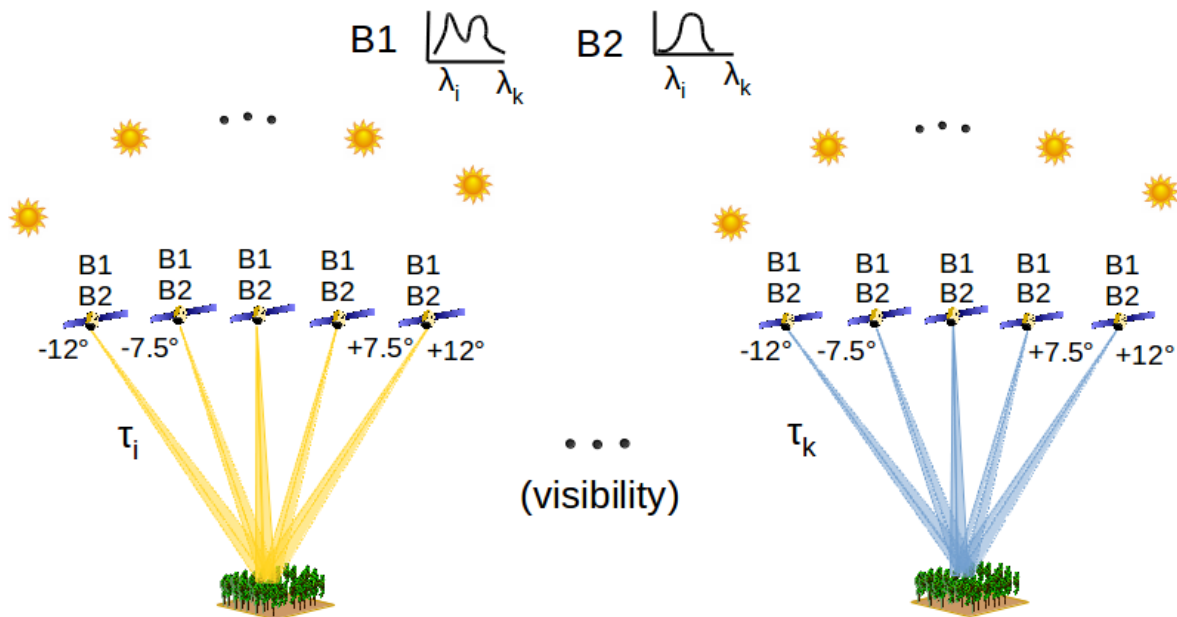
Figure 4.20: Pictorial representation of different DIRSIG simulations. For the same surface BRDF, different visibility conditions (τ), view conditions (5 angles), sun angles (10), and the RSRs (B1, B2) for a specific spectral band are simulated. The same combinations are simulated for different defoliated BRDFs for estimating the relative variations

.

The number of simulations from 10 sun angles, 5 across track angles, 4 visibility conditions, and two RSRs for 6 defoliation levels $(0\%, 10\%, 20\%, 25\%, 30\%, 40\%)$ result in 2400 simulations for each spectral band. Typically, red and NIR spectral bands are used for monitoring the changes in the canopy and so the simulations and analyses were only performed for these two spectral bands. The relative variation for a spectral band from these simulations can be determined

using the formula in Equation 4.18.

$$\Delta L_{ij} = 100 \left| \frac{L_{ij} - L_{(i=0\%)j}}{L_{(i=0\%)j}} \right| \quad \forall \quad i \in \{0\%, 10\%, 20\%, 25\%, 30\%, 40\%\} \quad , \quad j \in \{1, 2, \ldots, n\}$$

$$S(i)_L = \frac{\sum\limits_{j=1}^{n} \Delta L_{ij}}{n}$$

$$\sigma_{S(i)_L} = \sqrt{\frac{\sum\limits_{j=1}^{n} (\Delta L_{ij} - S(i)_L)^2}{n-1}}$$

$$(4.18)$$

where,

n is the number of sun, view angles, RSR and visibility combinations

$S(i)_L$ is the signal estimated based on TOA radiance for a specific defoliation level i

The differences are calculated for different BRDFs but with the same sensor and environmental conditions, and hence are devoid of any effects from these factors. Further, the standard deviation shows an estimate of the deviations across different simulation combinations. A small standard deviation for different levels of defoliation indicates the validity of the characterized signal.

4.5.2.2   Signal for reflectance data

TOA reflectance :   For clear scenes, a reduction in between-scene variability can be achieved through a normalization for solar irradiance by converting the at-sensor radiance to planetary reflectance (Landsat 7 Science Data Users Handbook, 2009) also known as TOA reflectance. This data typically contains both the surface and atmospheric reflectance of the Earth. This is computed using the formula shown in Equation 4.19.

$$\rho_p = \frac{\pi L_\lambda d^2}{\text{ESUN}_\lambda cos(\theta_s)} \tag{4.19}$$

where,

ρ_p is the planetary reflectance

L_λ is the spectral radiance at the sensor's aperture

d is the earth-sun distance in astronomical units

$ESUN_\lambda$ is the mean solar exoatmospheric irradiance

θ_s is the solar zenith angle

The mean solar exoatmospheric irradiance for red and NIR spectral bands are 1547 and 1044 respectively. The earth-sun distance (d) changes with the day of the year and can be obtained from Landsat 7 Science Data Users Handbook (2009). The variation in d is small within the growing season, but is still determined for each DIRSIG simulations to reduce any errors in the estimation of the signal. An approximate day of the year and the corresponding d for each solar zenith and azimuth angles (from Table 4.5) are determined by matching them with the expected solar angles over the Harvard forest, calculated based on the Landsat 8 and Sentinel 2 equatorial crossing time. The TOA radiance from the DIRSIG simulations is then converted to its corresponding TOA reflectance by using the above formula. The relative variations for the TOA reflectance data are then determined using Equation 4.18 by replacing the radiance with the corresponding TOA reflectance.

Surface Reflectance derived using ELM :

The at-sensor radiance can be converted to surface reflectance using any of the atmospheric compensation techniques. One such technique is to use the reflectance panels on the ground to compensate for the atmospheric effects using the ELM method. In this method, two panels of known Lambertian reflectances are used and their corresponding radiances are observed by the sensors. A linear relationship is established between at-sensor radiance and its reflectance for these panels. The slope and intercept of this line can then be used to convert any other radiance data to its surface reflectance. The formulations for ELM are shown in Equation 4.20.

$$L = (\frac{E_s cos(\theta_s)}{\pi} + L_d)\tau \; r_d + L_u$$
$$L = mr_d + b$$

$$(4.20)$$

where,

E_s is the mean exoatmospheric irradiance

r_d is the reflectance factor of the Lambertian panel

L_u is the upwelled radiance

τ is the transmission from the ground to the sensor

Thus, knowing the reflectance factor of the two panels and their radiance, the linear equation can

be solved to estimate m and b. In this method, it is assumed that the atmospheric parameters τ, L_u, and L_d are constant over the entire scene for all the view angles. In general, the ELM compensation method is performed for the real data by imaging well calibrated panels at the same time of image acquisition to avoid any effects due to atmospheric and solar angle differences. The entire image is then corrected with only one set of ELM measurements (i.e., one slope and intercept). But this can introduce errors as the non-nadir view angles changes the path length of the transmission. Further, the L_u and L_d parameters may change depending on the atmospheric conditions within the scene.

Hence, in this research work, two different ELM methods are employed for compensation; ELM (typical) and ELM (ideal). The DIRSIG simulation in both cases use Lambertian panels of known reflectance as the BRDF for the spheroid surface. In the ELM (typical) method, the slope and intercept is determined for every sun angle and visibility conditions for both RSRs in each spectral band. The across-track and nadir view angles are compensated by the same gain and intercept. In the case of ELM (ideal) method, the same simulation combinations as shown in Table 4.5 are performed, i.e, every canopy at-sensor DIRSIG simulation is compensated by a unique slope and intercept corresponding to the sensor and environmental conditions. In the ELM (ideal) method, the different view angles are compensated independently and hence this method will provide an ideal compensation for the atmospheric attenuations.

Similar to the TOA reflectance method, the relative variations are then determined using Equation 4.18 for the surface reflectance data derived from both the ELM methods.

4.5.2.3   Signal for the canopy BRDF data

The relative variation using different defoliated BRDFs can be estimated for a specific sun and view angle. But a better estimate of the variation in reflectance can be determined by using more than one sun and view angles that are expected during the growing season over the Harvard forest. The different sun and view angles that are used to estimate the relative variations are shown in Table 4.5. The relative variation for each level of defoliation can be estimated using

the formula shown in Equation 4.21.

$$\Delta\rho_{ij} = 100 \left| \frac{\rho_{ij} - \rho_{(i=0\%)j}}{\rho_{(i=0\%)j}} \right| \quad \forall \quad i \in \{0\%, 10\%, 20\%, 25\%, 30\%, 40\%\} \quad , \quad j \in \{1, 2, \ldots, n\}$$

$$S(i)_\rho = \frac{\sum\limits_{j=1}^{n} \Delta\rho_{ij}}{n}$$

$$\sigma_{S(i)_\rho} = \sqrt{\frac{\sum\limits_{j=1}^{n} (\Delta\rho_{ij} - S(i)_\rho)^2}{n-1}}$$

(4.21)

where,

n is the number of sun and view angle combinations

$S(i)_\rho$ is the relative variation based on the BRDF for a specific defoliation level i

Since the BRDF represents the anisotropic reflectance property of the surface, the relative variation computed as above indicates the mean relative change in the reflectance for different levels of defoliation. It is extremely difficult, if not impossible, to estimate these signals from the real data. This is because, the range of forest defoliation needs to be controlled, properly assessed, and the BRDF for these levels of defoliation needs to be accurately measured. This is one of the advantages of analyzing the data in a simulation environment where all these variables can be controlled and analyzed.

4.5.2.4   Signal for NDVI data

NDVI is one of the most widely used vegetation indices, and it describes the greenness or the relative density and health of the vegetation. Many applications rely on the NDVI data for the assessment of forest canopy cover, defoliation and biomass, and therefore it is used in this research as one of the data products to characterize the signal. NDVI is calculated from the red and NIR wavelengths and it is computed as shown in Equation 4.22. Healthy vegetation absorbs most of the light in the red spectral band and reflects a large portion in the NIR band. In general, NDVI saturates over dense vegetation, but, both the signal estimation and the analyses are performed using relative variations. The use of relative variation as a metric

makes the comparison consistent and makes the high value of NDVI a non-issue.

$$NDVI = \frac{\text{response}_{NIR} - \text{response}_{RED}}{\text{response}_{NIR} + \text{response}_{RED}} \quad (4.22)$$

As shown in the formula, NDVI is the ratio of difference in response to their summation between the red and NIR responses. The response here could be radiance (at-sensor) or reflectance (TOA or surface). Accordingly, the different types of NDVI products can be generated. Typically, reflectance data compensated for atmospheric attenuations are used to generate NDVI, but it is not uncommon to generate the same from TOA radiance data. In this research, NDVI is calculated for all the five types of product; at-sensor radiance, TOA reflectance, surface reflectance using ELM (typical), surface reflectance using ELM (ideal), and surface reflectance directly from the modeled BRDF. The relative variations are estimated for each of these five data products as shown in Equation 4.18 by replacing the radiance with the NDVI products.

4.5.2.5 Curve-fit

The previous section discussed the characterization of changes in the level of defoliation as the signal for different types of data products. This was not attempted earlier as it is nearly impossible to model with the real data. However, with the DIRSIG simulation environment, we were able to evaluate these signals at discrete signal levels (10%, 20%, 25%, 30%, 40%). This in itself is a useful result for the remote sensing community but this can be further improved by identifying a simple and continuous function to describe the relative variation for different levels of defoliation. In this research, two functions are considered: a line fit and an exponential fit.

In the case of a line fit, a linear regression using least squares is used to fit the relative variation with the levels of defoliation. The exponential curve fit also estimates the least square solution, but for a scaled exponential function with offset. The two functional forms are shown in Equation 4.23.

$$\text{Linear Fit}: \ Y = mX + c$$
$$\text{Exponential Fit}: \ Y = ae^{bX} + c \quad (4.23)$$

where,

m, c are the parameters for the linear fit

a, b, c are the parameters for the exponential fit

Y indicates the relative variation for different types of data products

X indicates the different levels of defoliation

For the curve fit, there are 6 sample points (levels of defoliation) including the case with no defoliation. The redundancy in the least squares fit for linear and exponential fits are 4 and 3, if all the data points are used. However, the validity of the fit can be inferred better by leaving one observation out from the least squares model. The signal at the 25% defoliation level is not used in model fitting, resulting in a redundancy of 3 and 2 for the linear and exponential fits respectively. The residual error between the estimated and measured signal at 25% defoliation provides a measure of the accuracy of the two models. The curve fit provides a way to interpolate the level of defoliation for a specific relative variation, for different data products.

Thus by modeling the relative variations in the data products as a function of defoliation, the detection of signal in the presence of noise (effects due to the sensor and environmental factors) can be studied. In this research, one of the objectives is to estimate the effect of these factors when there are no changes on the ground. To analyze these effects, we need to define a specification such as Noise Equivalent Power (NEP), which is defined as the minimum power required for an output signal-to-noise ratio of 1. In this case, we have defined a similar term called Noise Equivalent Defoliation (NED). This is defined as the level/amount of defoliation that is contributed only due to a specific sensor or environmental factor. In other words, the effect due to a specific factor is equivalent to the effect that would be observed when the forest defoliates by a certain amount. By using this definition, the effects are directly related to the actual changes on the ground. The NED is a useful term in understanding the sensitivity of various factors, as it indicates the uncertainty in measuring the actual changes in the forest.

4.6   Sensor and environmental factor analysis

The previous section discussed the use of defoliation as a method to represent the signal and the different levels of defoliation as signal levels. In this section, the methods used to analyze the effect of the different sensor and environmental factors are discussed. Many factors such

as slope of the terrain, aerosol differences in the atmosphere, cloud contamination, detector's responsivity, etc affects the measured radiance, but the analysis in this research is primarily focused on four dominant factors: RSR, solar zenith angle, across-track angle and visibility condition of the atmosphere. The solar azimuth and along track angles were found to be insignificant and were not included in this analysis (see Table 5.24 and Appendix F). The slope of the terrain does impact the sensor reaching radiance, but for time series analysis where the same region on the ground is observed by two or more sensors, the terrain does not have a direct impact. However, it can affect the sensor reaching radiance between the two sensors if their view angles are different which is captured by the effect in the across-track angles.

The sensitivity analysis of these four factors are studied by comparing their effects to the effect produced due to an actual change on the ground. Since it is impossible to control these factors with the real data, the entire analysis is performed using the simulated data generated from the DIRSIG tool. The spectral analysis is performed for the red and NIR spectral bands as they are widely used in the change detection studies.

4.6.1 RSR effects

The RSR for OLI and Sentinel-2 sensors are shown in Figures 3.15, 3.18, 3.19 and their relative shape comparisons for common spectral bands are shown in Appendix E. The center wavelength and the shape of the RSR are very similar between the OLI and MSI sensors in the NIR spectral band, but are dissimilar in the red band and may produce different responses while observing the same target.

The effect of RSR is studied under different visibility conditions, across-track angles, solar angles, and for different levels of defoliation. A pictorial representation of the simulation combinations is shown in Figure 4.21. The 10 different sun angles expected over the Harvard forest during the growing season, five view angles, and the four visibility conditions used in the simulations are listed in Table 4.5. Each of these simulations are performed using the RSRs from the MSI and OLI sensors for NIR and red spectral bands. The relative variation is estimated as shown in Equation 4.24 by taking the ratio of the absolute difference between the two responses with respect to the OLI response. It is important to note that the relative variation is estimated between the MSI and OLI response for the same exact visibility condition, view and sun geometry. This ensures that the calculated variations are only due to their differences

in the RSR and not due to any other effects. The response of OLI and MSI sensors are very similar such that the results do not vary using either of them as reference. The mean relative variation is computed by averaging the relative variations estimated for all the simulations (10 sun x 5 view x 4 visibility = 200 simulations). The mean relative variation indicates the effect of observing the same target using two different RSRs. The standard deviation of the 200 relative variations indicates the uncertainty due to the varying sensor (across-track) and environmental conditions (solar angles and visibility).

$$
\Delta RSR_{ij} = 100 \, \frac{|R(OLI)_{ij} - R(MSI)_{ij}|}{R(OLI)_{ij}}
$$

$$
\forall \quad i \quad \in \quad \{\text{defoliation levels}: 0\%, 10\%, 20\%, 25\%, 30\%, 40\%\} \quad , \quad j \in \{1, 2, \ldots, n\}
$$

$$
RSR_i = \frac{\sum\limits_{j=1}^{n} \Delta RSR_{ij}}{n}
$$

(4.24)

where,

n is the number of sun, view angle and visibility combinations

$R(OLI)_{ij}$ is the response for a specific defoliated BRDF i and a simulation j

RSR_i is the effect of RSR for a signal level i

As mentioned earlier, the level of defoliation changes the magnitude of the response, and so it can be assumed to represent different signal levels. Hence, the mean and STD of the relative variation is estimated for all the signal levels by simulating the same combination with different defoliated BRDFs. Consistency in the relative variation for different signal levels indicate that the effect due to RSR is independent of the signal level. The effects are also estimated for different data products, such as TOA radiance, TOA reflectance, surface reflectance (ELM-ideal, ELM-typical, BRDF) and their corresponding NDVI products for appropriate comparisons with the forest signal.

Figure 4.21: Pictorial representation of the DIRSIG simulations for analyzing RSR effects. For the same surface BRDF, different visibility conditions ($\tau_i...\tau_k$), view conditions ($\pm 12°, \pm 7.5°, 0°$) and the sun angles (10 angles) are simulated for OLI RSR (red) and MSI RSR (blue) (see Table 4.5). The ratio of the absolute difference between the two responses with respect to the OLI response is estimated as the relative variation for a single simulation and a specific BRDF. Average over all the simulations is used as an estimate of the RSR effects. Mean relative variation is also computed for all the signal levels by simulating the same combination with different defoliated BRDFs.

The functional relationship between the level of defoliation and the relative variation (see Section 4.5.2.5), and the mean relative variation for RSR differences are used to estimate the effects of RSR in the defoliation units (NED). ie, the mean relative variation is used to estimate the corresponding level of defoliation from the functional form of the forest signal. The effect of RSR in terms of changes on the ground is a useful metric that indicates the uncertainty when coincident OLI and MSI datasets are used in change detection applications. In some cases, compensation techniques are employed to reduce the impact of RSR differences between the two sensors. The performance of the compensation techniques and its residual errors can be estimated by evaluating the effect due to the RSR, before and after the compensation.

The compensation for RSR effect is typically performed using a Spectral Band Adjustment Factor (SBAF), which takes into account the spectral profile of the target and the RSR of the two sensors.

Two different SBAF techniques are widely used by the remote sensing community. In the first method, the actual shape of the RSR is not used, but the effects are adjusted by estimating the scale factor from the near-coincident and calibrated TOA reflectance product of the two

sensors (Teillet et al., 2006). This method may introduce errors if the two datasets are acquired under different atmospheric conditions, solar, and view geometries. In the second method, the TOA reflectance of the target is calculated from a hyperspectral data and the shape of the RSR is used to estimate the in-band reflectance for a given RSR. The SBAF is then calculated by taking the ratio of the in-band reflectances between the two sensors (Chander et al., 2013). This method is dependent on an accurate estimation of the TOA reflectance and has the same issues with the BRDF effects as with the other method. In both cases, the SBAF factor is band and target specific and in this research, is calculated for both the NIR and red spectral bands for the Harvard forest canopy. The formulation for SBAF is shown in Equation 4.25.

The use of TOA reflectances in SBAF estimation may introduce additional errors due to the differences in the atmospheric conditions between the datasets. An ideal compensation for the RSR effect is to use the target's surface reflectance. Since the BRDF of the forest canopy is available in this research, this can be used to estimate the SBAF for the two sensors. Additionally, the compensation technique is devoid of any errors due to the BRDF effects, since the same view and illumination angles are used in estimating the SBAF. In this research, both the compensation methods (TOA reflectance and BRDF reflectance) are used to verify their effectiveness in reducing the RSR effects.

$$
\begin{aligned}
\text{Method 1}: \quad SBAF_i &= \frac{\hat{\rho}_\lambda(OLI)_i}{\hat{\rho}_\lambda(MSI)_i} \quad \forall \ \ i \in \{1, 2, \dots, n\} \\
SBAF &= \sum_{i=1}^{n} \frac{SBAF_i}{n} \\
\text{Method 2}: \quad SBAF_j &= \frac{\bar{\rho}_\lambda(OLI)}{\bar{\rho}_\lambda(MSI)} = \frac{\frac{\int \rho_\lambda RSR_{\lambda(OLI)} \delta\lambda}{\int RSR_{\lambda(OLI)} \delta\lambda}}{\frac{\int \rho_\lambda RSR_{\lambda(MSI)} \delta\lambda}{\int RSR_{\lambda(MSI)} \delta\lambda}} \\
SBAF &= \sum_{j=1}^{k} \frac{SBAF_j}{k}
\end{aligned}
\tag{4.25}
$$

where,

$\hat{\rho}_\lambda$ is the TOA reflectance for a specific spectral band

$\bar{\rho}_\lambda$ is the effective BRDF reflectance (adjusted by the shape of the RSR)

ρ_λ is the BRDF reflectance of the canopy

n is the number of sun, view , and visibility combinations randomly selected

k is the number of sun, view combinations shown in Table 5.11

For method 1, about 100 simulations are randomly selected from different sun, view and visibility conditions (see Table 4.5). The TOA reflectance computed for these simulations are used to estimate the SBAF factor. For method 2, about 72 sun and view combinations, as shown in Table 5.11, are used to estimate the SBAF factor. Numerical integration is used to approximate the continuous integral for the in-band reflectance and RSR.

4.6.2 Across-track effects

The orbital parameters and the field of view of the two sensors are different. This reduces the possibility of imaging a target at the same view angle by both the sensors. The relative view angles between the two sensors to a target is also dependent on its geographic position. The along track angles for the two sensors are small and in general, any effect due to the differences in the along track is negligible (see Appendix F). However, in the across track direction, the field of view for the OLI and MSI sensors are $\pm 7.5°$ and $\approx \pm 10.5°$ respectively. Since the orbits of MSI and OLI are different, the same target could be imaged in the back-scattering direction by one sensor and in the forward scattering direction by the other. This can introduce differences in the at-sensor reaching response. For forest canopy, the difference in reflectance between the back-scatter and forward scatter direction in the NIR spectral bands could be as high as 30% and hence it is important to estimate their effects. The effect is dependent on the BRDF of the target and is typically not corrected in any change detection applications.

The effect of view angle in the across-track direction is studied by analyzing the DIRSIG simulated responses for different sensor RSRs, solar angles and visibility conditions. A pictorial representation of the simulation combinations for two extreme angles ($\pm 12°$) is shown in Figure 4.22. The 10 different sun angles expected over the Harvard forest during the growing season, the two sensor RSRs, and the four visibility conditions used in the simulations are listed in Table 4.5.
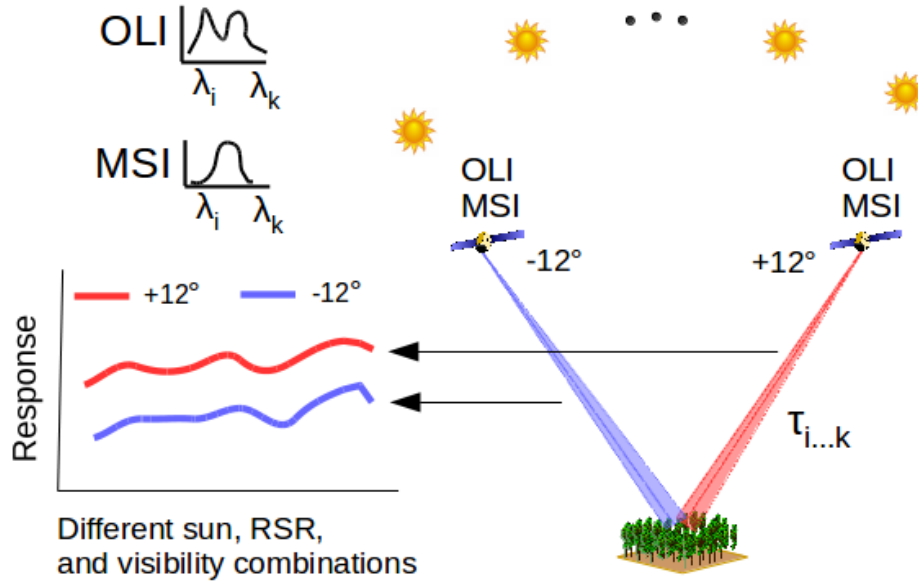
Figure 4.22: Pictorial representation of the DIRSIG simulations for analyzing the view angle effects. For the same surface BRDF, different visibility conditions ($\tau_i...\tau_k$), RSRs (MSI, OLI) and the sun angles are simulated for two extreme view angles ($\pm12°$). The ratio of the absolute difference between the two responses with respect to the ($+12°$) response is estimated as the relative variation for a single simulation and a specific BRDF. Average over all the simulations are used as an estimate of the across-track effects. Mean relative variation is also computed for all the signal levels by simulating the same combination with different defoliated BRDFs.

The two sensors can view the same target at different view angles and therefore, different combinations of view angles between the OLI and MSI sensors can be used to estimate their effect. For example, the OLI sensor may view a specific target at $-7.5°$ whereas the MSI sensor may view the same target at nadir. In this case, all combinations of the sun angles and visibility conditions are used to estimate the effects for the specific view angle differences. The relative variation is estimated as shown in Equation 4.26. It is similar to the relative variation computed for the RSR, except that the simulations used are constrained by specific view angle differences. The downside of this comparison is that the estimated effect is a combination of effects due to the sensor RSR and the view angle differences. The effects only due to view angle differences can be estimated for two extreme view angles ($\pm12°$) as shown in Equation 4.27. In this case, the relative variation is estimated between the two extreme view angles for a specific RSR, visibility condition, and sun geometry. This ensures that the calculated variations are only due to their differences in view angles and not due to any other effects. In this case, the relative variations from 40 simulations (10 sun x 1 view x 4 visibility) for each RSR is averaged together to estimate the mean effects of the view angle differences for two extreme angles. The STD of the relative variations indicates the uncertainty due to the different sun angle, RSR and

visibility conditions. The effects are also estimated for other defoliated BRDFs to evaluate the consistency across different signal levels.

$$\Delta XT_{ij}[kl] = 100 \frac{|R(OLI)_{ijk} - R(MSI)_{ijl}|}{R(OLI)_{ijk}}$$

$$\forall \quad i \quad \in \quad \{\text{defoliation levels}: 0\%, 10\%, 20\%, 25\%, 30\%, 40\%\} \quad, j \in \{1, 2, \dots, n\},$$

$$k \quad, \quad l \quad \in \quad \{-12, -7.5, 0, 7.5, 12\}$$

$$XT_i[kl] = \frac{\sum\limits_{j=1}^{n} \Delta XT_{ij}[kl]}{n}$$

$$(4.26)$$

where,

n is the number of sun and visibility combinations

$R(OLI)_{ijk}$ is the OLI response for a specific defoliated BRDF i and a simulation j for view angle k

$XT_i[kl]$ is the effect of view angle for a signal level i and for view angle difference between k and l

$$\Delta XT_{ij}[\pm 12°] = 100 \frac{|R_{ij}[-12°] - R_{ij}[+12°]|}{R_{ij}[+12°]}$$

$$\forall \quad i \quad \in \quad \{\text{defoliation levels}: 0\%, 10\%, 20\%, 25\%, 30\%, 40\%\} \quad, j \in \{1, 2, \dots, n\}$$

$$XT_i[\pm 12°] = \frac{\sum\limits_{j=1}^{n} \Delta XT_{ij}[\pm 12°]}{n}$$

$$(4.27)$$

where,

n is the number of sun, RSR and visibility combinations

$R_{ij}[+12°]$ is the response for a specific defoliated BRDF i and a simulation j for view angle $+12°$

$XT_i[\pm 12°]$ is the effect of view angle for a signal level i

The effects are also estimated for different data products, such as TOA radiance, TOA reflectance, surface reflectance (ELM-ideal, ELM-typical, BRDF) and their corresponding NDVI products for appropriate comparisons with the forest signal. As in the case with the RSR, the effects for each of the products are transformed to the corresponding NED units.

4.6.3   Visibility effects

The dissimilarity in the orbital parameters of the two sensors results in imaging the same target at different day and time. There is a very small possibility of imaging a specific target of interest on the same day, but the time of acquisition will still be different due to their differences in the equatorial crossing time. The atmospheric conditions can change even within half-an-hour and is dependent on several parameters and visibility is one such parameter that can be modeled easily within DIRSIG (MODTRAN). Since the two sensors are less likely to image the target on the same day, the differences in the atmospheric conditions can introduce a large variation in the sensor reaching radiance. The atmospheric differences can be a major source of error in any change detection application and hence it is important to estimate its effect. The effect is dependent on how different the conditions are and hence, in this research, different combinations of the visibility conditions are used to represent the variation expected between any two dates. The effect of differences in the visibility conditions is studied by analyzing the simulated responses for the two sensors at different solar and view angles. A pictorial representation of the simulation combination is shown in Figure 4.23. The configuration for the sun, view and the visibility conditions are listed in Table 4.5. For example, OLI sensor may image the target at 10 km visibility whereas the MSI sensor may image the same target a few days later when the visibility is 15 km. The effect due to the visibility differences can then be estimated from the relative variation for these sensor responses for different sun and view geometries, as shown in Equation 4.28. One of the issues with this method is that, the estimated effect for visibility inherently includes the effect due to the RSR differences. The effect due to the visibility differences, devoid of the RSR effect, can be estimated from the relative variation between any two different visibility conditions for a specific sensor and for different sun and view geometries. In this case, 50 simulations (10 sun x 5 view) are used for estimating the effect for each sensor and they are combined together to determine the net effect due to the visibility

difference (see Equation 4.29).

$$\Delta VIS_{ij}[k,l] \;=\; 100 \, \frac{\mid R(OLI)_{ij}[k] - R(MSI)_{ij}[l] \mid}{R(OLI)_{ij}[k]}$$

$$\forall \quad i \quad \in \quad \{\text{defoliation levels}: 0\%, 10\%, 20\%, 25\%, 30\%, 40\%\} \quad, j \in \{1, 2, \ldots, n\}$$

$$k \quad, \quad l \quad \in \quad \{7 \text{ km}, \, 10 \text{ km}, \, 15 \text{ km}, \, 20 \text{ km}\} \tag{4.28}$$

$$VIS_i[k,l] \;=\; \frac{\displaystyle\sum_{j=1}^{n} \Delta VIS(RSR)_{ij}[k,l]}{n}$$

where,

n is the number of sun and view combinations

$R(OLI)_{ij}[k,l]$ is the response for a specific defoliated BRDF i using OLI RSR, and a simulation j for visibility conditions k, l

$VIS_i[k,l]$ is the effect of visibility between the two sensors imaged at visibility conditions k, l for a signal level i

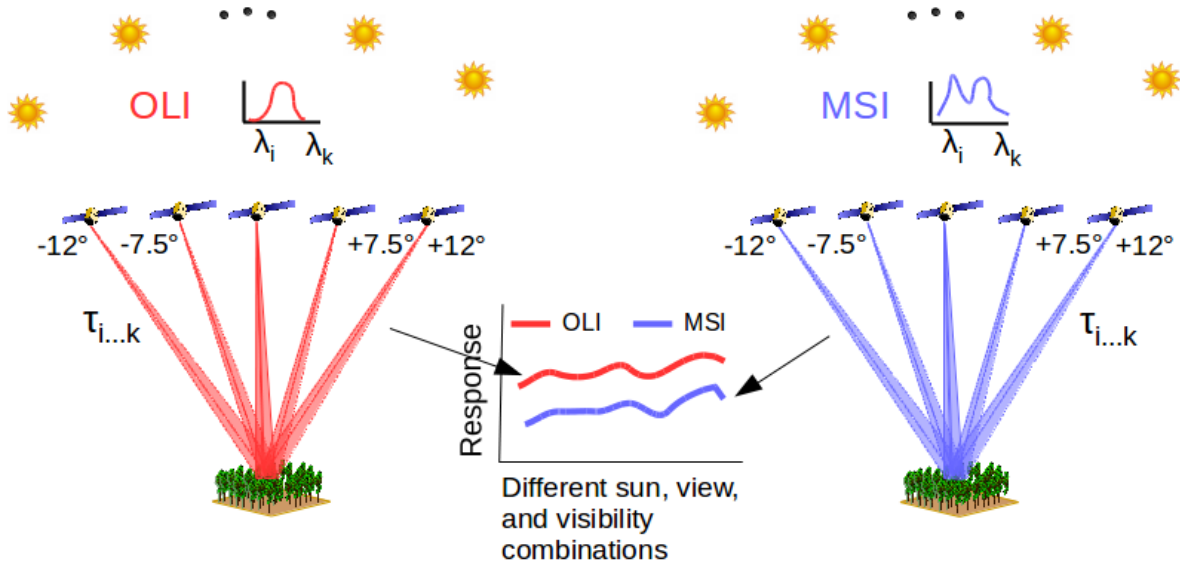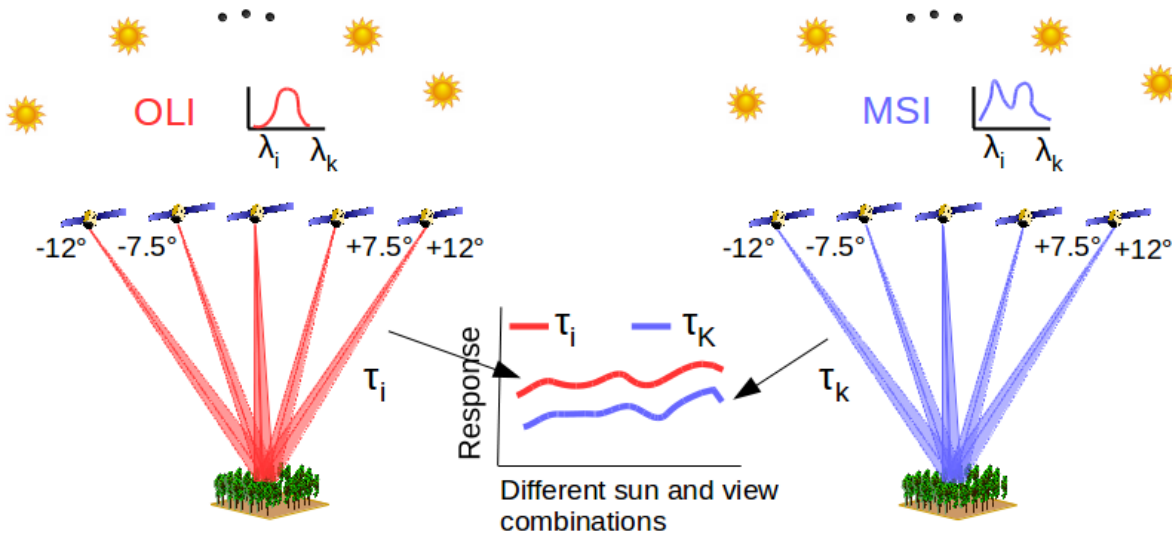
Figure 4.23: Pictorial representation of the DIRSIG simulations for analyzing the visibility effects. For the same surface BRDF, different view angles ($\pm 12°, \pm 7.5°, 0°$), and the sun angles are simulated for two different sensors (OLI,MSI). The ratio of the absolute difference between the two sensors' response at different visibility condition with respect to the OLI's response is estimated as the relative variation for a single simulation. Average over all the simulations is used as an estimate of the visibility effects for a specific difference in visibility. Mean relative variation is also computed for all the signal levels by simulating the same combination with different defoliated BRDFs.

$$\Delta VIS(RSR)_{ij}[k,l] = 100 \frac{\mid R(RSR)_{ij}[k] - R(RSR)_{ij}[l] \mid}{R(RSR)_{ij}[k]}$$

$$\forall \quad i \quad \in \quad \{\text{defoliation levels}: 0\%, 10\%, 20\%, 25\%, 30\%, 40\%\} \quad, \quad j \in \{1, 2, \ldots, n\}$$

$$RSR \quad \in \quad \{ MSI , OLI \} \quad k \quad, l \quad \in \{7 \text{ km}, 10 \text{ km}, 15 \text{ km}, 20 \text{ km}\}$$

$$VIS(RSR)_i[k,l] = \frac{\sum\limits_{j=1}^{n} \Delta VIS(RSR)_{ij}[k,l]}{n}$$

$$VIS_i[k,l] = \frac{VIS(OLI)_i[k,l] + VIS(MSI)_i[k,l]}{2}$$

$$(4.29)$$

where,

n is the number of sun and view combinations

$R(RSR)_{ij}[k,l]$ is the response for a specific defoliated BRDF i, OLI or MSI RSR, and a simulation j for visibility conditions k, l

$VIS_i[k,l]$ is the effect of visibility between two conditions k, l for a signal level i

As in the case with the other factors, the effects are also estimated for other defoliated BRDFs, and for different data products, such as TOA radiance, TOA reflectance, surface reflectance (ELM-ideal, ELM-typical, BRDF), and their corresponding NDVI products. The effects for each of the products are transformed to the corresponding NED units.

In the change detection applications, the large variation observed due to difference in atmospheric conditions is typically compensated by atmospheric compensation methods such as ELM. The two ELM techniques used in this research (ELM-ideal and ELM-typical) were discussed in Section 4.5.2.2. They are used to compensate for the visibility differences in the simulated data. The same process, as described in this section, is used to estimate the residual effects on the compensated data.

4.6.4 Solar zenith effects

Acquisition of images on two different dates or times not only alters the atmospheric conditions, but may also change the illumination conditions. For example, two images acquired within a span of 1 hour in June (summer) over Harvard forest can show as high as 10 degrees difference in

solar zenith angle. The solar azimuth angle can also be different, but typically, forest canopies are assumed to be azimuthally symmetric and hence their effects can be ignored. However, the zenith angle has a direct cosine effect on the incoming irradiance on to the scene. Hence, any difference in the zenith angle affects the sensor reaching radiance directly.

The two sensors are highly likely to image the target at a different day and time and hence, it is important to study the effects due to the solar zenith variations between them. The effect is dependent on the magnitude of the difference in the solar zenith (SZN) angles, the seasons, and the geographic location of the forest. In this research, the effects are estimated for the Harvard forest in the growing season (June - August). The effects can be analyzed in two different ways; based on solar zenith angle differences or from the simulated sensor reaching radiance directly.

In the first method, the differences in the SZN angles are estimated from the difference in the date of acquisition (date offset) between the OLI and MSI sensors over the Harvard forest. The acquisition time for the two sensors differs by about 30 minutes due to their differences in the equatorial crossing time. The formula shown in Appendix A is used to compute the SZN angles for the entire year for both the sensors. For different date offsets (0 to 32 days), their corresponding differences in the solar zenith angles and the ratio of the cosine of the two SZN angles are computed. The ratio indicates the relative variation between the two dates and this is shown in Equation 4.30. The relative variations at different SZN angles and the corresponding difference in the SZN angles are used to estimate the approximate effects expected between the two sensors.

$$SZN_{ij} \; = \; 100 \; \left| \; 1 \; - \; \frac{cos(i)}{cos(i+j)} \right| \quad \forall \;\; i \;\; \in \{20°, 25°, 30°, 35°, 40°, 45°\} \;\;, \;\; j \in \{1°, ..., 10°\}$$

(4.30)

where,

i is one of the ranges of solar zenith angles expected over Harvard forest

j is the increment or the difference between two SZN angles

SZN_{ij} is the relative variation at a specific SZN angle i for a difference in SZN angle of j

In the second method, the analysis is performed using the simulated data for the two sensors at different atmospheric conditions, view angles and solar zenith angles. A pictorial representation of the simulation combination is shown in Figure 4.24. The relative variations are calculated between two different SZN angles under different view and visibility conditions (see Table 4.5

). Several sets of SZN angles, with the same difference in SZN angle, are used to estimate the mean effects from the relative variation. For example, all view and visibility combinations for simulations with SZN=30° and their corresponding combinations for simulations with SZN=35° form a set for 5° difference. Another set could be for the same 5° difference but comes from simulations with SZN=25° and SZN=30°. Average over all the different sets provides the net effect due to 5° difference in SZN angles.

For the same day acquisition over Harvard forest, a difference of 30 minutes between the two sensors (10 AM vs 10:30 AM) can introduce about $2° - 5°$ difference in the SZN angles. The difference could be as high as 9° when the collection between the two sensors are 20 days apart over the Harvard forest. The acquisition of the same target after 20 days is not uncommon since the OLI sensor has a revisit period of 16 days and the MSI sensor in Sentinel-2 has a revisit period of 10 days. Hence, two cases are considered; 5° and 10° difference in SZN angles. The solar angles used for these simulations are shown in Table 4.6. The formula used to estimate the effects for 5° difference is shown in Equation 4.31. In this method, the estimates for each sensor are calculated separately and averaged together. Further, the relative variations are estimated from the corresponding simulations with the same view and visibility conditions. Hence, the effect estimate for SZN in this method is devoid of RSR, visibility or across track effects.

$$\Delta SZN(RSR)_{ij}[k,l] = 100 \frac{\mid R(RSR)_{ij}[k] - R(RSR)_{ij}[l] \mid}{R(RSR)_{ij}[k]}$$

$$\forall \quad i \quad \in \quad \{\text{defoliation levels} : 0\%, 10\%, 20\%, 25\%, 30\%, 40\%\} \quad , \quad j \in \{1, 2, \ldots, n\}$$

$$RSR \quad \in \quad \{ MSI , OLI \}$$

$$k \quad \in \quad \{30°, 30°, 27°, 23°, 33°\} \quad , \quad l \in \{35°, 25°, 33°, 27°, 38°\}$$

$$SZN(RSR)_i[k,l] = \frac{\sum_{j=1}^{n} \Delta SZN(RSR)_{ij}[k,l]}{n}$$

$$SZN_i[k,l] = \frac{SZN(OLI)_i[k,l] + SZN(MSI)_i[k,l]}{2}$$

$$SZN_i[k-l] = \frac{\sum_{k,l=1}^{5} SZN_i[k,l]}{5}$$

$$(4.31)$$

where,

n is the number of visibility and view combinations

$R(RSR)_{ij}[k,l]$ is the response for a specific defoliated BRDF i, OLI or MSI RSR, and a

simulation j for SZN angles (k, l)

$SZN_i[k - l]$ is the effect of SZN for a difference of $(k - l)$ degrees at a signal level i

Table 4.6: Sun angles used to estimate the effect of difference in SZN for $5°, 10°$ angles

$\Delta SZN \approx 5°$			$\Delta SZN \approx 10°$		
SZN , SAZ	SZN , SAZ	ΔSZN_5	SZN , SAZ	SZN , SAZ	ΔSZN_{10}
30 , 137	35 , 145	5	25 , 150	35 , 145	10
30 , 137	25 , 150	5	30 , 137	40 , 150	10
27 , 135	33 , 157	6	27 , 135	38 , 153	11
23 , 132	27 , 135	4	23 , 132	33 , 157	10
33 , 157	38 , 153	5	20 , 150	30 , 137	10
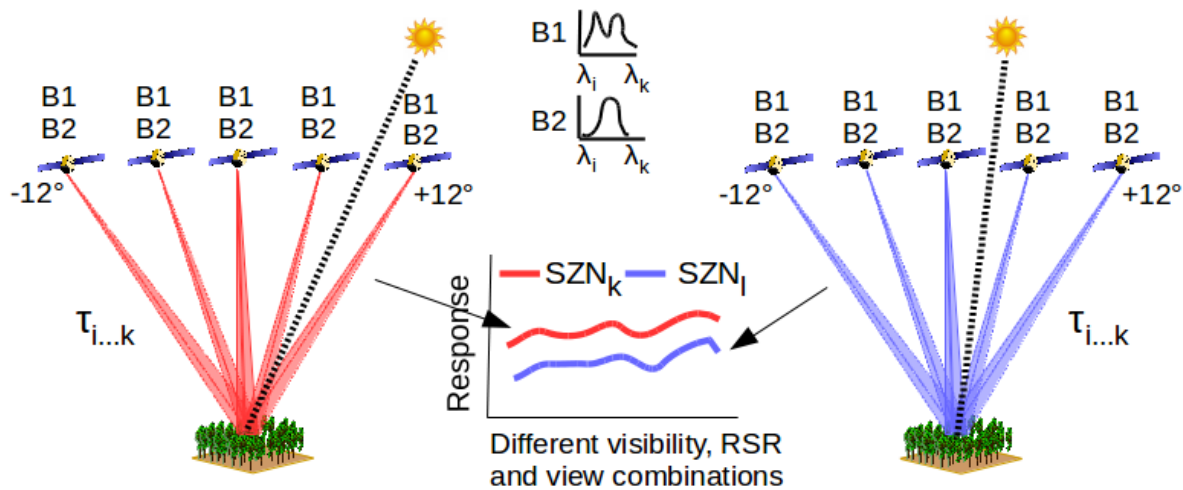
Figure 4.24: Pictorial representation of the DIRSIG simulations for analyzing the SZN angle effects. For the same surface BRDF, different view angles and the visibility conditions are simulated for two different sensors (OLI,MSI). The ratio of the absolute difference between the sensor's response at different SZN angles with respect to the response at one of the SZN angles is estimated as the relative variation for a single simulation. Average over all the simulations is used as an estimate of the SZN effects.

In this case, 20 simulations (4 vis x 5 view) are used for estimating the effect for each sensor and they are combined together to determine the net effect for a specific set of SZN angles. The SZN effect for 5° difference is calculated from the average of all the 5 sets of SZN angles. The effect will be different for each set depending on the SZN angle used. Further, in one of the sets, the difference is 6° and in the other, the difference is 4°. These may introduce some differences across all the 5 sets and it will affect the STD of estimated effects. Using both the

mean and the STD of the effect will help to assess the actual impact due to the differences in the SZN conditions.

As in the case with the other factors, the effects are also estimated for other defoliated BRDFs, and for different data products, such as TOA radiance, TOA reflectance, surface reflectance (ELM-ideal, ELM-typical, BRDF), and their corresponding NDVI products. The effects for each of the products are transformed to the corresponding NED units.

In the change detection applications, variations expected due to the differences in the illumination conditions are typically compensated by adjusting for the cosine effect of the irradiance. The TOA and the surface reflectance products are compensated for these effects. However, the differences in the illumination geometry can introduce residual effects due to the BRDF of the forest canopy. The magnitude of this effect can be inferred from the SZN effects of the reflectance data products.

4.7 Simulated vs Real data

It is important to validate the simulated data to gain more confidence in the analysis of the factors' effects. An ideal method of validation is to compare the real data with the simulated data using the same sensor and environmental conditions, but it is very difficult to reproduce the exact conditions. Therefore, reasonable assumptions and approximations are used where accurate sensor or environmental conditions are unavailable. In this research, the simulated BRDF of the Harvard forest is compared with the surface reflectance calculated from the MODIS and Landsat 8 products.

4.7.1 Reflectance comparison

The MODIS BRDF products rely on atmospherically corrected, cloud-cleared MODIS data measured over 16-day periods to generate the RossLi BRDF model coefficients. The methods used to generate these coefficients are found in MODIS Algorithm Theoretical Basis Document (A. H. Strahler et al., 1999). The MODIS BRDF product provides the model coefficients and it can be used to compare the reflectance for a given view and illumination geometry. The MODIS BRDF products from July and August (July 4 - 19 and July 28 - August 12) of 2015

are used to compare the simulated BRDF of the Harvard forest, as the OLI sensor collected two cloud-free scenes about the same dates. Since the MODIS BRDF is based on multi-date (16 days) observations, changes on the ground during this 16-day period may be lost. The two scenes from Landsat 8 are 23 days apart and may be useful for observing the changes during this period. However, the two scenes are collected from two different passes (WRS path 12 and 13) and so have different illumination and view geometry. Although the atmospheric conditions may also be different on the two days, atmospheric compensated products can be generated by USGS EROS data center. The Landsat 8 and the MODIS products used in the analysis were processed by USGS EROS data center and accessed from EarthExplorer (USGS, 2016).

Vermote et al. (2016) discusses the approach used in the atmospheric compensation of the Landsat 8 products. The atmospheric compensated data are provided in 16-bit unsigned integer format, which are converted to reflectance using the scaling coefficients provided in the metadata file (Surface Reflectance Product Guide, 2016). The surface reflectance is then determined from the image pixel corresponding to the Harvard forest in the atmospheric compensated products for the two dates. The surface reflectance for each date is considered as a BRDF measurement since the illumination and the view geometries are different. The view and the sun angles for the two dates are required for proper comparison of the simulated canopy BRDF with the measured BRDF.

The SZN and the SAZ angles can be calculated for the Harvard's geographic location using the Equation in Appendix A for the specific dates. The error in the estimation of the SZN and the SAZ angles are smaller than a degree. The computed solar angles were compared with the solar angles from the metadata file from the OLI surface reflectance product and was found to be within a fraction of a degree. The canopy reflectance from the simulated BRDF model varies smoothly and any effect due to an error of less than a degree in the illumination angle is negligible.

The view geometry is not provided in the metadata file for a specific pixel, but can be determined to a reasonable approximation. Knowing the Harvard forest's sample coordinate in the image, the number of samples in the across track direction, and the FOV of the OLI sensor ($\pm 7.5°$), the VZN angle can be computed as shown in Equation 4.32. Any error in estimating the VZN angle by this method is smaller than a degree and its effect in estimating the reflectance from the canopy BRDF is negligible. The OLI sensor's VAZ angle is estimated for the Harvard forest

using the Systems Tool Kit (STK) software (STK, 2016). The STK tool uses the Two-Line Elements (TLE) of the Landsat 8, and the ground coordinates of the Harvard forest to estimate the VAZ angle. In this case, the two scene acquisition dates are used in the STK tool to estimate the corresponding VAZ angles. The error in the estimation of VAZ is smaller than a fraction of a degree and its contribution to the canopy reflectance is assumed to be negligible.

After estimating the sun and the view geometries, the virtual Harvard forest model is validated by comparing the BRDF reflectance from the Landsat 8 products with the corresponding BRDF reflectance estimated from the MODIS BRDF products and the simulated BRDF model.

$$VZN \approx (Sample_{start} - Sample_{Harvard}) * \frac{(7.5 - (-7.5))}{Total_{samples}} - 7.5 \qquad (4.32)$$

where,

$Sample_{start}$ is the left most non-zero image sample number of the across track line that pass through the Harvard forest

$Sample_{Harvard}$ is the image sample number of the Harvard forest

$Total_{samples}$ is the number of samples in the across track line that pass through the Harvard forest

4.7.2   Effect analysis

Similar to the validation of the forest model, it is important to assess the change expected due to the sensor and the environmental factors. The change observed in the real data is due to the combination of the factors' effects and the actual changes on the ground, and hence an ideal comparison with the simulated data is extremely difficult, if not impossible. However, the contribution due to other factors can be estimated using the simulated data for the same illumination and view conditions as observed in the real data and this can be used to estimate the contribution of the effects in the real data.

The Landsat 8 scenes for the two acquisition dates, as mentioned earlier, are used to assess the effects of the factors discussed in Section 4.6. No Sentinel-2 data over the Harvard forest are available and therefore, contribution due to the RSR differences could not be studied. Secondly, the atmospheric visibility conditions for the two dates are not known and could not be estimated

accurately, so the analysis is performed for the following two cases. In the first case, the visibility for both the dates are assumed to be 20 km and in the other case, the visibility of the scene acquired in August is assumed to be about 15 km. These two cases provides an approximate range of variations expected due to the visibility differences. Since the view and the sun angles are known for the Landsat 8 data, the same angles are used to simulate the sensor reaching response.

In the previous analysis, the exact location of the Harvard forest was used. In this case, three ROIs (about 200 pixels each) are chosen to analyze the effect for different types of forest. Because the three ROIs are apart from each other ($> 5km$), the distribution and the types of the trees will be different, and as a result, their canopy BRDF will be different. Thus, using multiple ROI provides an opportunity to evaluate the changes in the Harvard forest at different locations. This is further extended by analyzing 3 more ROIs from Loyalsock state forest in Pennsylvania. Loyalsock forest is also a deciduous forest site with similar types of tree species as Harvard, but their tree characteristics and distributions are completely different. Similar to the Harvard scenes, the OLI sensor acquired two cloud-free scenes in late May and early June (May 29, June 5) of 2015 from two different passes (WRS path 15, 16). Although the chosen ROIs were cloud-free for Loyalsock forest, there were clouds in the vicinity, which may impact the results. Nevertheless, for lack of a better dataset, the Loyalsock forest is also used to estimate the observed changes on the ground.

For the Loyalsock and the Harvard forests, the ROIs are selected from the Level 1 and the surface reflectance products, processed by the USGS EROS data center. The Level 1 (TOA radiance and TOA reflectance) products provided by USGS consists of quantized and calibrated scaled Digital Number (DN) in 16-bit unsigned integer format, representing the multispectral image data acquired by the OLI and TIRS sensors. The DNs are scaled to the TOA radiance and TOA reflectance using the radiometric scaling coefficients provided in the metadata file (LANDSAT 8 (L8) DATA USERS HANDBOOK, 2016) for the OLI sensor's red and NIR spectral bands. Since the Level 1 and the surface reflectance products are accurately geo-referenced, the same ROIs are selected for both the acquisition dates and across all the product types. This ensures that the estimated radiance and reflectance for each ROI is consistent and comes from the same region on the ground. The mean reflectance (and radiance) for each ROI and the mean of all the ROIs for each forest are used to evaluate the relative changes observed in the real data.

In the DIRSIG tool, the view and the solar geometry from the real data is used to simulate the sensor response for the OLI sensor's RSR. The anisotropic properties of the ground are modeled using the Harvard canopy BRDF model. The atmospheric visibility for the Landsat 8 scenes are unavailable, and so the sensor responses are simulated assuming two visibility conditions (15 km and 20 km). The relative variations are calculated to estimate the effect due to the visibility conditions and also due to the combined effect of the SZN and the view angle differences. Similar to the factors' effect analysis in Section 4.6, the relative variations are computed for the TOA radiance, TOA reflectance, surface reflectance and the NDVI products. The relative variation estimated from the simulated data is then compared with the real data to evaluate the contribution of the factors' effects to the actual changes observed in the real data.

4.8   Factor screening experimental design

The factorial designs are widely used in experiments involving several factors to study the joint effect of the factors on a response. One of the important objectives of this research is to identify the significant factors that affect the sensor response. There are many factors that impact the sensor observed radiance, but some of the factor's effects may be insignificant in comparison to the other factors. Further, the complexity of the factorial experiment increases with more factors, and if not modeled properly, can lead to unreliable results. Hence, it is necessary to perform a factor screening experiment to identify the small number of significant factors that can be extensively studied using a more complex design (regression analysis). This section details the approach used to identify and screen the factors using 2^{k-p} fractional factorial experimental design. A detailed description for this type of design was provided in Section 3.5.3.

The typical procedure for setting up a 2^{k-p} design is as follows:

- Identify the factors

- Determine the number of experiments

- Choose a design with appropriate aliasing

The first task is to identify the factors. For the screening experiment, all the factors are identified ahead of time and the analysis is based on fixed factor effects. The number of experimental

runs required for the factor screening experiments are usually preselected based on the practical constraints. The number of experimental runs and the factors involved are used to setup the aliasing structure for the remaining factors. The aliasing structure is chosen such that the resolution of the design is as large as possible. The experimental design is analyzed based on the response variable, which is dependent on the statement of the problem. In this case, the objective is to identify those factors that affect the radiance measured by the sensor and hence, radiance observed by the sensor is used as the response variable. The task of determining the fixed factor levels and the appropriate design for screening experiment is discussed below.

4.8.1 Fixed factors

The radiance observed by a sensor is dependent on many factors, each of which can introduce variability at different levels. The factor screening experiment's main objective is to determine a few of the most significant factors. Although, there are many factors that could affect the response, only seven factors are considered for factor screening. The factors and their levels are listed in Table 4.7. The rationale behind the choice of levels for each of the factors is briefly described below.

Table 4.7: Factors and their levels used in the factor screening experiments

Factor Name	Factor	Levels	
		Low (-1)	High (+1)
RSR	A	RSR - OLI	RSR - MSI
Along track angle	B	$-2°$	$+2°$
Across track angle	C	$-15°$	$+15°$
Atmosphere visibility	D	5 KM	50 KM
Ground Slope	E	$-25°$	$+25°$
Sun azimuth angle	F	$90°$	$180°$
Sun zenith angle	G	$1°$	$50°$

4.8.1.1 Sensor factors

OLI RSR and MSI RSR :

The RSR for the OLI sensor is different from the MSI sensor (Section 3.4), and they both deviate from an ideal RSR (rectangular function). This introduces a difference in radiance as

observed by the sensor. The variabilities between sensor RSRs are captured by a factor with two levels. The prelaunch measured RSRs from the MSI and OLI sensors for one of the detectors is chosen to represent the two levels of this factor. There is also variability within a sensor as each detector will have different RSRs, but they are expected to be very small.

Line Of Sight (LOS) angles:

The detectors in these sensors have varying LOS angles to cover the entire ground swath in the across track direction. A sensor measures the radiance from the ground in different spectral regions with its focal plane module oriented in the across and along track direction (Figure 3.14). The forest BRDF can introduce differences to the radiance reaching the sensor as a function of LOS angles. The LOS angle in the along-track direction is much smaller than the across-track direction, and they are considered as two different factors. The two levels for each of these factors are determined from the design of the sensors. For the two-level experiments, the levels are chosen as widely spaced as possible, to reduce any errors with extrapolation in the design space. Since both the MSI and OLI sensors have LOS angles less than $\pm 2°$ in the along track and less than $\pm 11°$ in the across track directions, the levels are chosen as $(-2°, +2°)$ and $(-15°, +15°)$ for along and across track direction respectively.

4.8.1.2 Atmospheric factor

The atmosphere scatters both, the incoming radiation from the sun and the outgoing radiance from the target towards the sensor. There are many parameters that can be used to model the atmosphere in MODTRAN. However, for the screening experiment, only one of the parameters (visibility) is chosen. The assumption is that the variability caused by the atmosphere can be sufficiently expressed within the two levels using visibility criteria. While performing the complex design with multi-factor and multi-levels, more levels of the visibility are used.

The low visibility condition due to thick aerosol or fog can be useful to represent the diffusivity in the sky, which increases the upwelled radiance reaching the sensor. The high visibility condition (clear sky) can be used to approximate a near ideal day for sensor measurements. The two levels chosen for visibility are 5 KM and 50 KM.

4.8.1.3 Ground terrain factor

The terrain on the ground can change the apparent LOS or the view angles observed by the sensor with respect to the ground normal. Forests are typically found on hill slopes with varying gradient. The slope of the terrain can be oriented in any direction with respect to the sun and the sensor. The maximum variation is observed in the radiance reaching the sensor when the terrain is oriented towards and away from the sun.

The sensor observes the forest scene, modeled as BRDF, under varying factor levels in the factor screening experiments. The variation in the slope, in essence, affects the BRDF of the scene and hence, the terrain changes are represented by changes in the BRDF of the scene. Section 4.4.2.1 briefly discussed the approach used to generate the BRDF of a forest on a sloped terrain. The two levels representing the ground slope are obtained by generating two representative BRDFs when the slope of the ground is rotated by $+25°$ and $-25°$ about the $+Y$ axis.

4.8.1.4 Sun angle factors

The sun's position is described by its zenith and the azimuth angles. The position of the sun relative to the ground affects the incoming irradiance and also affects the apparent reflectance of the forest scene. For example, high solar zenith angle reduces the irradiance by the cosine of that angle, and also causes long shadows from the trees, affecting the amount of illumination in the forest.

The morning equatorial crossing time of the two sensors limit the azimuth angle of the sun to less than $180°$ angle. For most locations in the northern hemisphere, the azimuth variation throughout the year is much smaller than $180°$. For the factor screening experiments, the two levels are spaced as far away as possible, and the effects are assumed linear between the two levels. Although, the choice of $0°$ and $180°$ are valid, the observed response to these angles will be equal, since the forest scene is assumed to be azimuthally symmetric. The choice of $90°$ for the low level ensures that the response is dissimilar between the high and low levels, and the effects can be assumed to be linear within this range. Further, choosing the azimuth angles as $90°$ and $180°$ allows for maximum difference in response between these angles, when the ground is rotated about the $+Y$ axis.

The maximum zenith angle theoretically possible is $90°$, when the sun is illuminating the scene from the horizon but the remote sensing community is interested in observing the forest reflected energy when the sun zenith angle is low. Hence, the high level for the zenith angle is chosen as $50°$. The lowest possible zenith angle is $0°$ when the sun is directly above the scene. At this angle, the azimuth of the sun is undefined. In the experimental runs, setting zenith angle to $0°$ may cause a redundant observation, and can affect the estimated residuals and potentially impact the orthogonality of the design. Hence, the low level for zenith angle is set to $1°$. Any extrapolation of the factor effects to $0°$ zenith should still be valid, as the reduction in irradiance is less than $\frac{1}{1000}$th of the irradiance from the nadir sun.

4.8.2 2^{k-p} experiment design

The experimental design parameter 'p' in 2^{k-p} design is chosen based on the number of experimental runs. The number of experiments are chosen such that the main effects and two-way interaction effects are free and clear of low order aliasing. For k factors, there are k main effects and $\frac{k(k-1)}{2}$ two-way interactions. The degrees of freedom needed to satisfy the above criteria is given in Equation 4.33.

$$
\begin{aligned}
df &= 1 + k + \frac{k(k-1)}{2} \\
df &= 1 + 7 + \frac{7(7-1)}{2} = 29
\end{aligned}
\tag{4.33}
$$

There is a need for 29 experiments and the design that is nearest to it is 2^{7-2} which requires 32 experiments. The degrees of freedom will be greater than required and hence the parameter of fractionation 'p' can be 2 for this case. However, the design resolution is only IV and is not a nodal design, i.e., the design does not provide the highest resolution with the least number of experiments for the given factors. The nodal design with resolution greater than IV for seven factors requires 64 experiments ($p = 1$). The resolution for this design is VII, and the two-factor interaction effects are not aliased even with a three-factor interactions in this design. Typically, the resolution chosen for the factor screening experiments are either III or IV, but in this case, resolution greater than IV is chosen as it ensures no aliasing between the two-factor interaction effects. The design parameters, defining equations and generators are shown in

Table 4.8 and the design matrix table is shown in Appendix D.

Table 4.8: The parameters and properties of the 2^{7-1} fractional factorial design

Factors	Runs	Resolution	Fraction	Replicates	Center pts	Block
7	64	VII	$\frac{1}{2}$	1	0	1

Defining Relations :	I = ABCDEFG
Generators :	G = ABCDEF

Typically, the experimental design requires that the response is a random variable, and the experiments conducted are completely randomized, replicated and blocked to satisfy the basic principles of the DOE techniques. In this research, the experiments are conducted using DIRSIG as the experimental engine.

In reality, many random factors affect the geometry (structure, distribution, etc.) and the radiometry (mature, young, infected, etc.) of the forest, thereby affecting the radiance reaching the sensor. These disturbances cause randomness in the observed radiance. However, in DIRSIG, the scene is modeled as BRDF and no such random factors can be introduced directly, making DIRSIG a deterministic tool.

For this reason, the randomness will be introduced in the forest BRDF as it would approximate reality. The approach explained in Section 4.4.2.2 is used to generate a large number of BRDF for each terrain slope. Since the forest BRDF is different for each simulation, the sensor reaching radiance will also be different. These differences will cause the response to be a random variable.

For the factor screening experiment, there are 64 experimental runs, and hence 32 random BRDFs are generated for each ground slope. Each experimental run chooses one of the 64 BRDFs to represent a forest scene in DIRSIG. This approach ensures that the response is a random variable. Further, the 64 runs can also be performed for non-random BRDFs (e.g center of the forest BRDF). Repeating the experiment with and without random BRDFs provides a measure of consistency in identifying the significant factors.

The blocking for experiments is performed to isolate the effects due to variation in the experimental conditions or remove any specific effects due to nuisance factors which are

typically small. Since all the simulations are performed in DIRSIG, there are no changes in the experimental conditions and any nuisance factor in the experiments are uncontrolled and unknown. In general, any effect due to these nuisance factors are not analyzed in a factor screening experiments. Hence, blocking is not necessary for this design.

Replicates are useful to assess the lack of fit for the model. It provides a measure of true error and is useful to avoid the possibility of fitting the measurements to noise. However, the deterministic aspect of DIRSIG limits the possibility of running a replicated design. Further, the replicates for factor screening experiments are seldom performed and are avoided by spacing the two levels as far apart possible. This requirement was considered in this design as stated earlier. Hence no replicates are used for the factor screening experiment design.

4.8.2.1 DIRSIG experiment set-up

The experiments are conducted by simulating each factor combination, as shown in the design matrix table (refer to Appendix D), using the levels shown in Table 4.7.

Platform :

A camera with a single detector is used for measuring the radiance reaching the sensor. The LOS angles (for modeling view angles) and the RSR for the detector are determined based on the experimental configuration from the design matrix. An appropriate number of sub-samples within the detector are used to get an average radiance observed by the sensor. The spectral samples are chosen with a width of 5 nm. The reflectance of the forest is predominant in the NIR and SWIR bands, but the vegetation index in forest application often uses the NIR and red spectral bands. Since the response variable is dependent on the chosen spectral bands, the analysis is performed for all the spectral bands (VIS-NIR-SWIR) independently. The corresponding RSRs were used for the detector.

Atmosphere:

The visibility parameter for the atmosphere is updated in the tape5 file used by MODTRAN. The default mid-latitude summer atmospheric profile with rural aerosol is used with two different visibility conditions. The file used to define the atmosphere in DIRSIG is also used to define the position of the sun. The visibility, sun azimuth and sun zenith angles are updated based on the experimental run in the design matrix.

Platform motion and tasking:

For these experiments, the platform and the sensor is at a fixed position (700 KM altitude) and oriented for a nadir-view of the scene. The camera with a single detector is instantly exposed since there is no platform motion. The view angle effects are simulated by changing the LOS angles of the detector for a static platform.

Scene :

The earth is represented in the scene geometry by constructing a sphere with an associated material file. This is equivalent to draping the entire globe with a single BRDF model. The material file will use a RossLi BRDF model to represent the optical properties of the forest. For each experimental run in the design matrix, corresponding BRDF model (according to the slope of the terrain, random or non-random) is used in the material file.

Scripts are used to setup the DIRSIG configuration files for each experimental runs and to perform the simulation in DIRSIG. The procedure for analyzing these designs were discussed in the section 3.5.2. The analysis for the factor screening experiment is performed using Minitab (Minitab, 2015) software package and the results are shown in Section 5.8.

4.9   Regression analysis

While the previous section discussed the methods used to screen the important factors, in this section the use of regression analysis to understand the relative importance of the factors will be explored. In general, the regression model is used to understand the functional relationship as the true functional form of the response variable is not known. In this research, the factors that are found to be significant from the factor screening experiment are used to construct the regression model. Similar to the factor screening experiment, the at-sensor radiance is considered as the response variable in the regression model. Most of the forest canopy studies are performed using the red and the NIR spectral bands, so these responses are modeled independently.

Typically, transformations for both the factors and the response variables are performed to simplify the regression model. The factors' levels are transformed depending on the functional relationship of the factor to the response variable. For example, the at-sensor radiance is dependent on the logarithm of the visibility and so the factor levels are transformed to the logarithm units. But, for factors whose functional relationships are unknown, the transformation

that makes the main effect linear, are useful in simplifying the regression model. These models with linear main effect can provide a good overall summary measure and such models are likely to provide meaningful results, even when there are interactions between the factors. In this research, the following functional variables, as appropriate, are attempted: regular variable, variable in radians, cosine, square, product of cosine and sin, and natural logarithm. The transformed variable whose main effect plot shows a linear trend is chosen for the regression model. Similar to the factors, the response variable can also be transformed to reduce the complexity of the regression model. In this case, two different response variables are used: the at-sensor response and the natural logarithm of the at-sensor response. The regression analysis is performed independently for these two response variables. Since an ideal functional relationship may not be possible, polynomial regression analysis is used in this research.

The polynomial regression analysis methods are useful when the transformations cannot linearize the relation between the response and the predictors. In this type of regression, higher order polynomials and the interaction terms are used as regressors. One of the issues with this technique is that, the powers of the same variable could be highly correlated (non-orthogonal) and this could make it difficult to understand how the predictors impact the response. But in an orthogonal design, there are no correlations and therefore the main and the interaction effects can be estimated independently. The problem of correlation due to the powers of the same variables can be solved by using orthogonal polynomial contrasts.

The orthogonal polynomial contrasts are used to evaluate the polynomial trends in the response when the designed experiments involve quantitative factor levels. This is an extension of the two-level design where the coding ($\pm$ 1) is used to model the quantitative factors. The coefficients from the polynomial contrasts are used to partition the sum of squares (SS) in to linear, quadratic, cubic, etc. contributions (Kuehl, 2000). The highest order and the number of contrast coefficients are dependent on the number of levels of the factor. For example, a factor with 5 levels has 4 degrees of freedom (5-1=4), hence the contrast coefficients can be estimated for linear, quadratic, cubic and quartic terms. If the levels are equally spaced, the contrast coefficients are available in statistical textbooks, however coefficients for unequally spaced contrasts are not readily available. In such cases, Proc IML of Statistical Analysis System (SAS) can be used to generate the coefficients. As there will be more than one factor in the regression model, polynomial coefficients are estimated for each factor independently. In this research, the orthogonal polynomial coefficients are estimated using the example script provided

by Michael Galyean (2016). The orthogonal polynomial contrasts though makes the contrasts for linear and higher order polynomial terms orthogonal, the contrasts for the interaction terms between the variables may not be orthogonal if it is a fractional factorial design. The impact of this non-orthogonality is assumed to be negligible, as the number of experimental runs in the model estimation are almost as high as the full-factorial runs.

The orthogonal polynomial contrasts are estimated for the transformed levels of each factor. These contrasts are then used to perform the regression analysis. The number of terms (regressors) are dependent on the number of factors, their levels and the order of the polynomial terms used in modeling the response variable. The number of terms for k factors with n degree of a polynomial can be calculated from the multinomial sum (Feller, 1968). For example, let us assume there are 4 factors with 5 levels each, and a 4th order polynomial is used in the regression model with all the possible interaction terms (up to degree 4). The number of terms in this case is given by $\binom{4+1-1}{1} + \binom{4+2-1}{2} + \binom{4+3-1}{3} + \binom{4+4-1}{4} = 69$ terms (for degree = 1,2,3,4). The number of experimental runs in a full factorial design in this case is $5^4 = 625$ runs. Therefore, a fractional factorial design is sufficient to estimate the model coefficients using the least squares approach. For the real-world datasets, the number of observations are usually limited, but in this research, the DIRSIG tool is used to simulate the response and so, all the observations required to run a full-factorial experiment are collected. The experimental setup in DIRSIG and the configuration files are similar to those used for the factor screening experiment, except that a subset of those factors are used to simulate the observations. A fractional factorial design is used to construct the model and the remaining observations are used to validate the accuracy of the model. This is achieved by dividing all the experimental runs into training and validation datasets. The training datasets are further sub-divided into multiple sets or K-folds depending on the level of fraction. For example, for 5^4 observations, a $\frac{1}{5}^{th}$ fraction result in 5 independent sets (A,B,C,D,E) of 125 observations each. One of these (E) is assigned as a validation dataset and the remaining observations are used for training the model. The training set can be divided into 4-folds (A,B,C,D) with 125 observations in each fold. From the four fold, one of them is used as a test set and the remaining as model sets. Thus, four different models can be generated using four modeling sets (ABC, ABD, ACD, BCD) and are verified using the corresponding test set. Then, the four models are combined to build a single regression model, which are evaluated using the validation dataset (E). Although there are more observations than the parameters even in a fractional design, small number of regressors are preferred because of its simplicity

to explain and understand. This is achieved using stepwise regression method which can sift through large number of independent variables to identify a few but required variables.

Stepwise regression is a process of constructing a model by successively adding (forward selection) or removing (backward elimination) the variables or a combination of both based on some criteria. In this research, forward selection method is used which involves the following steps: a) start with no variables in the model, b) test the addition of each variable using a criteria, c) add that variable that improves the model the most, and d) repeat this process for all the terms. The $adjR^2$ is used as the criteria to test and select the best variable at each step. The process is continued till all the terms are included in the model and the summary statistics such as the RMS, order of variable selection, etc. are stored for analysis. The actual number of terms are estimated using the cross-validation approach.

Cross-validation (CV) is a way of measuring the predictive performance of a statistical model. Although the model fit statistics like R^2 may act as a good guide on the model prediction, they may tend to over-fit the model, especially when there are large degrees of freedom in the model. In the CV approach, the predictive accuracy of a model is typically measured by the mean squared error on the test dataset (Hastie et al., 2009) as it is likely to be larger than the training dataset which is used to estimate the model. In the K-fold CV method, the data is divided into K equal parts. For the ith part in K (i=1,..,K), a model is fit to the other (K-1) parts and the CV error, as shown in Equation 4.34, is computed for the ith part of the data. This gives the ith prediction error for a fitted model and the procedure is repeated to get all the other CV errors. The K CV errors are then combined together to estimate its mean and the standard deviation from which the standard errors are computed.

$$
\begin{aligned}
CV_K(\theta) &= \frac{1}{n_k} \sum_{i \in F_k} (y_i - \hat{f}_\theta^{-k}(x_i))^2 \\
CV(\theta) &= \frac{1}{k} \sum_{k=1}^{K} CV_K(\theta) \\
SD(\theta) &= \sqrt{\frac{\sum\limits_{k=1}^{K} (CV_K(\theta) - CV(\theta))^2}{K-1}} \\
SE(\theta) &= \frac{SD(\theta)}{\sqrt{K}}
\end{aligned}
\tag{4.34}
$$

where,

n_k is the number of data points in the kth fold

F_k is a set of all the index in the kth fold

$\hat{f}_\theta^{-k}(x_i)$ denotes the fitted function computed with the kth part of the data removed

y_i is the ith observation of the kth fold

K is the number of parts in the K-fold validation

θ is the tuning parameter

The purpose of calculating the CV error is to estimate the best tuning parameter ($\hat{\theta}$), which in our case, is to find the required number of terms in the regression model. This is determined by calculating the number of terms required in the model ($\hat{\theta}$) for which the CV error is minimum ($CV(\hat{\theta})$), as shown in Equation 4.35. Then, the one standard error rule is applied wherein, the simplest model whose error is within one standard error of the minimal error is considered as the optimal model (Hastie et al., 2009).

$$
\begin{aligned}
\hat{\theta} &= \underset{\theta \in \{\theta_1, \ldots \theta_m\}}{\operatorname{argmin}} \ CV(\theta) \\
CV(\bar{\theta}) &\leq CV(\hat{\theta}) + SE(\hat{\theta})
\end{aligned}
\tag{4.35}
$$

where,

$\hat{\theta}$ is the parameter at the minimal CV error

$CV(\bar{\theta})$ is the optimal CV error

$\bar{\theta}$ is the optimal model

The number of parameters used in the optimal model is of more interest than the optimal model itself. So, the forward selection regression method is used again to fit the best regression model, but the forward progression is stopped when it reaches the number of terms estimated from above. In this modeling, all the observations in the training dataset are used and the model is evaluated using the validation dataset. To maintain the hierarchy principle in the model, all the excluded lower order terms whose higher order polynomials were selected, are also included in the final model.

One of the objectives of this research is to determine the relative significance of the factors. This is determined by using the sum of squares (SS) calculated as part of the regression fit. The relative contribution of each term in the regression model is calculated by taking the ratio of the SS of each term to the SS explained by the model. The relative contribution of a factor is

determined by aggregating all the higher order polynomials and the interaction terms of that factor. In the case of interaction terms, the relative contributions are equally divided between the factors. The relative contributions can also be separated into first order and higher order contributions for each factor to assess the contribution of the non-linear terms to the model. Finally, the effects due to the interaction terms can be evaluated from the values in the relative contribution table and also from the interaction plots.

4.10   Summary

The purpose of this chapter is to explain the methodology required to achieve the objectives defined in Chapter 2. The chapter began with the qualitative techniques used to validate DIRSIG radsolver algorithms. This was followed by the approach used for a quantitative comparison against the published RT models, which provides further assurance to the accuracy of the DIRSIG model. The detailed approach to the modeling of forest canopy in the virtual environment was presented in Section 4.2. The sensitivity study on DIRSIG, sampling strategy and the tools necessary to measure the BRDF of the virtual forest scene were explained in Section 4.3. Section 4.4 addressed the approach used for modeling the BRDF measurements to the RossLi canopy BRDF model. This section also discussed the techniques that were used to generate auxiliary BRDF models. The approach used in the characterization of the forest defoliation in a simulated environment was discussed in Section 4.5. Section 4.6 explained the method used to perform the sensitivity analysis of the factors in terms of the changes on the ground. The method used to test the accuracy and the validity of the simulated data in representing a real forest was discussed in Section 4.7. This was followed by Section 4.8, which described the strategy used to screen the factors for its significance using the fractional factorial design. One of the main objectives of this research is to identify the relative significance of the factors. The regression analysis technique used to accomplish this objective was discussed in Section 4.9.

The next Chapter will present the results and provide the discussion for different strategies proposed in this chapter. These results include validation of DIRSIG, modeling of forest scene, BRDF measurements and model sensitivities, forest characterization and the evaluation of the factors' effects using one factor at a time technique, the comparison of the real and simulated data and finally, the use of DOE techniques to estimate the relative significance of the factors.

Chapter 5

Results

This chapter aims to summarize the results. In Chapter 4, the methodology used in the validation of DIRSIG was detailed, and the results are presented in Section 5.1. The consistency of DIRSIG with other RT models enable us to use DIRSIG to model and measure the reflectance for forest canopies. The synthetic forest scene, modeled based on the practical approach mentioned in Chapter 4, displays the closeness to the real-world scene. The results for the simulated forest are shown in Section 5.2. Section 5.3 presents the results of the sensitivity analysis. Section 5.4 discusses the results attained in fitting the BRDF measurements to the RossLi BRDF model. This section also discusses the results of the auxiliary BRDF models generated for different terrain slopes. The modeling of forest defoliation and its characterization are explained in Section 5.5. One of the main objectives of this research is to determine the effects of different factors, and the results of this analysis is shown in Section 5.6. The results from the comparison of the simulated BRDF with the reflectance measured by the OLI sensor are summarized in Section 5.7. This is followed by the discussion of results from the factor screening experiment in Section 5.8. Lastly, the results from the statistical analysis, performed to estimate the relative significance of the different factors are detailed in Section 5.9.

5.1 DIRSIG validation

The results for qualitative and quantitative validation are shown in this section.

5.1.1 Qualitative analysis

The approach used for qualitative verification of the Geodensity radsolver was presented in Section 4.1.1. The DIRSIG simulations of the trees are performed using the Geodensity radsolver for all the tree elements. The simulations employ the scaled irradiance approach as shown in Equation 4.4, to generate images in reflectance factor units. Figure 5.1a shows tree A which is at the center of the image and Figure 5.1b shows the same tree model used to represent tree B, but with an offset in its position. The two trees with tree A casting its shadow on tree B are shown in Figure 5.1c. The three images are captured in an identical viewing and solar condition, with the only difference being their relative offset. These images can be used to illustrate the effect of shadows.

(a)   (b)   (c)
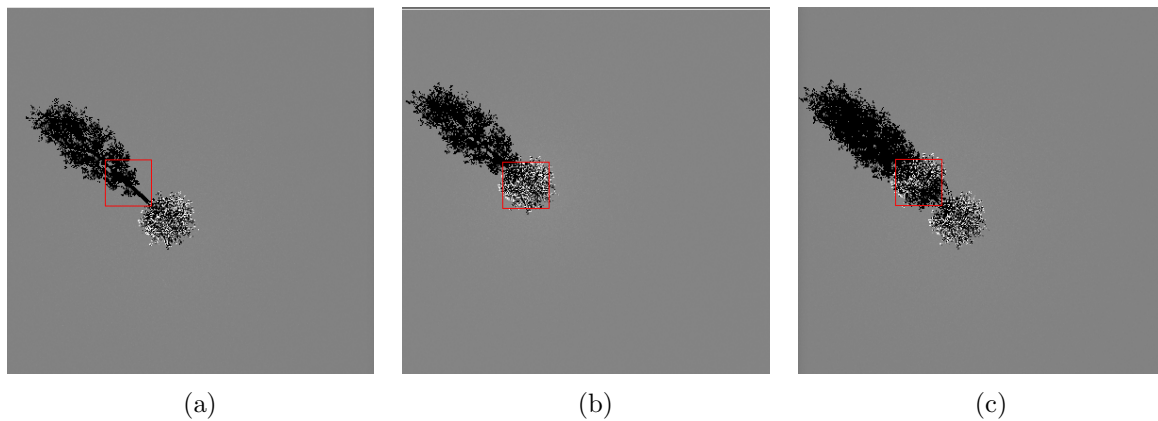
Figure 5.1: Qualitative analysis of shadow using two trees; tree A in Figure (a) and tree B in Figure (b). Figures (a), (b) show a single tree illuminated by the sun. The tree A casts its shadow on tree B in Figure (c). Their relative positions are highlighted by a red square.

An ROI can be used to compare the reflectance from the two similar trees. For example, shadow pixels of tree A in Figure 5.2b are compared with the corresponding pixels over tree B in Figure 5.2c. The difference in reflectance between the shadow and non-shadow pixels is shown in Table 5.1.

Table 5.1: The apparent reflectance difference between shadowed tree and illuminated tree

Illuminated pixels (Figure 5.2b)	Shadow pixels (Figure 5.2c)
22.9 %	18.8 %

(a)   (b)   (c)
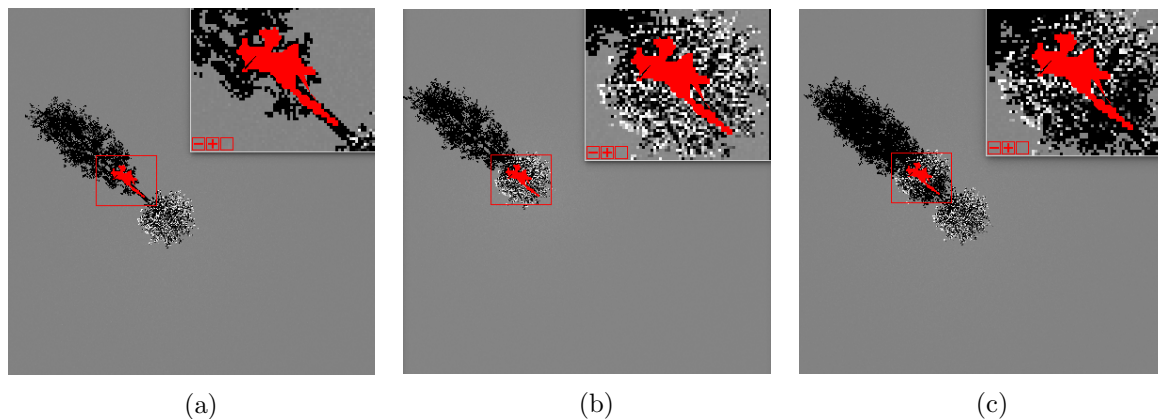
Figure 5.2: Qualitative analysis to show the shadow effect. Figure (a) shows the shadow pixels (red) selected using ROI. The average reflectance of the corresponding pixels (red) in Figure (b) and Figure (c) are compared to verify the performance of the radsolver.

The average reflectance of the shadowed pixels is lower (18.8%) than the non-shadowed pixels (22.9%). This reduction is attributed to the shadows. If there were no illumination over the shadowed pixels, the reflectance observed should be near 0%. However, the average of the shadowed pixels indicate much higher reflectance. This is owing to the scattering of photons from adjacent leaves, which acts as a secondary source of illumination.

The multiple scattering effect tends to increase the secondary illumination to both the shadow and non-shadow pixels. The additional illumination to non-shadow pixels would increase its apparent reflectance. This is illustrated in Figure 5.3 and in Table 5.2. In this case, a similar approach is used to identify a ROI for tree B that are not under the shadow of tree A. The apparent reflectance for the ROI in Figure 5.3c is higher than Figure 5.3b. This small increase in reflectance is attributed to an increase in the illumination caused by multiple scattering effect from tree A.

Table 5.2: The apparent reflectance difference between single and two trees due to multiple scattering effect

Single tree B (non-shadow)	Two trees A,B (non-shadow)
(Figure 5.3b)	(Figure 5.3c)
37.4 %	38.1 %

(a) (b) (c)
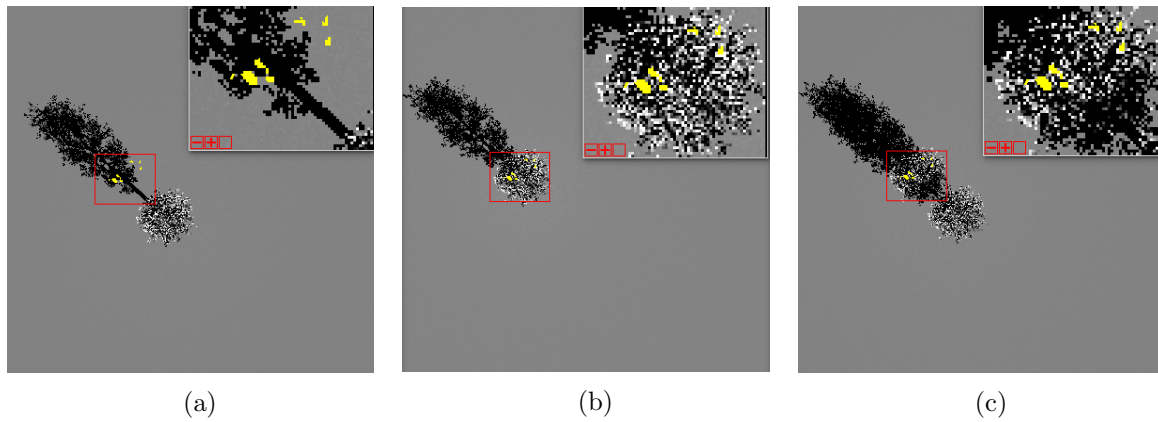
Figure 5.3: Qualitative analysis to show the multiple scattering effects. The ROI of pixels (yellow) in tree B are illuminated by sun and not under the shadow of tree A. The average reflectance of the corresponding pixels (yellow) in Figure (b) and Figure (c) are compared to verify the performance of the radsolver.

The other approach used for qualitative validation compares the reflectance observed from a typical forest canopy to the DIRSIG simulated results. For this evaluation, the forest scene is modeled based on the approach explained in Section 4.2. The simulated image of the modeled forest scene is shown in Figure 5.4. A nadir viewing camera captured the image when the sun was located at 30° zenith and 130° azimuth.
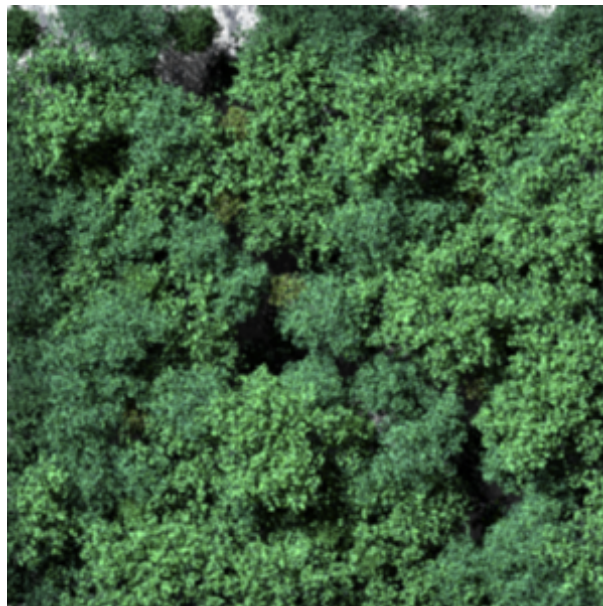
Figure 5.4: Simulation of Harvard forest scene modeled in DIRSIG using OnyxTree. A nadir viewing camera captured the image in RGB bands. The sun is located at 30° zenith and 130° azimuth.

The validation is performed for three different illumination and viewing conditions. In the first case, the camera and the sun are placed at nadir. In the second case, the camera is at nadir,

while sun is positioned off-nadir. In the third case, the camera and the sun lie along the principal plane, and the hot-spot effect is observed by the camera when the sun is off-nadir. In all three cases, the NIR spectral band ($\lambda = 0.865$) is chosen for evaluating the reflectance. The DIRSIG simulation of the three conditions as stated above are shown in Figure 5.5.

(a)   (b)   (c)
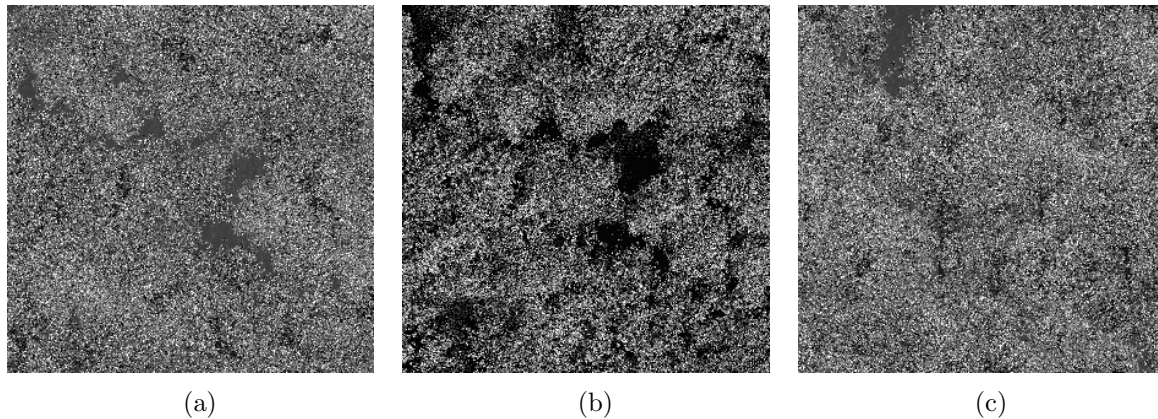
Figure 5.5: Evaluation of total reflectance from the forest canopy for three different cases when imaged in NIR spectral band ($\lambda = 0.865\mu m$). Figure (a) shows the simulation when the sun and the camera view are at nadir. Figure (b) shows the simulated image when the sun is at $30°$ zenith and $130°$ azimuth and view is at nadir. Figure (c) show the simulated image when the sun and the camera are in the principal plane with $30°$ zenith to observe the hot-spot effect.

The reflectance observed by the image in Figure 5.5a is found to be 59%. This is a unique case where the sun and the camera are at nadir position. It is a special case of hot-spot effect, since the backscatter and forward scatter are in the same direction. The reflectance observed in this case is much higher than the reflectance of any of the tree species used to model the forest scene. The higher reflectance is attributed to the back scattering (stacking effect) from the canopy.

The reflectance for Figure 5.5b is found to be 38%. In this case, the sun is off-nadir with a $30°$ zenith angle, and the image is acquired by a nadir viewing camera. The reflectance observed in this case is very similar to what is typically expected from the forest in this spectral region. For a similar illumination zenith angle and NIR spectral band, the reflectance found from the MODIS BRDF product and Landsat reflectance product varied from 34% to 40% over the Harvard forest.

The reflectance for Figure 5.5c is found to be 58%. Similar to the first case, the hot-spot effect is observed but from an off-nadir angle. The reflectance is very similar to the first case, as expected and the high reflectance is attributed to the back scattering characteristics observed in the forest canopy.

From the two tree analysis for multiple scattering, it is qualitatively shown that the Geodensity radsolver in DIRSIG reproduces shadows and multiple scattering effects correctly. The canopy reflectance analysis has shown that the reflectance estimated by DIRSIG matches closely to the expected canopy reflectance. The rudimentary analysis from these two approaches rule out any gross error with the radsolver, but the accuracy cannot be verified without a quantitative analysis. The DIRSIG validation with the RAMI III dataset will provide more assurance that the DIRSIG model is consistent with other radiative transfer algorithms.

5.1.2 Quantitative analysis

The quantitative analysis of the radsolver algorithm using the RAMI III dataset is performed for two different experimental conditions: the homogeneous and heterogeneous experiments. The methodology used for the validation was presented in Section 4.1.2. The results for the two experiments are shown in this section.

5.1.2.1 RAMI Homogeneous experiments

The homogeneous experiments are validated for the red and NIR spectral bands for total scattering (single + multiple) BRF along the principal plane. There were two different leaf arrangements in the homogeneous experiment. The result for planophile leaf arrangement is shown in Figures 5.6 and 5.7. For these experiments, the zenith angle is negative when the view and the sun have the same azimuth. This is indicated by the hot-spot effect in the negative zenith direction.

(a) Red (SZN=20)

(b) Red (SZN=50)

(c) NIR (SZN=20)

(d) NIR (SZN=50)
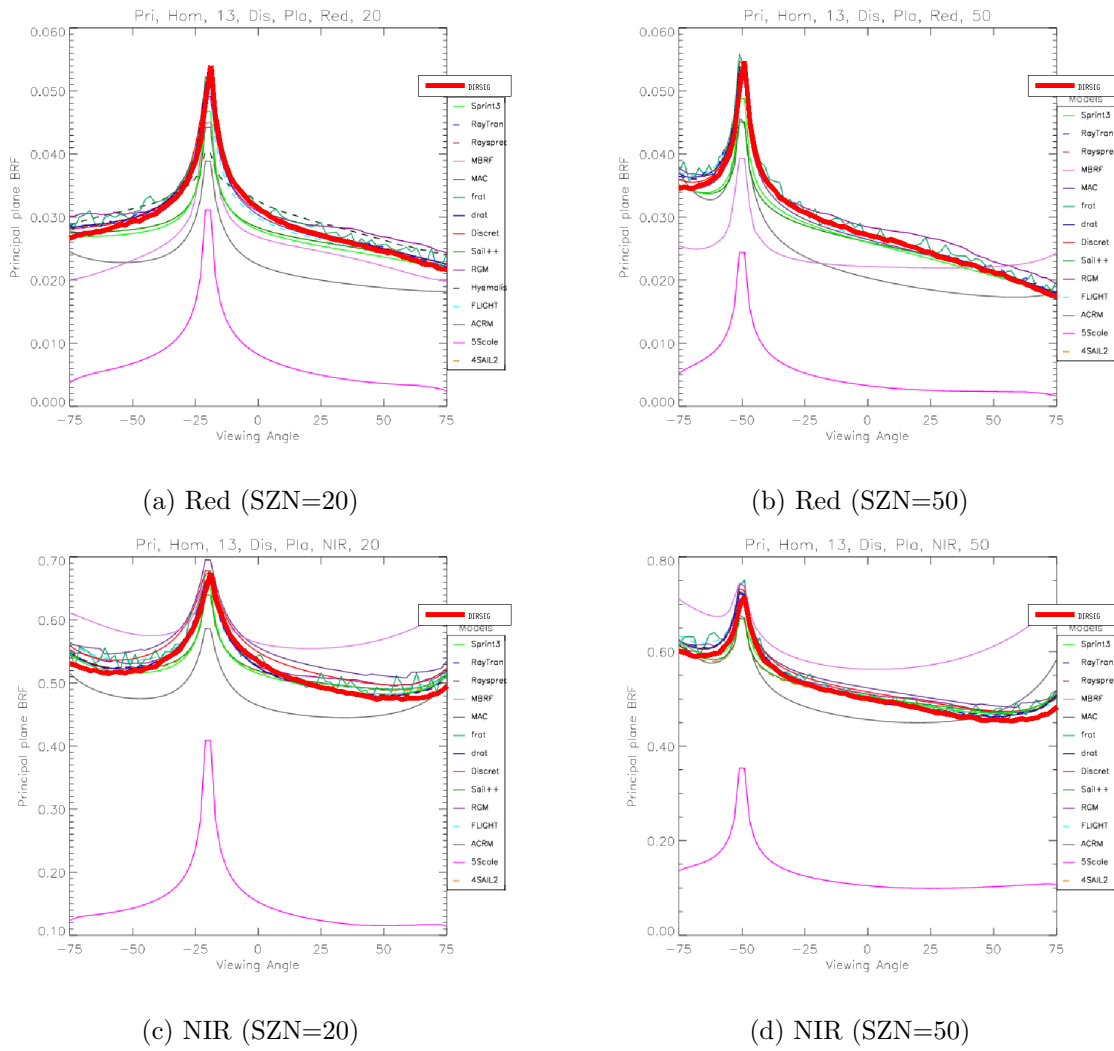
Figure 5.6: Principal plane BRF (total scattering) for Homogeneous experiments with planophile leaf arrangement. The DIRSIG BRF results (red) are over-plotted on the RAMI published results (RAMI, 2015). Figure (a) and (b) shows the BRF results for red spectral band when the sun is at 20° and 50° zenith angle respectively. Figure (c) and (d) shows similar BRF results, but for NIR spectral band.

(a) Red (SZN=20) (b) Red (SZN=50)

(c) NIR (SZN=20) (d) NIR (SZN=50)
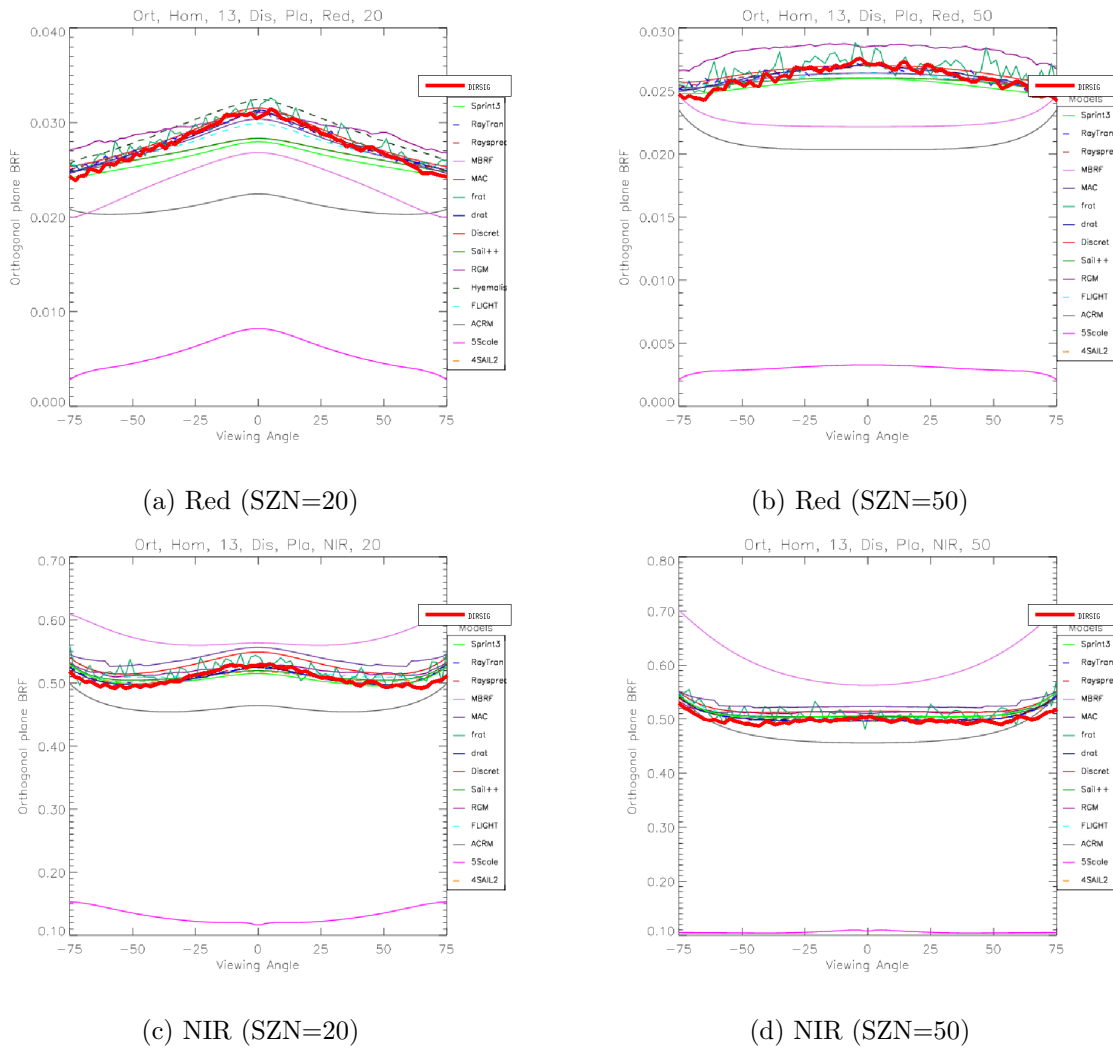
Figure 5.7: Cross plane BRF (total scattering) for Homogeneous experiments with planophile leaf arrangement. The DIRSIG BRF results (red) are over-plotted on the RAMI published results (RAMI, 2015).Figure (a) and (b) shows the BRF results for red spectral band when the sun is at 20° and 50° zenith angle respectively. Figure (c) and (d) shows similar BRF results, but for NIR spectral band.

From Figure 5.6, it is clearly evident that the BRF results from DIRSIG are consistent with most of the RT models in both the red and NIR spectral region.

5.1.2.2 RAMI Heterogeneous experiments

The heterogeneous experiments are validated only for the "Real Zoom-In" case. The results for the Het04a1 test case is shown in Figure 5.8. Similar to the homogeneous experiments, the zenith angle is negative when the view and the sun have the same azimuth angle. The DIRSIG results for the test case are over-plotted on the published results to show their consistency

against other RT models. The nine different test cases (c1-c9) for the 3x3 grid of the central region (b5) of the "Real Zoom-In" experiments are shown in Figure 5.9. From Figures 5.8 and 5.9, it is evident that the DIRSIG results are consistent with many of the published RT models for the heterogeneous experiment.
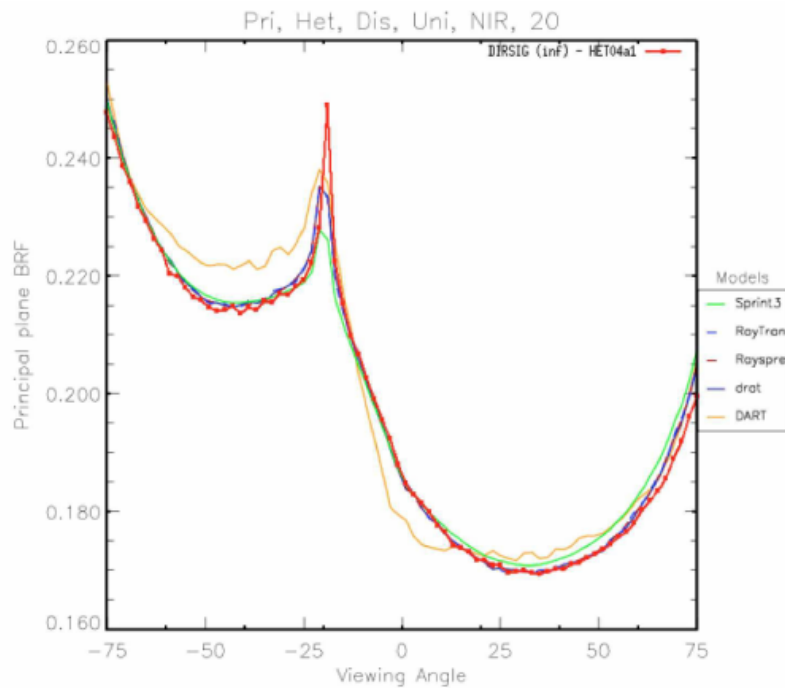
Figure 5.8: Principal plane BRF (total scattering) of the entire scene for the "Real Zoom-in" heterogeneous experiment. The DIRSIG BRF results (red) are over-plotted on the RAMI published results (Goodenough and Brown, 2015, RAMI, 2015) for the test case HET04a1.

A small increase in the hot-spot reflectance is observed in DIRSIG results in comparison with other RT models. This is attributed to the fine sampling of the zenith angles in DIRSIG. The RT model results seem to show coarser zenith resolution with a linear interpolation for the intermediate reflectance. Further, the uncertainity at the hot-spot region between the published RT models are high and it is unknown as to which models are accurate. A small but noticeable noise is observed in the DIRSIG results and it is assumed to be an outcome of the reduced photon sampling and/or undersampling of the detector. This is not observed in the homogeneous experiment results shown in Figure 5.6, where the scene extent is small (25 m x 25 m). The scene extent for heterogeneous experiments is large (270 m x 270 m), and it would lead to a substantial increase in processing time if higher sub-samples and photon bundles were to be used. The increase in the processing time to achieve higher fidelity for large scenes is one

of the main reasons that the size of the virtual Harvard forest constructed in DIRSIG is limited to 50 m x 50 m.
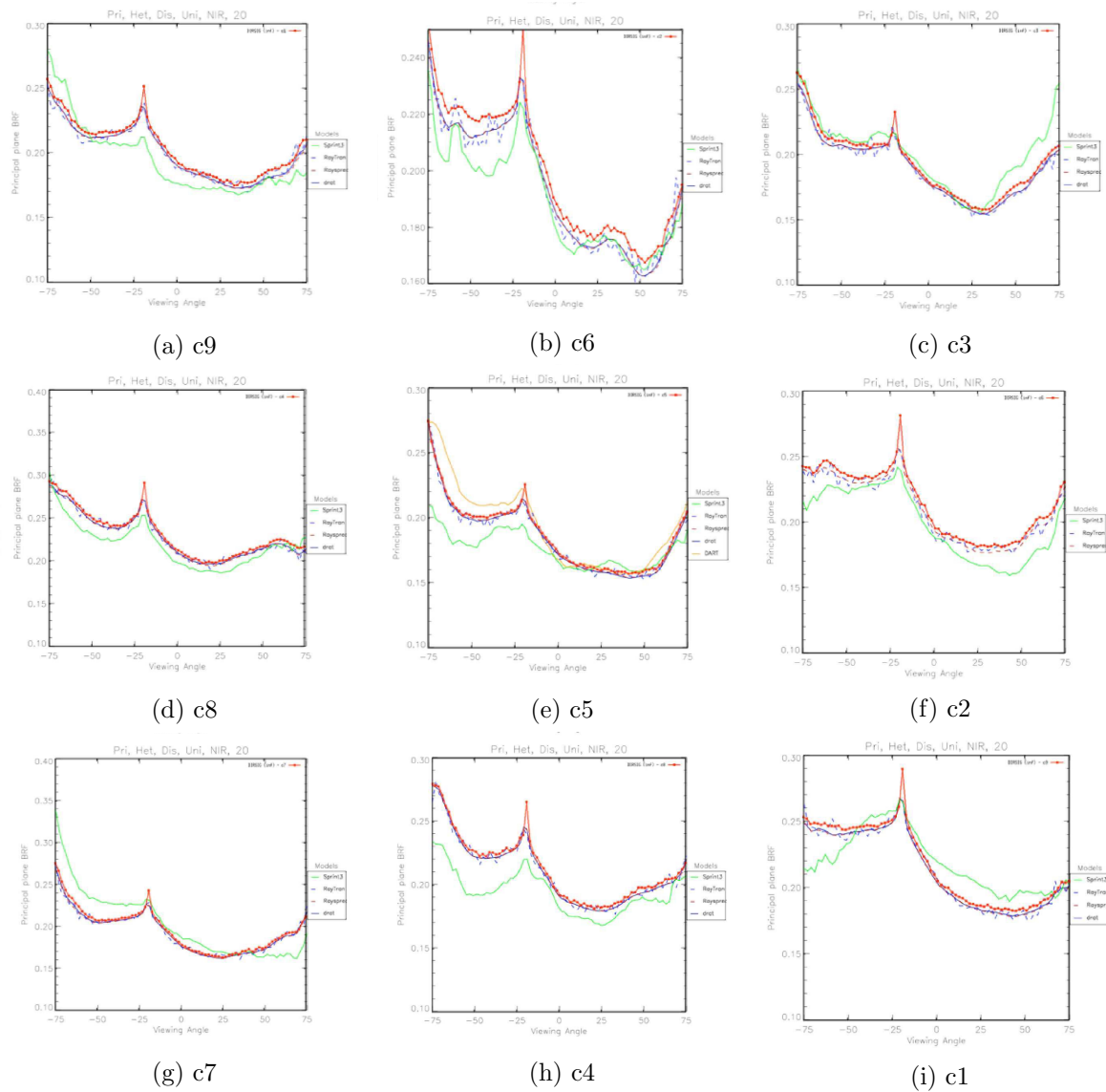
Figure 5.9: Principal plane BRF (total scattering) for test case HET04c1 - HET04c9. The DIRSIG BRF results (red) are over-plotted on the RAMI published results (Goodenough and Brown, 2015, RAMI, 2015)

The quantitative validation using RAMI III homogeneous and heterogeneous experiments indicate that the DIRSIG modeling of the forest canopy produces BRF that are consistent with most of the RT models used in RAMI studies. It is important to note that many of the models used in the RAMI studies were developed specifically for forests. But DIRSIG is a general model that incorporates radiometric, atmospheric and sensor models that are not specific to a type of scene. The consistency between the DIRSIG model and other canopy radiative transfer models

indicate that all the other DIRSIG proven capabilities can also be used over the simulated forest canopy. The qualitative and quantitative validation results have provided enough confidence in using DIRSIG for accurate modeling of forest canopy BRDF.

5.2 Canopy geometry and optical properties

The tree geometry is derived from OnyxTree based on the methodology explained in Section 4.2.2. The tree model from the OnyxTree tool provides an option to export the tree geometry elements such as branches, trunk, leaf, etc. accordingly to the level of complexity needed to represent the tree. The trees used in the Harvard virtual scene consists of facetized tree geometries for trunk, primary branches, secondary branches, tertiary branches, bough and leaves. The geometry for twigs are not included since their contribution to canopy reflectance is negligible. Ten unique tree models are generated using the OnyxTree tool such that the modeled trees match the LAI from the forest inventory. These tree models are shown in Figures 5.10 and 5.11. The color and the relative size of the trees shown in Figure 5.10 are chosen for illustration purposes. The true color of the modeled trees in the VIS-NIR-SWIR region is dependent on the associated tree and ground optical properties. The actual tree model parameters determined from the OnyxTree tool and DIRSIG simulations (LAI), are shown in Table 5.3.

Table 5.3: Parameters for ten unique tree models used to represent the Harvard forest canopy. LAI is calculated from DIRSIG. The number of trees replicated to simulate the forest canopy is shown in the last column.

Tree Name	Dominant/ Co-dominant	LAI	Height	Crown width	Crown height	# of trees
BlackBirch (BB)	D	5.12	20.03	5.59	12.21	10
PaperBirch (PB)	C	4.63	19.77	3.1	12.85	16
RedMaple 1 (RM-1)	C	2.65	10.57	5	6.65	12
RedMple 2 (RM-2)	C	2.93	23.5	7.3	13.63	13
RedMaple 3 (RM-3)	C	2.24	23.73	8.84	13.6	23
RedMaple 4 (RM-4)	D	5.38	23	5.8	13.57	25
RedOak 1 (RO-1)	D	2.66	21.96	6.24	8.78	18
RedOak 2 (RO-2)	D	3.93	23.15	11.66	15.97	29
RedOak 3 (RO-3)	D	3.21	30.13	14.24	9.94	15
YellowBirch (YB)	C	5.28	10.81	4.31	6.48	5

The geometry of the tree is coupled with its trunk and leaf optical properties in DIRSIG to represent a tree in the virtual environment. The ground is modeled as a flat plane and the ground litter spectra, shown in Figure 5.14, is used to model the texture of the ground based on a grass texture map. The field measured leaf reflectance spectra is inverted to get the transmittance spectra using the PROSPECT inversion tool. The inversion for a Red Oak leaf spectra is shown in Figure 5.12. It is observed from Figure 5.12 that the PROSPECT inverted reflectance spectra (red) is in good agreement with the field measured reflectance spectra (blue). The RMS between the measured and PROSPECT inverted spectra is found to be 0.005 ($< 0.5\%$ in reflectance units). The noise in the field measured spectra near $2.5\mu m$ is due to low SNR, but is smooth in the PROSPECT inverted spectra as the PROSPECT model uses a modeling approach to derive its reflectance spectra. Note that the spectral bands of the OLI and MSI sensors are limited to $2.3\mu m$ and noise in the data near $2.5\mu m$ is not an issue. The transmittance spectra is shown as a green curve in Figure 5.12. The inversion tool is used to generate the transmittance and reflectance spectra for other leaf species, and their results also match the field spectra accurately. The spectral reflectance and transmittance for all the leaf and trunk species, used for modeling the optical properties of the Harvard forest scene are shown in Figure 5.13. The trunk reflectance spectra is used for all the branches and boughs. The leaf reflectance spectra for black birch is used for yellow birch leaf as they were not collected in the field. This should not impact the canopy reflectance since the number of yellow birch trees in the site is small.

(a) BB   (b) PB   (c) RM-1   (d) RM-2   (e) RM-3

(f) RM-4   (g) RO-1   (h) RO-2   (i) RO-3   (j) YB
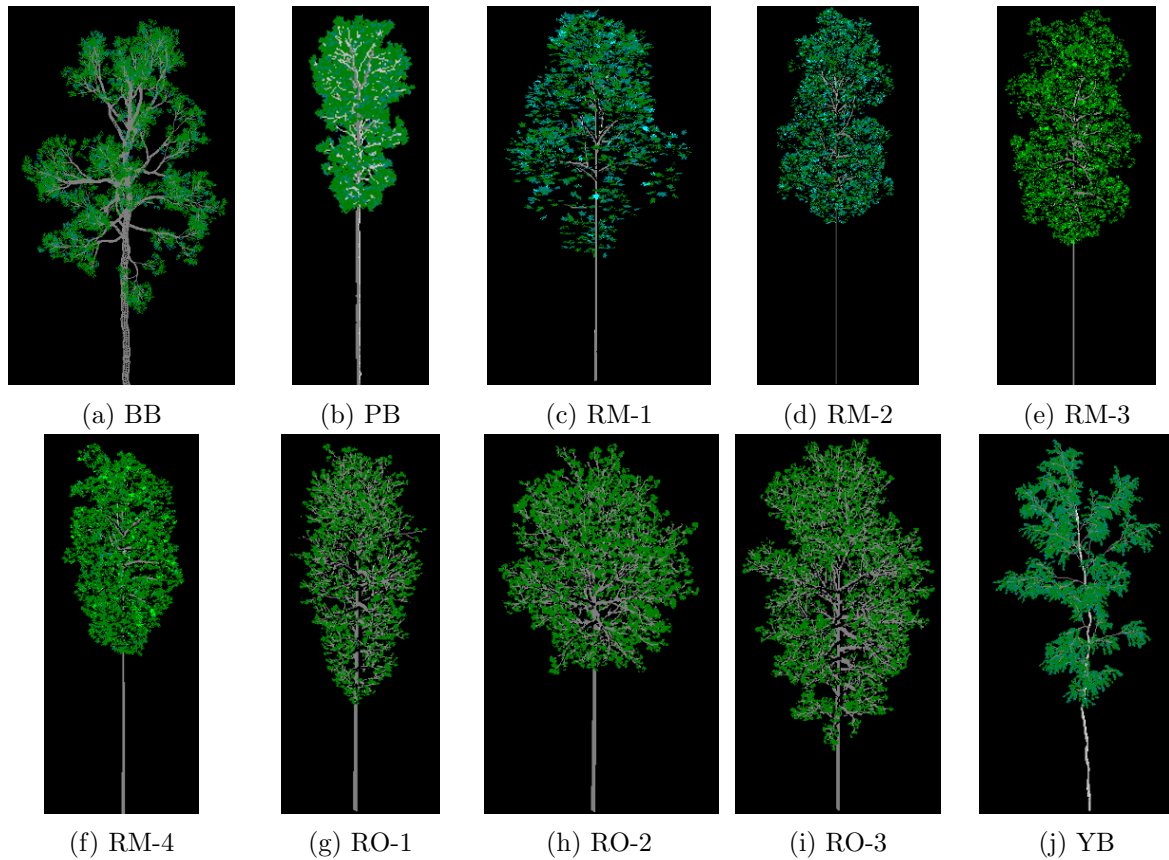
Figure 5.10: Side-view of ten unique trees modeled using OnyxTree based on the Harvard forest inventory data. The color of the leaves Figures (a)-(j) are chosen only for illustration. The optical properties for these tree elements affects the true color of the trees.
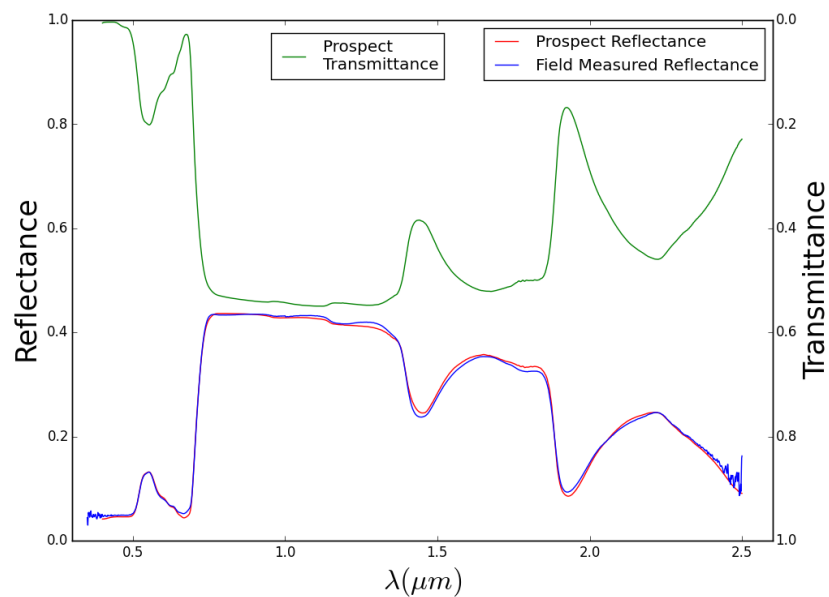
Figure 5.12: Inversion of Red Oak leaf reflectance spectra to generate reflectance and transmittance spectra using PROSPECT. The field measured spectra (blue) is in good agreement with PROSPECT derived reflectance spectra (red). The field measured spectra was collected in Sep 2013.

(a) BB

(b) PB

(c) RM-1

(d) RM-2

(e) RM-3

(f) RM-4

(g) RO-1

(h) RO-2

(i) RO-3

(j) YB
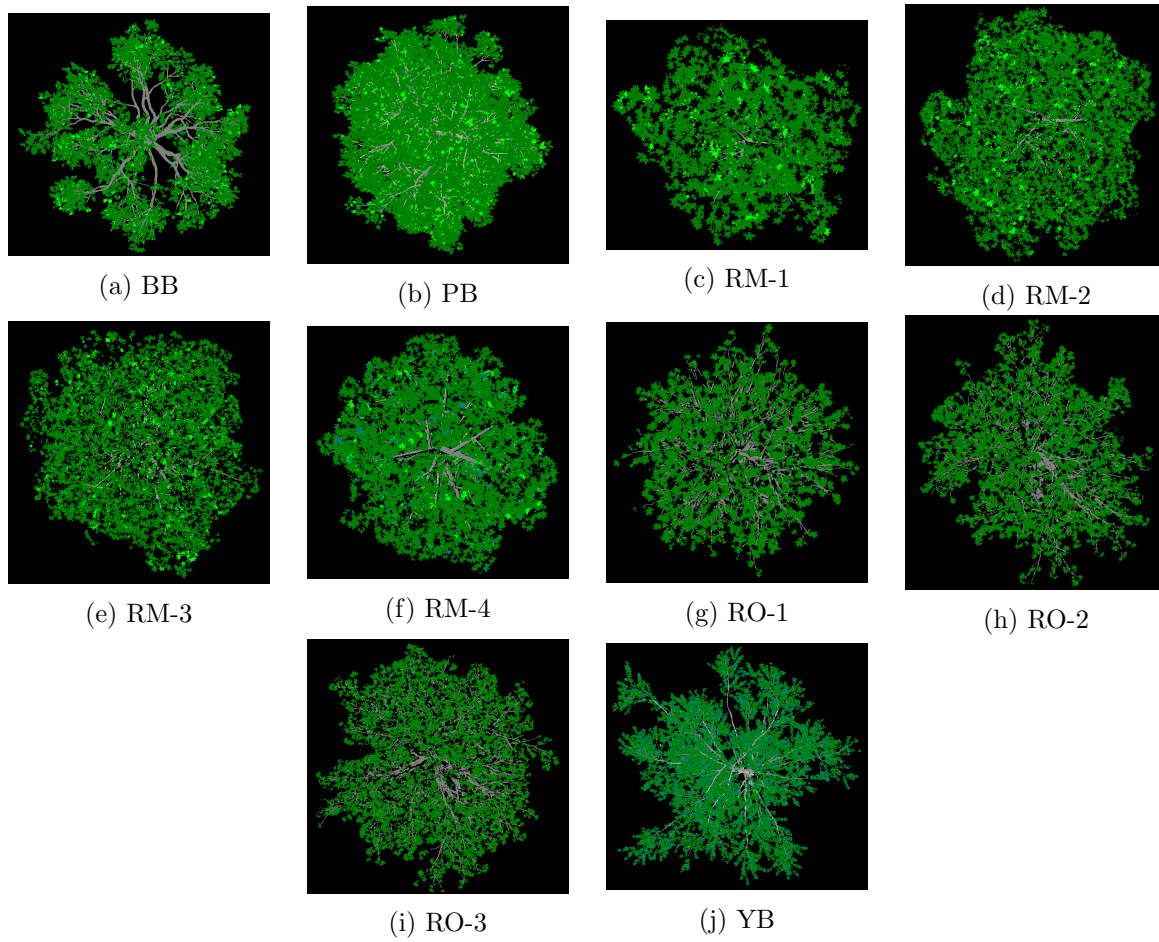
Figure 5.11: Top-view of ten unique trees modeled using OnyxTree based on the Harvard forest inventory data. The color of the leaves Figures (a)-(j) are chosen for illustration.
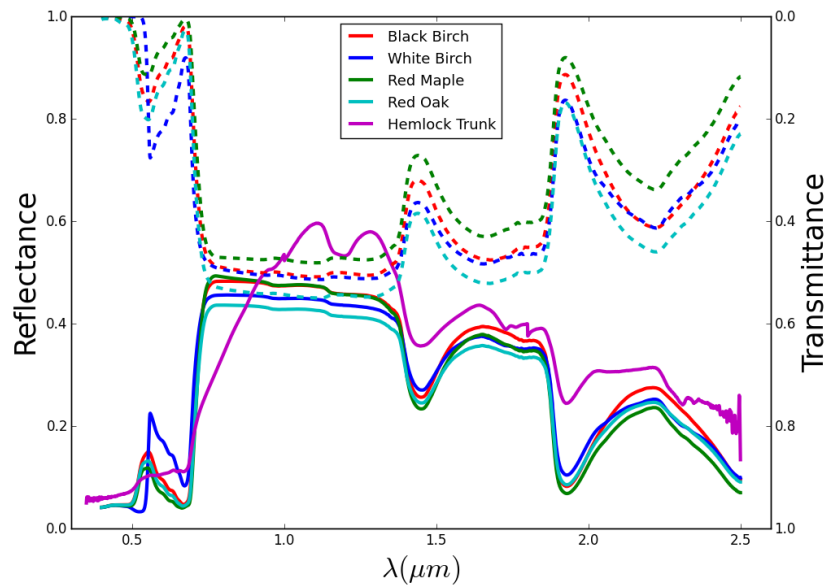
Figure 5.13: The reflectance and transmittance spectra for different tree elements modeled in DIRSIG.
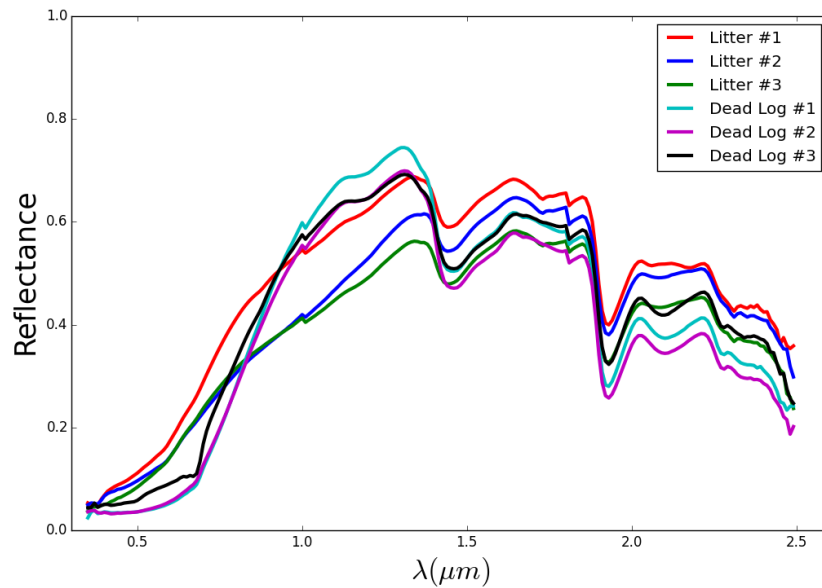
Figure 5.14:   Ground litter spectra collected by the research teams from UMB and BU for Harvard hardwood forest. The field measured spectra was collected in Sep 2013.

In the forest canopy modeling, a broadleaf of a deciduous tree is modeled with multiple facets, as illustrated in Figure 5.15. The surface normal of the facetized geometry is used by the radsolver for radiometric computations. The normal for each facet is oriented in a different direction according to the 3D shape of a leaf. In DIRSIG, each facet of a leaf is assumed to have Lambertian BRDF and Lambertian Bi-directional Transmittance Distribution Function (BTDF). Although all the facets of a leaf have Lambertian BRDF, their outgoing radiance varies by the cosine of its normal's orientation with respect to the illumination angle. This causes a varying outgoing radiance for a leaf when viewed under different viewing and illumination conditions, leading to a non-Lambertian BRDF and BTDF for each leaf. Hence, in the modeling of forest canopy, the leaves are considered to exhibit non-Lambertian BRDF and BTDF, and its directional reflectance or transmittance is dependent on the shape and orientation of the leaf.
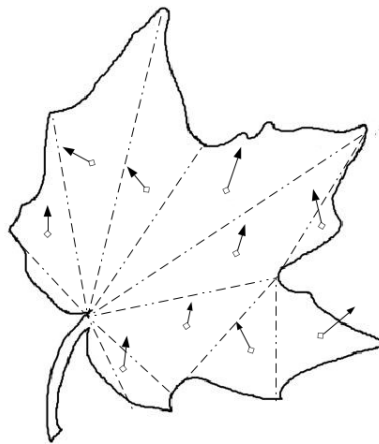
Figure 5.15:   An example showing facetized leaf geometry. The normals for each facets are oriented in different directions.

The tree placement strategy discussed in Section 4.2.2 is used to place trees randomly. The Poisson disc sampling method provides the horizontal coordinates for tree placements and the trees are randomly rotated about the Z (up) direction, to remove any specific pattern. The radius value, $r = 3.5$ for random sampling generated 166 tree positions within a 50 m x 50 m extent and 52 trees in the central 30 m x 30 m scene extent. The virtual forest scene simulated in DIRSIG is shown in Figure 5.16 for nadir and non-nadir view.

The tree trunks are visible in the side view of the forest scene as shown in Figure 5.16a. In the side-view image, the top of the canopy receives direct illumination from the sun and is very bright, while the lower portion of the canopy receives less illumination leading to a dark image. The multiple scattering effect inside the canopy illuminates the lower portion, but it is considerably lower in the VIS spectrum than in the NIR spectrum. The nadir view and off-nadir image shows the texture in the forest canopy due to the variation in height and spectral reflectance of the trees. The shadows from adjacent trees are visible at the center and are also observed at a few darker regions in the image. The rotation of off-nadir image (VAZ=270°) with respect to the nadir image indicates the difference in the location of the shadows due to view angle differences.

(a) side-view   (b) nadir   (c) off-nadir
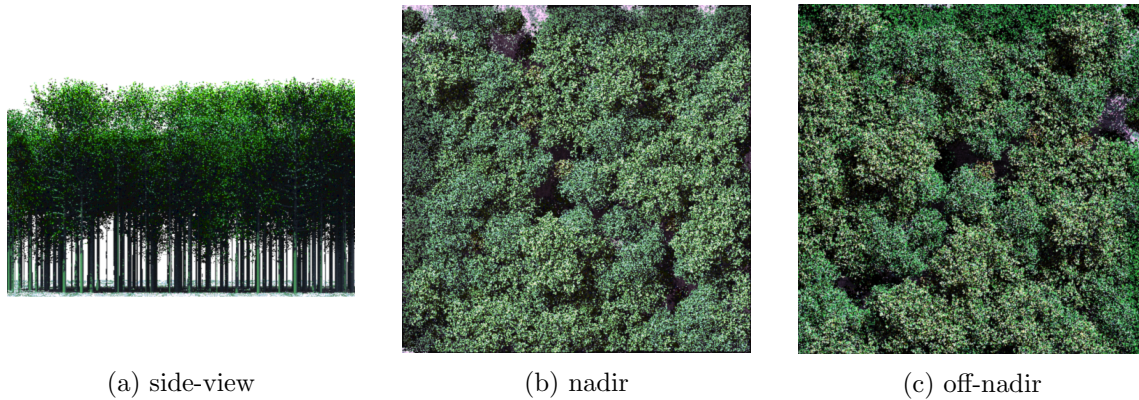
Figure 5.16: Simulation of Harvard forest scene in DIRSIG. Figures (a),(b), and (c) shows the virtual scene when viewed from side view (azimuth (VAZ) = zenith (VZN) = 90°), nadir view (VZN = 0° , VAZ=0°), and at off-nadir view (VZN = 5° and VAZ=270°) respectively. The sun illuminates the scene from 30° zenith angle and 130° azimuth angle.

Two distinct locations at the Harvard forest site observed by a space-borne sensor (Google Earth, 2015) are shown in Figure 5.17 for a side-by-side comparison of the real and the virtual forest. It is important to note that no quantitative comparison is attempted between the real and the synthetic scene. As mentioned earlier, it is difficult to simulate a virtual forest that matches very close to the real forest using the OnyxTree tool, but Figure 5.17 indicates the similarities that can be achieved using the modeling approach discussed in this research. It can be seen that the thick and closed canopy structure and the textural pattern, observed in the synthetic scene is similar to a typical deciduous forest, although not a perfect match to the Harvard site. It is important to understand that the goal is to model a plausible deciduous forest and not an exact match of the real Harvard site.

It is extremely difficult to simulate a real forest without a high resolution multimodal dataset such as ground LIDAR, aerial and forest inventory data. The research in this domain is not mature enough to provide a process that can seamlessly integrate all the datasets to generate a virtual scene (Disney et al., 2013, Hildebrandt and Iost, 2012, Kelbe et al., 2013, Reitberger et al., 2009, Rosell et al., 2009). Lack of such dataset and processes lead to the development of the synthetic forest scene using the OnyxTree and Poisson disc sampling tools. The results indicate that the virtual scene modeled in DIRSIG is a reasonable representative of a deciduous forest.

(a) Harvard forest site   (b) Harvard forest site   (c) Synthetic forest scene
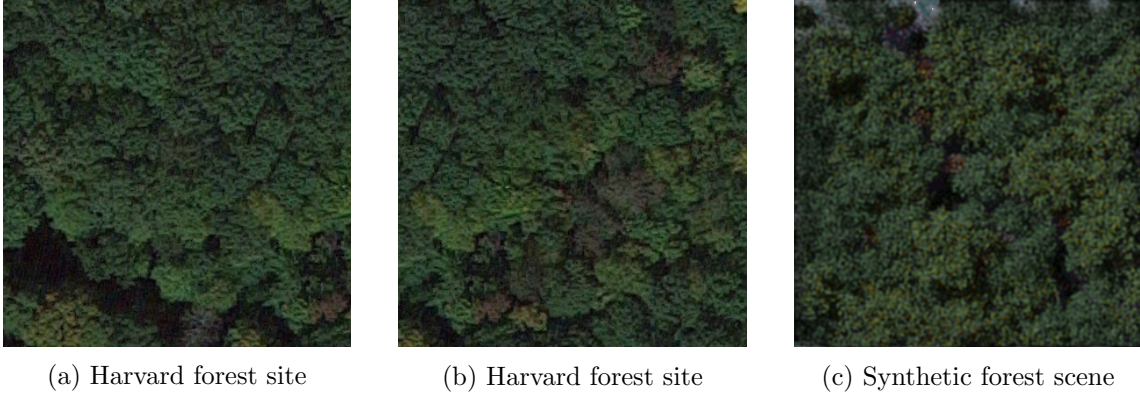
Figure 5.17: Simulation of virtual forest scene using DIRSIG and an image of real Harvard forest site captured by high resolution satellite (Google Earth, 2015). The closed canopy structure and the textural variation are very similar between the synthetic and the real forest images. Figures (a), (b) shows the image of Harvard forest from Google Maps and Figure (c) shows the synthetic forest scene simulated in DIRSIG.

5.3 BRDF measurements

5.3.1 Sensitivity study

A sensitivity study was conducted to determine the optimal number of photon bundles and sub-samples necessary to represent the canopy interactions accurately. The study is conducted by varying one factor at a time and each factor is studied for different levels. This is shown in Tables 5.4 and 5.5. The highest number of photon bundles and detector samples that are practically feasible to run a DIRSIG simulation is used as the reference. In this case, the configuration with ten million photon bundles and 33 sub-samples ($1000/m^2$) is considered as the reference. This configuration is expected to produce the most accurate and noise-free result that can be achieved within a reasonable time. However, the processing time required for this configuration is impractical to use it for measuring BRDF. The photon bundles and the ray sub-samples are propagated randomly and not in a uniform grid as presented in Tables 5.4 and 5.5. However, it is expected that for large samples, as in this case, the approximation of sample density to a uniform grid is valid and reasonable.

As stated in Section 4.3.1, the difference in the radiance between the reference and a specific configuration (ΔL) is compared to the expected noise from a space-borne sensor. For the spectral region of interest (NIR), $SNR = 600$ is assumed to be the higher limit expected from the remote sensing systems. The noise calculated for this SNR at the average radiance is found to be 0.16 $Wm^{-2}sr^{-1}\mu m^{-1}$.

Table 5.4: Sensitivity study for estimating the optimal photon bundle configuration. Each photon bundle illuminates an entire voxel of the scene, but their horizontal sampling density is shown here.

# bundles (in millions)	Photon bundle density (# bundles per unit area)	Bundle spacing in one dimension (square grid)
10.0	$\sim 4,000$ per m^2	1 every 1.5 cm
6.25	$\sim 2,500$ per m^2	1 every 2 cm
5.00	$\sim 2,000$ per m^2	1 every 2.25 cm
4.00	~ 1600 per m^2	1 every 2.5 cm
2.80	~ 1120 per m^2	1 every 3 cm
1.60	~ 640 per m^2	1 every 4 cm
1.00	~ 400 per m^2	1 every 5 cm
0.51	~ 204 per m^2	1 every 7 cm
0.25	~ 100 per m^2	1 every 10 cm
0.12	~ 48 per m^2	1 every 15 cm

Table 5.5: Sensitivity study for estimating the optimal detector sub-samples configuration.

# sub-samples (in one dimension)	sub-sample density (# bundles per unit area)	sub-sample spacing in one dimension (square grid)
2	~ 4 per m^2	1 every 50 cm
3	~ 9 per m^2	1 every 33 cm
4	~ 16 per m^2	1 every 25 cm
5	~ 25 per m^2	1 every 20 cm
7	~ 49 per m^2	1 every 14 cm
10	~ 100 per m^2	1 every 10 cm
14	~ 196 per m^2	1 every 7 cm
20	~ 400 per m^2	1 every 5 cm
25	~ 625 per m^2	1 every 4 cm
33	~ 1000 per m^2	1 every 3 cm

The DIRSIG simulations for different configurations are performed with the same Harvard forest scene modeled earlier. All the simulations are imaged under the same atmospheric condition (mid-latitude summer) and illumination angles. The illumination angles are represented by the position of the sun on a summer solstice day, for Northern Hemisphere, at 11 AM over the Harvard forest site. The camera captures the forest scene at $1m$ GIFOV at nadir from 100 KM altitude. Each simulation is performed by varying the photon bundles and the detector sub-samples as shown in Tables 5.4 and 5.5. The flux density map from the Geodensity radsolver is dependent solely on the number of photon bundles. Therefore, all the simulations with varying detector sub-samples use the pre-computed flux density map from a specific photon

bundle configuration. The absolute difference in radiance is compared to the noise for different configurations and are shown in Figures 5.18 and 5.19.
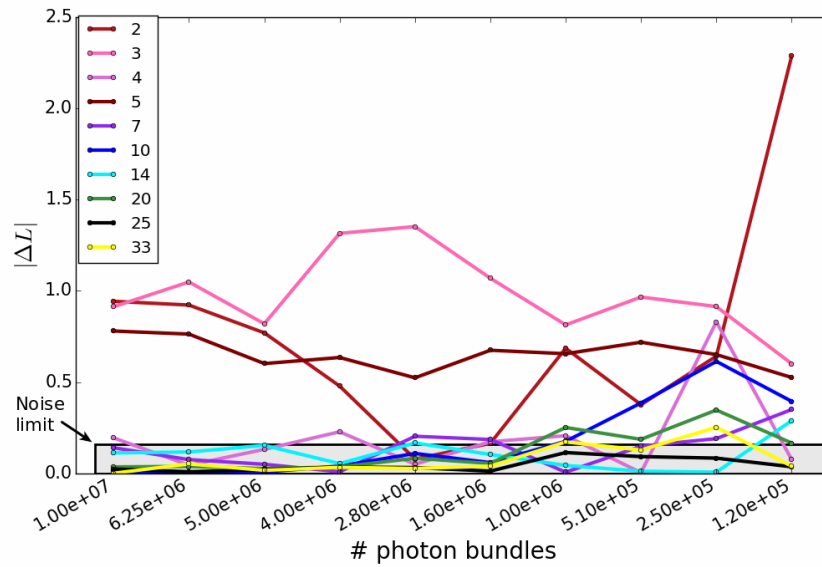
Figure 5.18:   Plot to show the difference in radiance vs photon bundles for different detector sub-sample configurations. The limiting criterion for optimal selection (noise limit) is shown in the gray box.
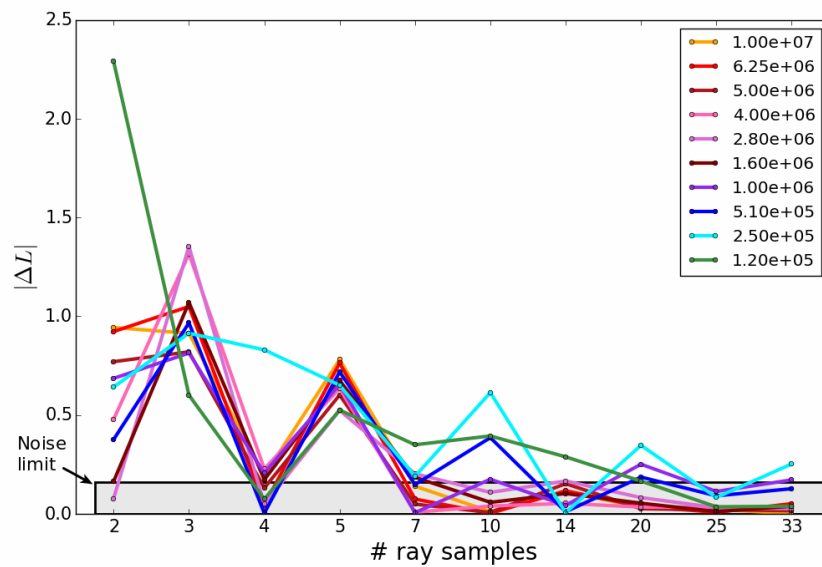
Figure 5.19: Plot to show the difference in radiance vs detector sub-samples for different photon bundle configurations. The limiting criterion for optimal selection (noise limit) is shown in the gray box.

The plot in Figure 5.18 shows noticeable noise in the results for low sub-samples and does not follow the exponential behavior as expected. This indicates that the low detector sub-samples

do not provide any gain irrespective of the number of photon bundles. The plot in Figure 5.19 shows the expected behavior but with a noticeable noise affecting the optimal selection. Certain configurations indicate a low ΔL with moderate sub-samples, but it increases for the very next higher sub-sample configuration. This is attributed to the correlation between the facet size, ray sampling and photon bundle density. It is possible that the sample distance for sub-sample rays may correspond closely with the facet size, leading to a better estimation of the outgoing radiance from the scene. Hence, an optimal value is chosen when the configuration for photon bundle and ray sub-sampling shows a consistent behavior.

Figures 5.20 and 5.21 show the useful configurations for photon bundles and sub-samples. There is more than one choice that satisfies the required criteria, but the emphasis is given to the detector sub-samples over photon bundles. The number of photon bundles, illumination angle, and wavelength impact the incoming radiation and the flux density map. However, the flux density maps can be stored and reused for different view angles, reducing the computational time for photon bundle propagation. Secondly, processing time increases quadratically with increase in sub-samples, but linearly for an increase in photon bundles. Thus, finding the smallest number of sub-samples reduces the processing time for BRDF simulations. The smallest number of sub-samples required for meeting the noise criteria is found to be 7 sub-samples, and the corresponding photon bundles should be higher than 4 million bundles.
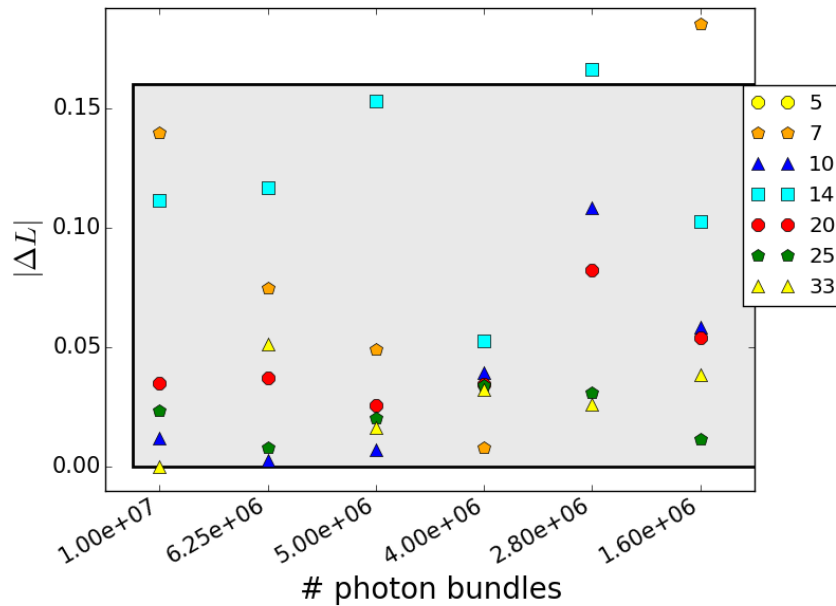
Figure 5.20: Plot of the difference in radiance with respect to photon bundles. The best and the smallest detector sub-sample is found to be 7 sub-samples.

For BRDF measurements, 5 million photon bundles and 10 ray sub-samples are chosen, as their results are consistently better and higher than the minimum requirement. This translates to approximately 1 ray sub-sample every 10 *cm* and 1 bundle every 2.25 *cm*.
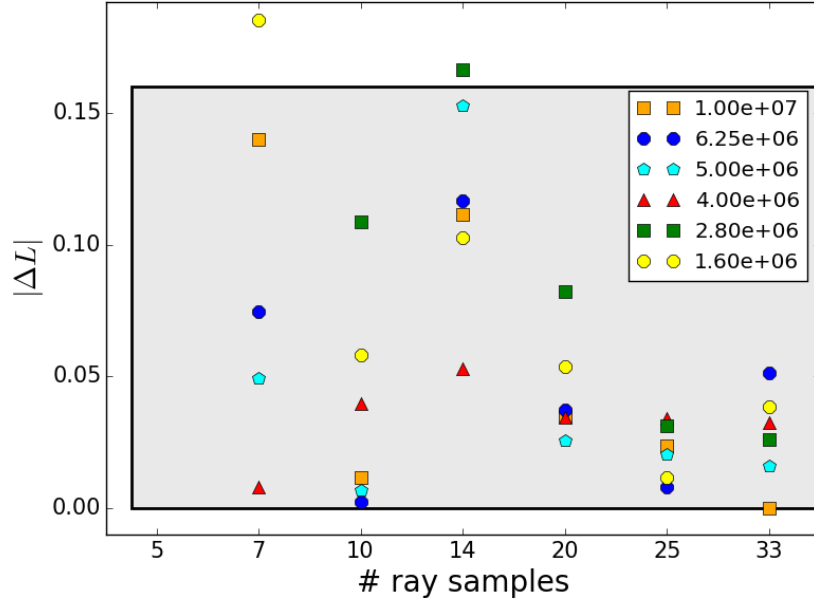
Figure 5.21: Plot of the difference in radiance with respect to sub-samples. The best and optimal number of bundles is found to be 4 million photon bundles.

The radiance results from the sensitivity study are converted to reflectance to estimate the error in reflectance units. This is achieved using the ELM method. The ELM method requires two distinct reflectances and their corresponding radiances, which are used to establish a linear relationship between the reflectance and the radiance. The linear coefficients can then be used to convert the radiance image into a reflectance image.

A scene with two Lambertian reflector panels (10 % and 50 %) is modeled in DIRSIG. The scene is simulated for the same atmospheric condition, illumination and viewing geometry as used in the sensitivity study.

The simulated image of the panel in radiance units and its associated reflectance is used to estimate the linear coefficients. These linear coefficients are used to convert all the radiance images from the sensitivity study to reflectance factor units. The absolute difference between the reference and all the other configurations are calculated as before, but in reflectance units and are shown in Figure 5.22. The relative error in percent is shown in Figure 5.23.
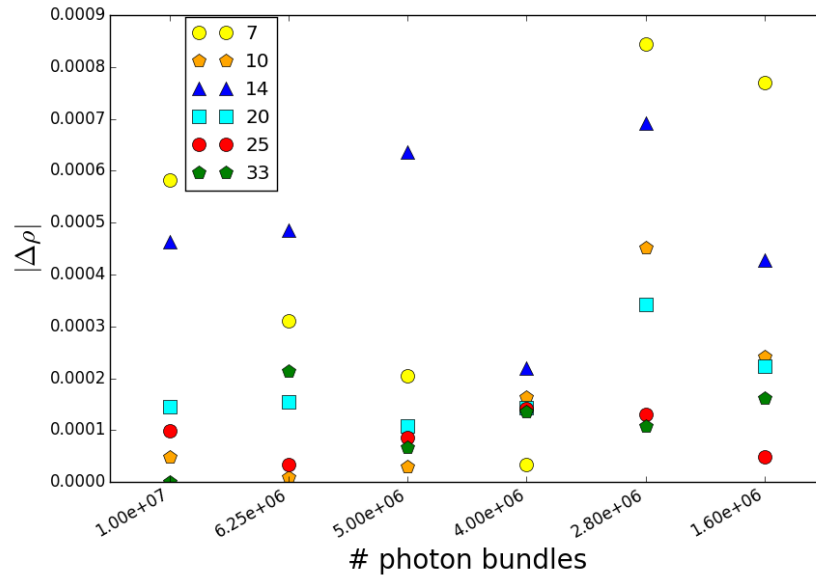
Figure 5.22: Sensitivity study of photon bundles and detector sub-samples in reflectance factor units.
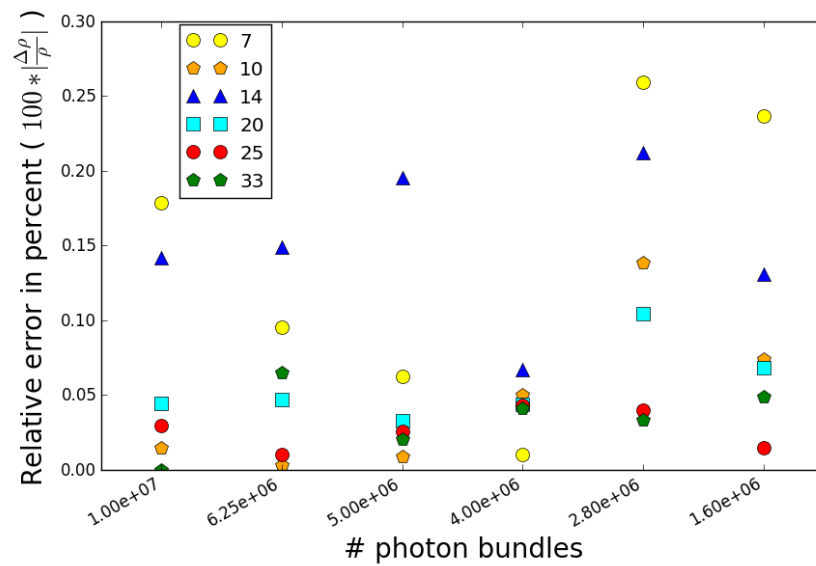
Figure 5.23: Relative error for different photon bundle and sub-sample configurations used in the sensitivity study.

The error introduced in the chosen configuration (10 ray sub-samples and 5 million bundles) is less than $\frac{1}{1000}$th in reflectance units. The results clearly indicate that the chosen configuration will introduce no more than $\frac{3}{10}$th of a percent of error for the forest canopy reflectance, and is thereby considered negligible for the BRDF measurements.

5.3.2 BRDF sampling

The BRDF for the Harvard scene is measured in DIRSIG for specific wavelength, illumination and view angles. The samples chosen for these measurements are shown in Tables 5.6 and 5.7, and are represented in Figures 5.24, 5.25, and 5.26. The illumination samples are chosen such that they span the season for leaf-on conditions. The FOV of L8 and S2 are padded and adjusted for the curvature of the earth, and appropriate view samples are determined.

Table 5.6: Illumination and view angles sampled for BRDF measurement.

Illumination angles		View angles	
Azimuth	Zenith	Azimuth	Zenith
140	35	0	90
132	24	2	180
148	41	2	0
153	20	2	104
151	34	7	98
158	26	13	86
161	35	17	90
166	40	15	98
142	21	3	284
166	71	8	279
240	75	12	263
300	70	17	270
		15	278
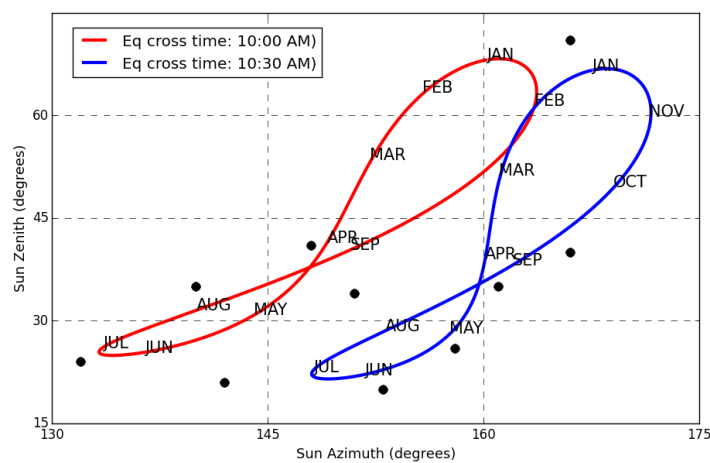
Figure 5.24: Samples for illumination angle plotted in black dots
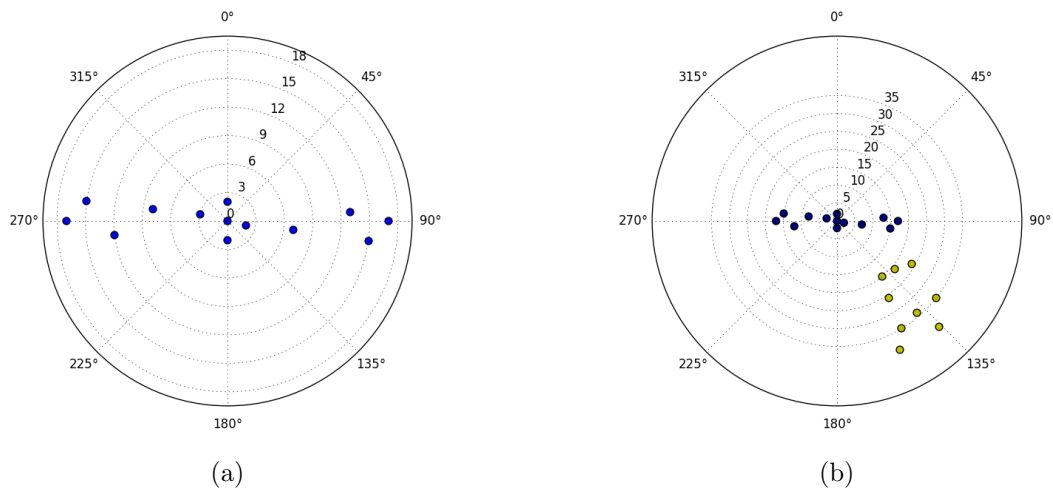
Figure 5.25: Samples chosen for illumination and view angles in polar plot. Figure (a) shows the view angle samples (blue) and Figure (b) shows both sun (yellow) and view angle samples (blue).
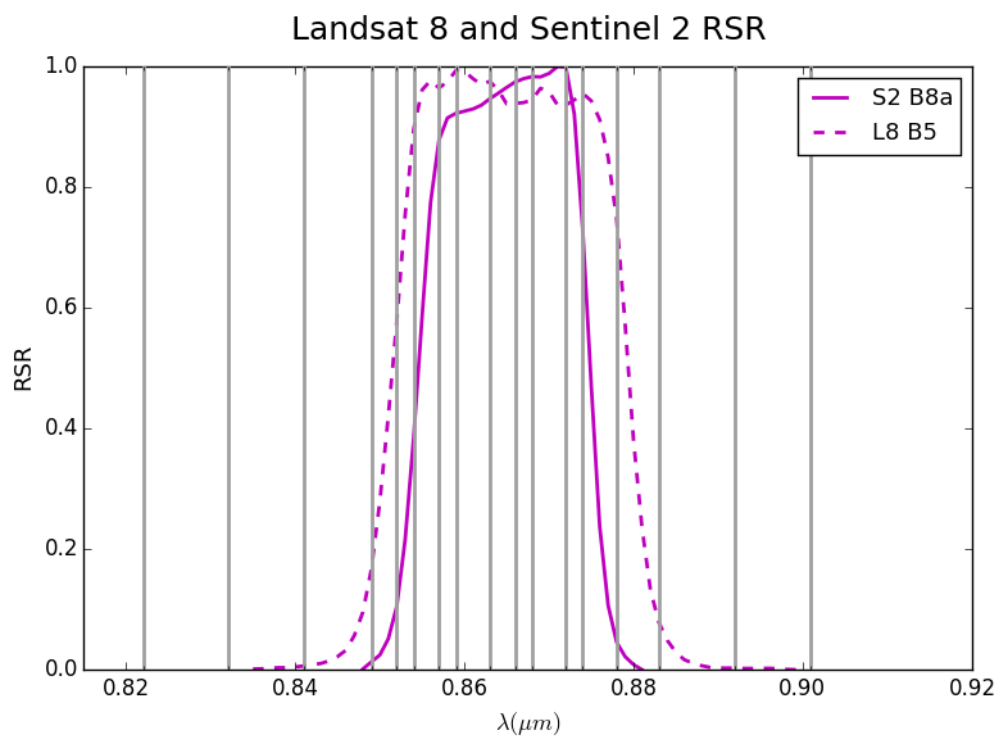
Figure 5.26:   Wavelength samples chosen for NIR spectral bands of OLI and MSI sensor. The gray line indicates the sampled wavelength. The samples for the remaining bands are shown in Appendix E

Table 5.7: Wavelength samples chosen for BRDF measurement

Spectral Bands			λ samples (μm)
Bands	OLI Band	MSI Band	
Blue	Band 2	Band 2	0.450, 0.455, 0.463, 0.475, 0.487, 0.498, 0.512
Green	Band 3	Band 3	0.520, 0.527, 0.537, 0.541, 0.544 0.552, 0.561, 0.569, 0.576, 0.580, 0.585, 0.588, 0.592
Red	Band 4	Band 4	0.636, 0.640, 0.649, 0.654, 0.662 0.671, 0.675, 0.680
Red Edge 1	-	Band 5	0.696, 0.699, 0.705, 0.709, 0.711
Red Edge 2	-	Band 5	0.732, 0.733, 0.735, 0.740, 0.743 0.746, 0.748
Red Edge 3	-	Band 7	0.772, 0.775, 0.777, 0.780, 0.784 0.791, 0.794
NIR	Band 5	Band 8a	0.800, 0.810, 0.822, 0.832, 0.841 0.849, 0.852, 0.854, 0.857, 0.859, 0.863, 0.866, 0.868, 0.872, 0.874, 0.878, 0.883, 0.892, 0.901
SWIR 1	Band 6	Band 11	1.560, 1.565, 1.570, 1.575, 1.580 1.590, 1.600, 1.620, 1.630, 1.640, 1.650, 1.655, 1.660, 1.665
SWIR 2	Band 7	Band 12	2.100, 2.105, 2.110, 2.120, 2.125 2.130, 2.150, 2.170, 2.190, 2.210, 2.230, 2.240, 2.250, 2.270, 2.280, 2,285, 2.290 2.300, 2.310

There are 12 samples for the illumination angles, 13 samples for the view angles, and 100 samples for the wavelength. In total, this requires $12 * 13 * 100 = 15,600$ simulations. An automated script is used to construct the required files for DIRSIG simulations, and the simulations were processed at the RIT's research computing facility. As stated earlier, the flux density map is dependent on the wavelength and illumination angles. It requires $12 * 100 = 1200$ simulations to generate all the distinct flux density maps and they are processed ahead of the other simulations. The saved maps are used by all the other simulations, reducing the processing time significantly ($\sim 80\%$ reduction).

The 3D ROI process explained in Section 4.3.2.3 is unique to a specific view angle. The 13 distinct ROIs are used to estimate the reflectance across all the simulated images. The ROI overlay of two different view angles are shown in Figure 5.27 and the rest are shown in Appendix E.

(a) VAZ : 90° , VZN : 0°   (b) VAZ : 263° , VZN : 12°
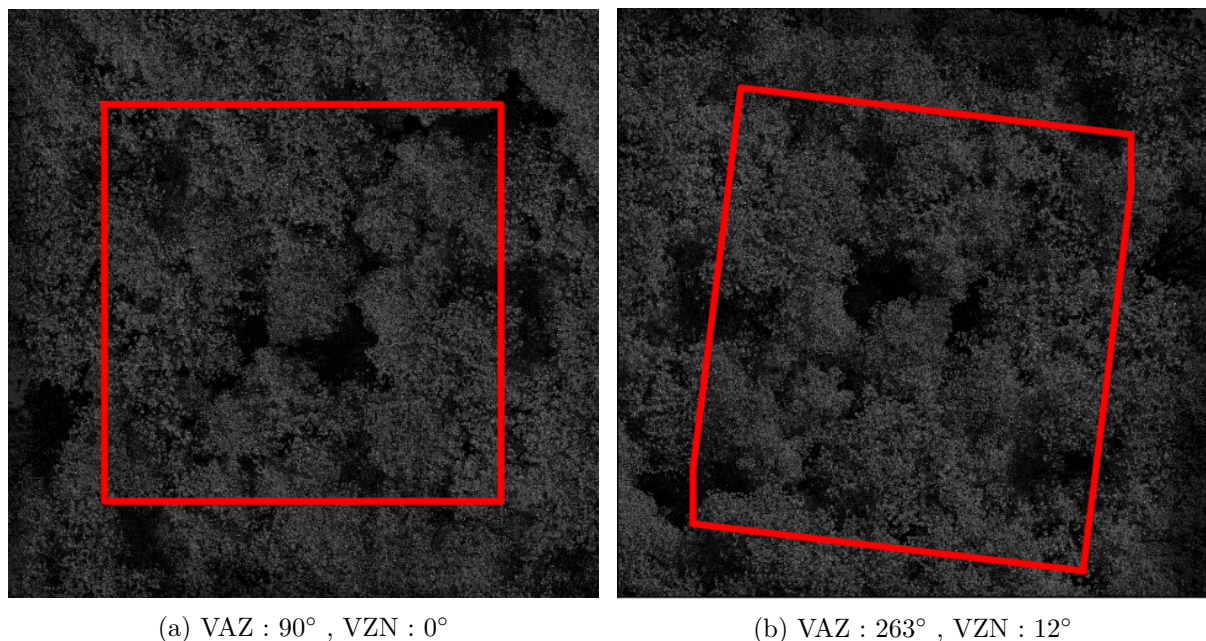
Figure 5.27: The 3D ROI overlay of two different view angles when illuminated by the sun from 30° zenith angle and 130° azimuth angle ($\lambda = 0.866\mu m$). The 3D ROI ($-15 < X < 15$, $-15 < Y < 15$, $15 < Z < 30$) represents the central section of the scene with height estimated from the top of the canopy.

The measurements for BRDF are generated for different forest conditions and are modeled independently. For example, a forest on a sloped terrain will run all the $15,600$ simulations independently and these measurements will be used for BRDF modeling for sloped terrain. The

samples and the 3D ROI results shown in this section are applicable for all the different forest conditions, such as slope and seasonal changes.

5.4   BRDF modeling

The BRDF measurements for a specific forest condition, such as forest on a flat terrain, is used to fit a RossLi BRF model and the model coefficients are estimated. The least squares model is used for fitting the measurements to derive the three RossLi model coefficients. The model coefficients fully describe the BRF model of the forest canopy. The RossLi BRF model uses BR and HB model coefficients, as explained in Section 4.4. The HB and BR parameters for each tree and the weighted HB and BR of the canopy is shown in Table 5.8.

Table 5.8: The RossLi BRDF parameters (HB and BR) for each tree modeled in DIRSIG.

Tree Name	Vertical crown radius	Horizontal crown radius	Height to crown center	HB	BR	weight
BlackBirch (BB)	6.1	2.8	13.9	2.2	2.1	0.06
PaperBirch (PB)	6.4	1.5	13.3	2.0	4.1	0.1
RedMaple 1 (RM-1)	3.3	2.5	7.2	2.1	1.3	0.07
RedMple 2 (RM-2)	6.8	3.6	16.6	2.4	1.8	0.08
RedMaple 3 (RM-3)	6.8	4.4	16.9	2.4	1.5	0.14
RedMaple 4 (RM-4)	6.7	2.9	16.2	2.3	2.3	0.15
RedOak 1 (RO-1)	4.3	3.2	17.5	4.0	1.4	0.11
RedOak 2 (RO-2)	7.9	5.8	15.1	1.9	1.3	0.17
RedOak 3 (RO-3)	4.7	7.2	25.1	5.0	0.7	0.09
YellowBirch (YB)	3.2	2.1	7.5	2.3	1.5	0.03
Canopy HB / BR				2.5	1.9	

The kernel functions in the RossLi BRF model are dependent on the viewing and illumination geometry, and the Li kernels are dependent on HB and BR parameters as well. The HB and BR parameters of the canopy, and the viewing geometry for each measurement, are used to generate the corresponding kernel functions. The BRF reflectance for each viewing geometry is determined using the kernel functions and the model coefficients as shown in Equation 4.9.

As mentioned in Section 4.4, two independent RossLi model coefficients are generated depending on the zenith angle of the sun. The one-to-one plot of the measured and model derived

reflectance shows the fit accuracy of the model. Figure 5.28 shows the fit between the measurement and the model reflectance when the illumination geometry is restricted to low zenith angles (SZN < 50°).
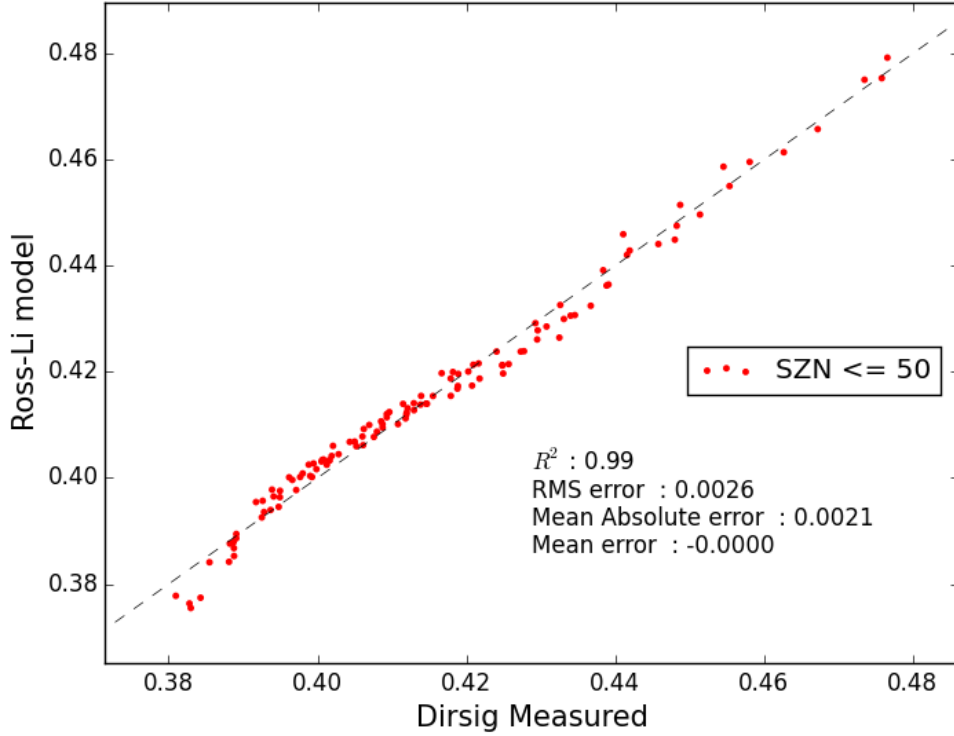
Figure 5.28: Plot of measured BRF against modeled BRF for $\lambda = 0.866 \ \mu m$, SZN < 50°, and the ground slope is 0°.

The regression coefficient ($R^2 = 0.99$) for the model fit is very high indicating an accurate fit to the measurements for low solar zenith angles. The RMSE of the fit is less than $\frac{3}{10}$th of a reflectance unit. Figure 5.29 shows the fit between the measured and the modeled reflectance for high zenith angles. The number of data points (3 illumination * 13 view =39 simulations) used to fit the model is small as compared to the case of low zenith angles. The plot shows that the model does not correctly fit the data as its regression coefficient is very low and the fit RMSE is more than 2 reflectance units. Since these model coefficients are used only for estimating the skylight contributions to at-sensor radiance, the error introduced by the model is considered small. A sensitivity analysis is performed to determine the skylight contribution and is discussed in the following section.
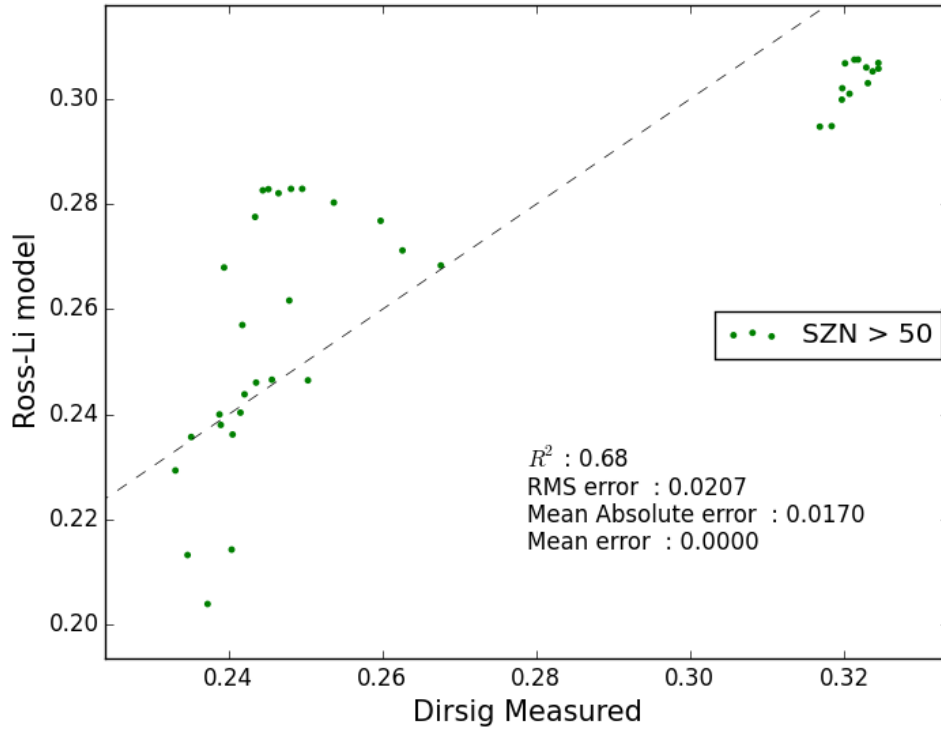
Figure 5.29: Plot of measured BRF against modeled BRF for $\lambda = 0.866\ \mu m$, SZN > 50°, and the ground slope is 0°.

The regression plot, as a function of wavelength, is shown in Figure 5.30. The consistency in the RossLi model fit across the wavelengths for low solar zenith angles clearly indicate the spectral invariant nature of the RossLi BRF model and its accuracy to fit the measurements. However, for large solar zenith angles, the regression coefficient shows a pattern similar to the forest spectral reflectance, i.e., low regression coefficient when the canopy reflectance is low. This indicates a bias in the model fit as a function of reflectance for large zenith angles. This could be due to the noise in the measurement for low reflectance in DIRSIG, especially, when the solar zenith angle is high.

The incorrect fit between the measured and modeled data when the solar zenith angle is high may be attributed to a few reasons. Firstly, the RossLi BRF model is assumed to be an approximate representation of the canopy BRF and is valid for low zenith and view angles. Secondly, the RossLi BRF model assumes that the forest is azimuthally symmetric, whereas the virtual forest is modeled using the random tree placement technique. The asymmetric nature of the synthetic scene can introduce differences in reflectance as a function of azimuth angle, particularly when the trees cast large shadows for high solar zenith angles. The results indicate that the RossLi BRDF model cannot describe the measurements for high zenith angles accurately. However,

as shown in Section 4.4.1.3, any error in the modeling of BRF for high solar zenith angles will only affect the skylight contribution, and these modest errors in a small term will result in low overall errors.
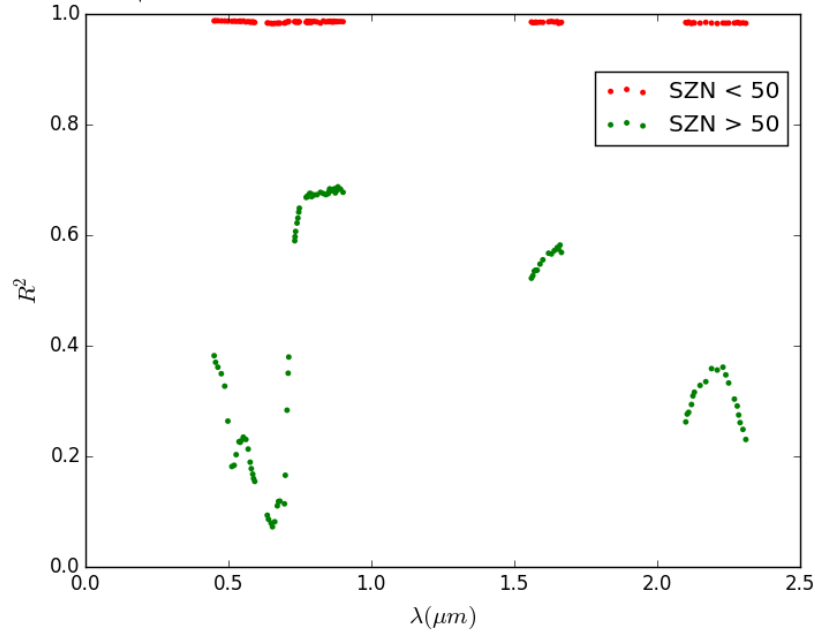
Figure 5.30: Plot of regression coefficient for the RossLi model fit as a function of wavelength.

The RossLi model coefficients are used to generate the BRF in polar plot, as shown in Figure 5.31. The polar plot shows the reflectance for a specific wavelength ($\lambda = 0.866\mu m$) and solar illumination angle (SAZ = 139° , SZN = 30°). The view angles are restricted to 50° and the hot-spot region is seen at the back-scatter direction.

Similar to Figure 5.31, a BRF plot is generated for all the illumination angles for a given view angle. The polar plot of BRF for $\lambda = 0.866\mu m$, and nadir view is shown in Figure 5.32. The reflectance across all the azimuth angles for a given solar zenith angle is the same, due to the azimuthally symmetric assumption of the RossLi BRDF model.

The BRF in the principal and the cross plane is shown in Figure 5.33. The back-scatter convention is different in this plot than the RAMI validation. The zenith angle is positive when the view and illumination azimuth angles are equal. The principal plane shows the hot-spot effect observed at 30° zenith angle. The cross plane shows the peak reflectance at nadir, and the variation in reflectance is small, as expected.
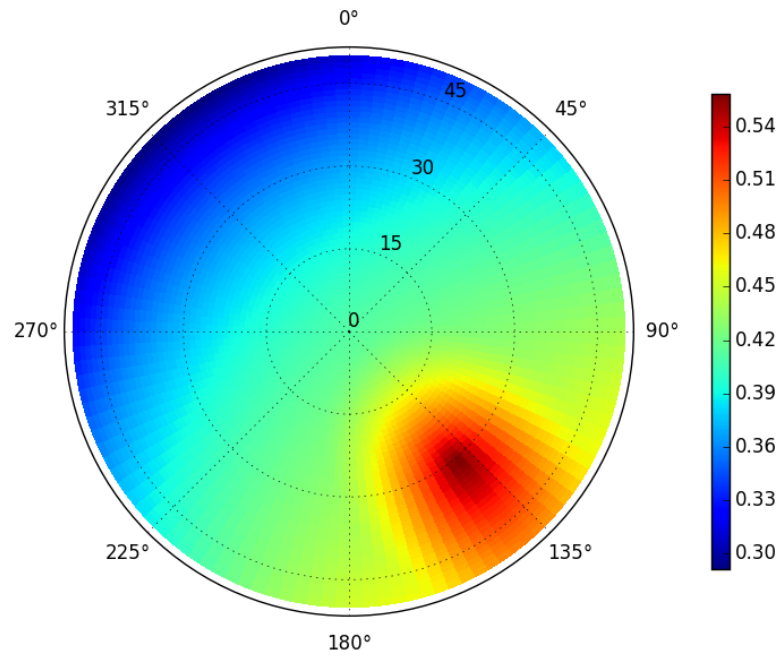
Figure 5.31: Polar plot of the spectral BRF for different view angles for $\lambda = 0.866\mu m$, SAZ = 139°, SZN = 30°, and the ground slope is 0°. The hot-spot due to the back-scatter effect is clearly visible at the same view angle as the illumination angle.
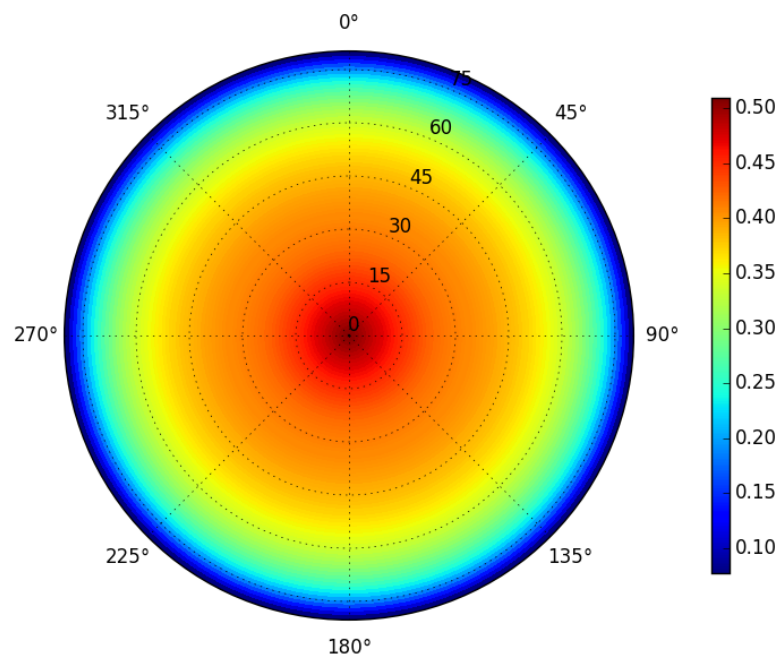
Figure 5.32: Polar plot of the spectral BRF for different illumination angles for $\lambda = 0.866\mu m$, VAZ = 0°, VZN = 0°, and the ground slope is 0°. The hot-spot effect is clearly visible at the nadir.

(a) principal plane
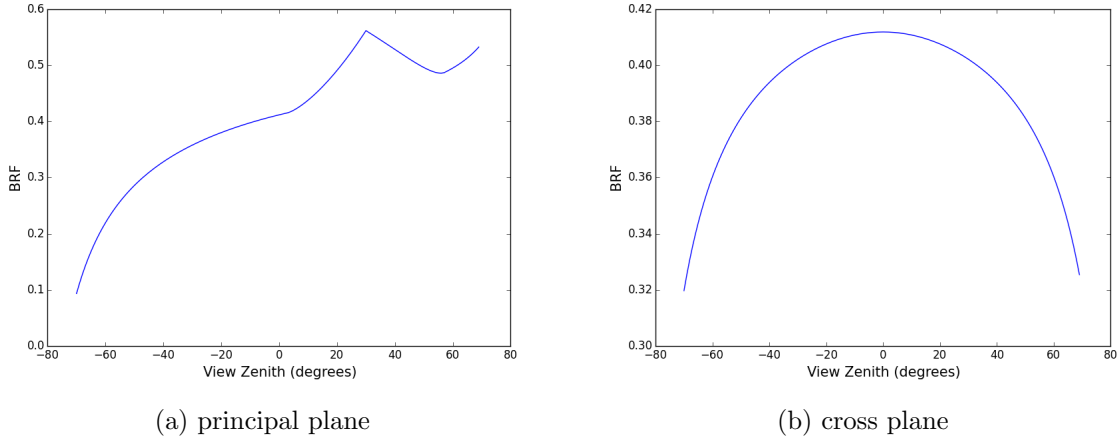
(b) cross plane

Figure 5.33: The principal plane and cross plane BRF for $\lambda = 0.866\mu m$, SAZ = 139°, SZN = 30°, and the ground slope is 0°. The hot-spot effect is observed at VZN = 30° in the principal plane.

5.4.1 Sensitivity study

5.4.1.1 Measurement sensitivity

As discussed in Section 4.4.1.1, sensitivity of the RossLi BRDF model to measurements can be evaluated based on both the RossLi model coefficients and using the reflectance measurements averaged over several illumination and view angles. The LOOC validation is performed by a random selection of six sun angles and ten view angles, as shown in Table 5.9. Each of these angles is removed one at a time and its relative error in the RossLi model coefficients with respect to the reference model is evaluated. In total, 16 different RossLi model coefficients are estimated and compared against the reference. Table 5.10 shows the summary statistics of the relative difference for the 6 sun and 10 view angles for a specific wavelength ($\lambda = 0.866\mu m$). It can be seen that the relative error for the view samples are much smaller than the sun samples. This indicates that the RossLi model has a very low sensitivity to the view angles and are likely to introduce a larger effect for variation in the sun angles. Further, the mean and RMSE of the RossLi geometric model coefficient (f_{geo}) for the sun and view angles are much higher than the other coefficients, but its effect in the BRDF is small since the geometric coefficient is very small and closer to zero. Figure 5.34 shows that the relative error for the three RossLi model coefficients are independent of the wavelength.

Similar to the model coefficients, the sensitivity of the measurement in reflectance is assessed by the relative error in reflectance. In this case, for each of the 16 model coefficients determined

Table 5.9: Sun and view angles used in measurement sensitivity analysis

	Sun angles			View angles	
ID	Zenith	Azimuth	ID	Zenith	Azimuth
S1	21	129	V1	2	104
S2	41	135	V2	12	263
S3	24	119	V3	2	0
S4	35	127	V4	15	278
S5	40	153	V5	17	270
S6	34	138	V6	13	86
			V7	7	98
			V8	0	90
			V9	3	284
			V10	17	90

Table 5.10: RossLi model sensitivity to measurements based on its coefficients using LOOC validation technique for $\lambda = 0.866 \mu m$.

Samples	RossLi coefficient	STD	Mean	RMSE
	f-iso	0.06	0.24	1.65
Sun samples	f-geo	10.53	21.10	23.58
	f-vol	0.50	1.65	1.72
	f-iso	0.06	0.11	0.13
View samples	f-geo	1.18	3.13	3.35
	f-vol	0.35	0.62	0.71

earlier, the reflectance (BRF) for different sun and view angles are computed (see Table 5.11). These BRF values are compared with the reference BRF to estimate the relative error as shown in Equation 5.1. The mean relative error for all the sun and view samples are estimated across VIS-NIR-SWIR region to check for spectral independence. As in the case with the coefficients, the relative error is larger for the sun angles than the view angles and is independent of wavelength (see Figure 5.35). It can be seen that the relative error is less than $\frac{1}{5}^{th}$ of a percent which is extremely small and beyond the sensitivity of many instruments.

Table 5.11: Sun and view angles used in the RossLi model sensitivity analysis

	Sun angles		View angles	
ID	Zenith	Azimuth	Zenith	Azimuth
1	35	127	0	90
2	24	119	2	180
3	41	135	7	98
4	20	140	13	86
5	34	138	17	90
6	26	145	15	98
7	35	148	3	284
8	40	153	8	279
9	21	129		

(a) Sun samples

(b) View samples
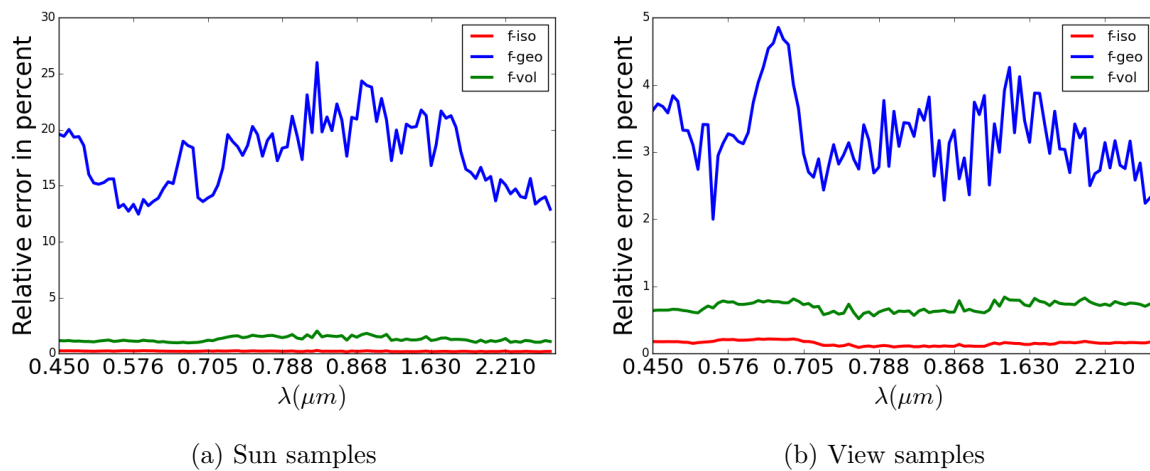
Figure 5.34: Relative error of the RossLi model coefficients using LOOC validation technique. Figure (a) and (b) shows the relative error for sun and view angles respectively. The relative error is in percent units and found to be very small for isotropic and volumetric coefficients but much higher for the geometric coefficient.

$$\Delta r_{ij}(percent) \;=\; 100 * \frac{\rho_{ij} - \rho_{ir}}{\rho_{ir}}$$

$$\mu_{r_j} = \frac{\sum\limits_{k=1}^{n} \Delta r_{ij}}{n} \tag{5.1}$$

where,

ρ_{ij} is the BRF reflectance for i^{th} sun or view angle for j^{th} LOOC model coefficient

ρ_{ir} is the BRF reflectance for i^{th} sun or view angle for the reference model coefficient

i is the sun or view angle shown in Table 5.11

j corresponds to one of 16 models derived by removing one view or sun angle from Table 5.9

n is the number of sun and view angles shown in Table 5.11

(a) Sun and view samples ($\lambda = 0.866\mu m$)

(b) Mean of all samples
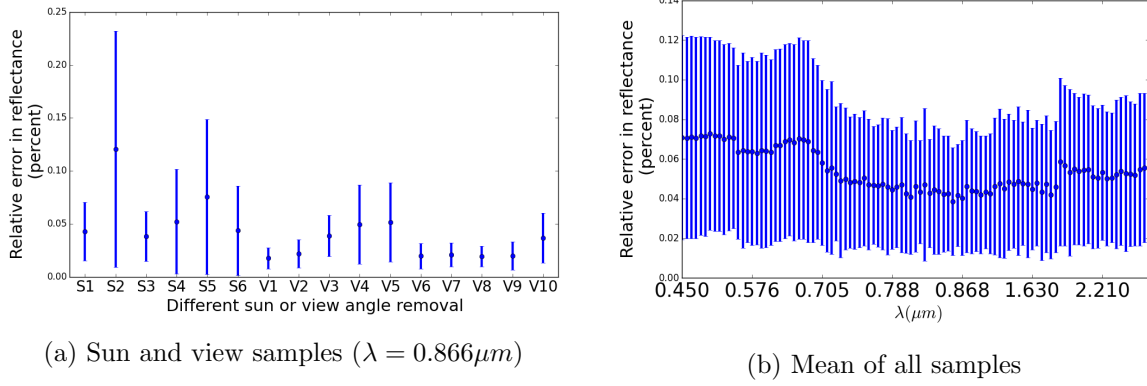
Figure 5.35: Relative error of the BRF reflectance using LOOC validation technique. Figure (a) shows the relative error with its 1 standard error for different sun (S1-S6) and view (V1-V10) samples for $\lambda = 0.866\mu m$. Figure (b) shows the mean relative error with 1 standard error as a function of wavelength. The relative error is less than $\frac{1}{5}^{th}$ of a percent.

Thus, from the LOOC validation technique, it is concluded that the sensitivity of the measurements to generate the RossLi model coefficients is extremely small and its effect on the model is negligible.

5.4.1.2 Modeling sensitivity

The model sensitivity analysis is performed by measuring the reflectance for varying sun, view and spectral samples, similar to the canopy BRDF measurements. In this case, the spectral samples are limited to the center wavelength of OLI multispectral bands (Blue, Green, Red, NIR, SWIR-1, SWIR-2). The view angles are sampled at every $3°$ and $10°$ in the zenith and azimuth directions respectively. The solar zenith samples are chosen such that one of them is $70°$ and the rest are less than or equal to $50°$. DIRSIG is used to measure the BRDF for all these samples and the measurements are categorized to fit four different RossLi models based on the view and solar zenith angle constraints. The RMS of the fit residuals and the R^2 of the model fit are then used to evaluate the validity of each model. Figure 5.36 shows the fit between the DIRSIG and the RossLi model for the four different sun and view conditions in the NIR band ($\lambda = 0.866\mu m$).

From Figure 5.36a, it is seen that the measurements match the RossLi model when the solar zenith angles are less than or equal to $50°$ and view zenith angles are less than $40°$. For the majority of remote sensing sensors and leaf-on conditions, the RossLi model is capable of modeling the measurements accurately. The model fit is also reasonable when the solar zenith angle is $70°$, as shown in Figure 5.36b. However, when the view zenith angles are greater than

40°, the model derived reflectance does not match well with the DIRSIG measurements, as evident in Figures 5.36c and 5.36d. This mis-match can be attributed to the measurement and modeling issues. The limited size of the forest canopy and a fixed-size frame camera used for BRDF measurement reduces the number of forest pixels within the 3D ROI for larger view zenith angles, as discussed earlier in Section 4.3.2.3. Secondly, the RossLi model assumes a homogeneous forest where the height of the trees are assumed to be the same, but the virtual forest has trees of uneven height. This causes more issues especially when the tree shadows are large for high solar and view zenith angles.

(a) SZN <= 50° , VZN < 40°

(b) SZN > 50° , VZN < 40°

(c) SZN <= 50° , VZN > 40°

(d) SZN > 50° , VZN > 40°
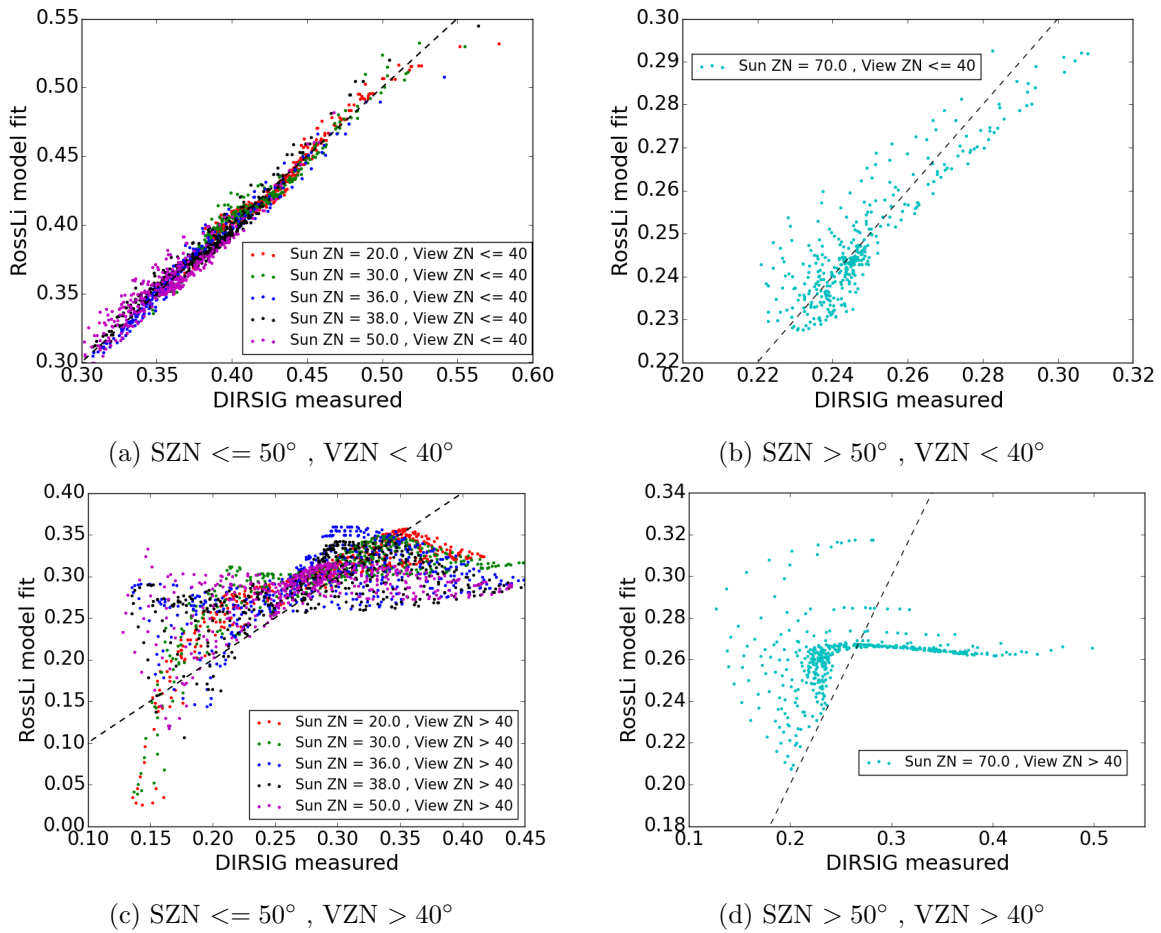
Figure 5.36: Fit between the DIRSIG and the RossLi model for different sun and view conditions. The DIRSIG measurements matched the RossLi model when the view zenith angles are less than 40° (a,b), but the fit is poor when the view zenith angles are more than 40° (c,d).

Figure 5.37a shows the regression coefficient (R^2) for the different view and solar zenith angle conditions. The 'red' line in the plot shows the R^2 value when the solar zenith angles are less than or equal to 50° and the view angles are less than the three cut-off view angles (20°, 30°, 40°). Similarly, the 'blue' line indicates the R^2 when the solar zenith angle is 70°. The high R^2 value

(≈ 0.85) for both the cases (red and the blue line) indicate that the model is consistent and accurate for low view zenith angles. The low R^2 value for the other two cases when the view angles are greater than the cut-off angle (green and black lines) indicate that the model is inconsistent for high view zenith angles.

Figure 5.36 shows that the model fits well when the view zenith angles are less than $40°$, but Figure 5.37a shows that the R^2 value is small when the measurements whose VZN angles are less than $40°$ (but $> 20°$) are also included with the other measurements ($> 40°$). This discrepancy is due to the large residuals observed when VZN is greater than $40°$ that the overall model fit is poor. This is also observed in the RMS plot shown in Figure 5.37b. It is somewhat unexpected that the RMS is marginally lower when the SZN angle is high ($70°$) than when the SZN angle is less than or equal to $50°$. This is because, the SZN $<= 50°$ case consists of 5 different SZN angles which introduces a lot more variability within the model, and as a result reduces the model fit accuracy (R^2 and RMS). The variation of the reflectance due to the sun angles (zenith and azimuth), though small, is primarily due to the azimuthally symmetric assumption of the forest for the RossLi BRDF model. The RMS error is about 3% in the NIR band suggesting that the RossLi BRDF model is acceptable and valid for modeling deciduous forest canopies. The spectral consistency of the model across the VIS-NIR-SWIR regions is shown in Figures 5.38a and 5.38b.

(a) R^2 (b) RMSE
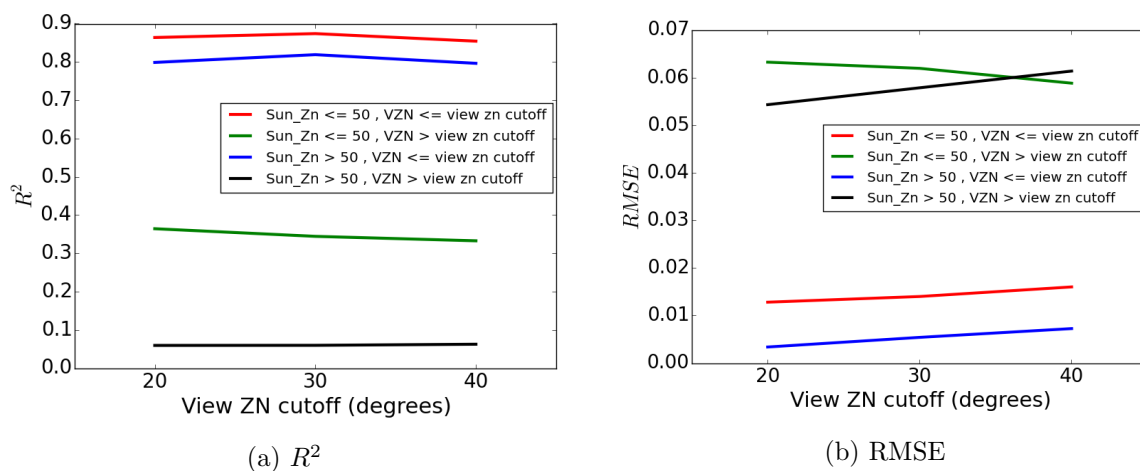
Figure 5.37: Regression coefficient (R^2) and RMSE for different view and solar zenith angles. Figure (a) shows the R^2 of the fit using the measurements whose solar zenith angle is either less than $50°$ (red, green) or greater than $50°$ (blue and black) and view zenith angles is lesser (red,blue) or greater (green,black) than the cutoff view zenith angles. Figure (b) shows the corresponding RMSE from the model fit.

From the sensitivity analysis, it is concluded that the RossLi model approximates the

measurements very well for view zenith angles less than 40° and for nominal solar zenith angles (< 50°). The RossLi model may be applicable even for larger view zenith angles, but could not be correctly validated in this research. Since many remote sensing sensors including the two sensors of interest in this research (OLI and MSI) have smaller field of view (< 20°), the RossLi BRDF model is more than adequate to model the canopy BRDF.

(a) R^2 (b) RMSE
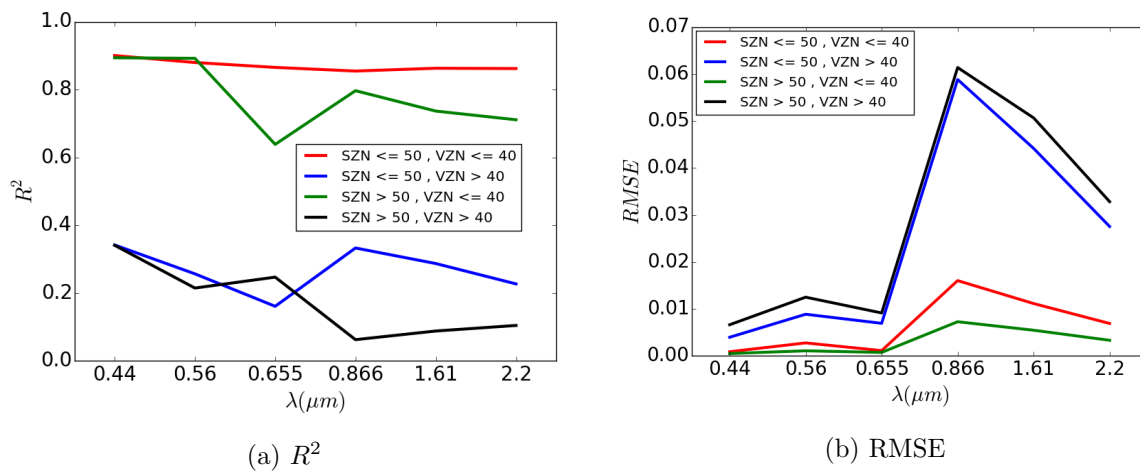
Figure 5.38: Regression coefficient (R^2) and RMSE for different view and solar zenith angles as a function of wavelength. Figure (a) shows the R^2 of the fit using the measurements whose solar zenith angle is either less than 50° (red, blue) or greater than 50° (green and black) and view zenith angles is lesser (red,green) or greater (blue,black) than 40° . Figure (b) shows the RMSE from the model fit. RMSE varies with increase in reflectance, but the R^2 is uniform for low solar and view zenith angles.

5.4.1.3 Sensitivity of BRDF for high solar zenith angles

It is important to assess the effect of the BRDF model for high solar zenith conditions since the measurements did not fit the RossLi model as accurately as for the low solar zenith conditions. Since most of the remote sensing sensors observe the ground targets at low solar zenith angles, the effect is limited to the diffused skylight contributions. This effect is studied by estimating the contribution of the downwelled radiance to the total sensor-reaching radiance, as discussed in the Section 4.4.1.3.

Figure 5.39a shows the contribution of downwelled radiance to the total radiance reaching the sensor for a nadir view sensor at an altitude of 705 KM, and observed using mid-latitude summer atmosphere with rural aerosol, and visibility of 20 KM on a summer solstice day over the Harvard forest. The modeled BRDF for low zenith and high zenith angles are used with the MODTRAN propagated atmospheric attenuation parameters to estimate the downwelled and

sensor-reaching radiance as shown in Equations 4.15 and 4.15. The total downwell contribution is found to be no more than 12% and it is about 6% when only the skylight contributions from high solar zenith angles are considered. When the atmospheric visibility is reduced to 10 KM, the skylight contribution increases marginally to 7% as shown in Figure 5.40b. Therefore, the expected error in total sensor reaching radiance due to the error in modeling the BRDF for large SZN angle is still very small. For example, if the error in the BRDF fit is 0.5% (0.0025 in reflectance units, see Figure 5.28) for low SZN angles and about 6% for high SZN angles, the net error in the BRDF modeling is about 0.88% (0.93*0.5% + 0.07*6%). Thus, the modest errors (high solar zenith BRDF) in a small term (downwelled) result in low overall errors in estimating the total sensor reaching radiance.

(a) Total skylight contribution

(b) Skylight contribution for high SZN angles
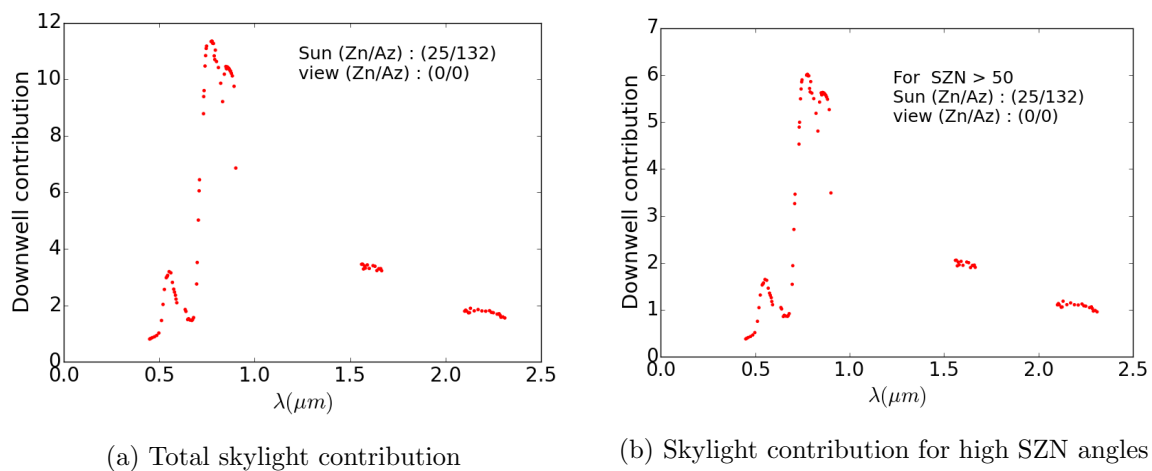
Figure 5.39: Contribution of downwelled (skylight) radiance to the total sensor-reaching radiance for a mid-latitude atmosphere with rural aerosol and 20KM visibility on a summer solstice day over Harvard forest. Figure (a) shows the total downwelled radiance contribution and Figure (b) shows the contribution only for the skylight whose solar zenith angles are more than 50°.

(a) Total skylight contribution

(b) Skylight contribution for high SZN angles
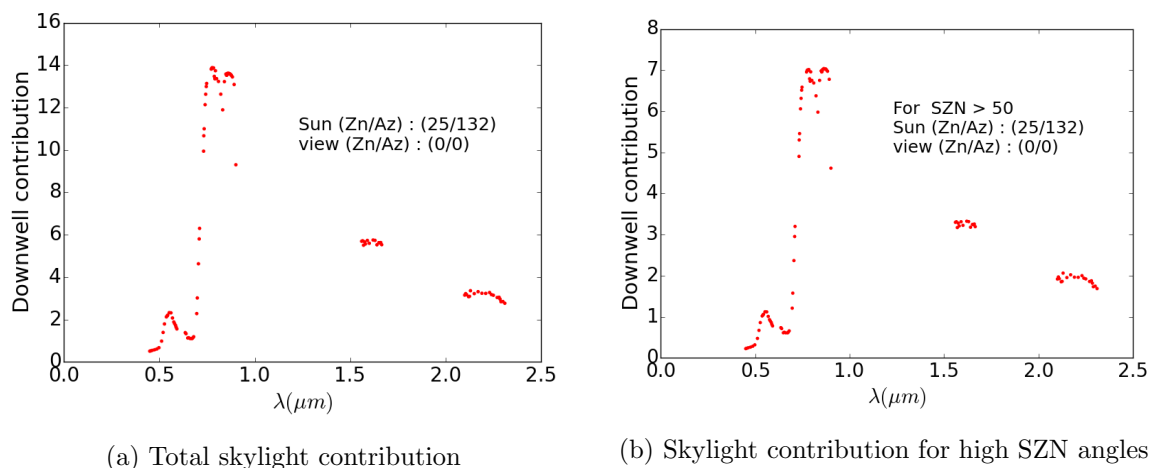
Figure 5.40: Contribution of downwelled (skylight) radiance to the total sensor-reaching radiance for a mid-latitude atmosphere with rural aerosol and 10KM visibility on a summer solstice day over Harvard forest. Figure (a) shows the total downwelled radiance contribution and Figure (b) shows the contribution only for the skylight whose solar zenith angles are more than 50°.

5.4.2 Auxiliary BRDF

5.4.2.1 Terrain slope BRDF models

The terrain slope BRDF is generated, as discussed in section 4.4.2.1, for five different slopes ($-25°, -15°, 0°, 15°,$ and $25°$). The differences between the five BRDFs are shown in Figure 5.41 for the principal and cross plane directions. The extent of view angles in the plot are limited to the nominal angles expected for medium resolution remote sensing sensors such as OLI and MSI.

(a) Principal plane
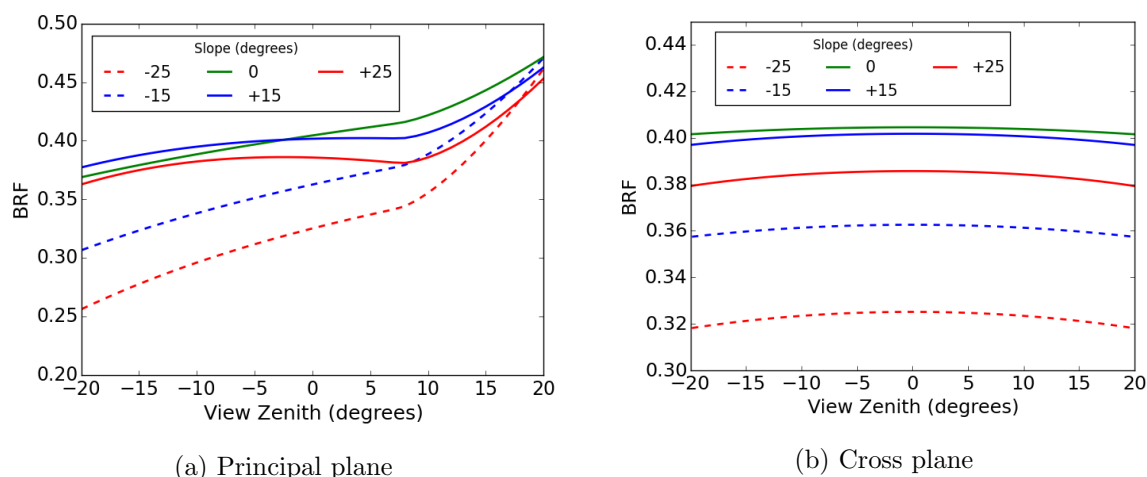
(b) Cross plane

Figure 5.41: Principal and cross plane BRF for the five terrain slopes modeled in DIRSIG for $\lambda = 0.866\mu m$. There are measurable differences in the BRF for different terrain slopes.

It is observed from the principal and cross plane BRF plots that the reflectance varies as high as 20% at nadir (8% in reflectance factor units) when compared with a flat terrain BRDF. The differences are small in the principal plane near the back-scatter direction (VZN = 35°). However, in the forward scattering direction, the differences are very high when the terrain is oriented away from the sun (for slope = −25°, −15°). This is expected since the trees at the top of the slope casts long shadows on to the trees near the bottom of the slope when they are oriented away from the sun. This also results in reduced reflectance (due to shadows) as is evident in the plot. In the cross plane, the reflectance differences are somewhat uniform in both the forward and back-scatter directions for different slopes, but the reflectance varies by about the same magnitude as in the principal plane.

The plot shows very high differences in the principal and cross plane BRF for terrain slopes, but the real differences for a typical remote sensing sensor could be small as it may not image in either plane. To estimate the variations for Landsat like sensors, relative differences between the different slope BRDFs with respect to the flat terrain BRDF can be determined using Equation 5.2. The different sun and view angle combinations are shown in Table 5.11. For each sun angle, there are 8 view angles and in total, there are 72 reflectance values computed from the sun and view angle combinations.

$$
\begin{aligned}
\Delta\rho_{ij} &= 100|\frac{\rho_{ij} - \rho_{(i=0)j}}{\rho_{(i=0)j}}| \quad \forall \quad i \in \{-25, -15, 0, +15, +25\} \quad , \quad j \in \{1, 2, \ldots, n\} \\
\mu_{\rho_i} &= \frac{\sum\limits_{j=1}^{n} \Delta\rho_{ij}}{n} \\
\sigma_{\rho_i} &= \sqrt{\frac{\sum\limits_{j=1}^{n}(\Delta\rho_{ij} - \mu_{\rho_i})^2}{n-1}}
\end{aligned}
\tag{5.2}
$$

where,

n is the number of sun and view angle combinations

The relative differences for a specific sun angle (SZN = 35°, SAZ = 127°) for different view angles are shown in Figure 5.42a. It is clear that the relative variation for the negative terrain slopes (−25°, −15°) are consistent in magnitude with the principal and cross plane directions. The positive terrain slopes (+15°, +25°) induce very small relative differences as they are oriented towards the sun, similar to the effect observed in the principal plane of Figure 5.41. Near the back-scatter direction (appx. $VZN = 17°, VAZ = 90°$), the relative variation is large for the

positive slopes but small for the negative slopes. The relative differences of the positive slopes vary inversely with the negative slopes for different view angles, as observed in the shape of the curve (see Figure 5.42a).

Figure 5.42b shows the mean and standard deviation for the different terrain slopes for all the sun and view angle combinations (72 BRF). The large standard error indicates that the reflectance variations are highly dependent on the solar and view angle conditions. From these plots, it can be concluded that the reflectance varies as a function of terrain slopes and the magnitude of the effect is dependent on how the terrain slope is oriented with respect to the sun and view angles, i.e, the slope of the terrain can introduce major differences in the BRDF.

(a) Different view angles

(b) Average over all sun and view angles
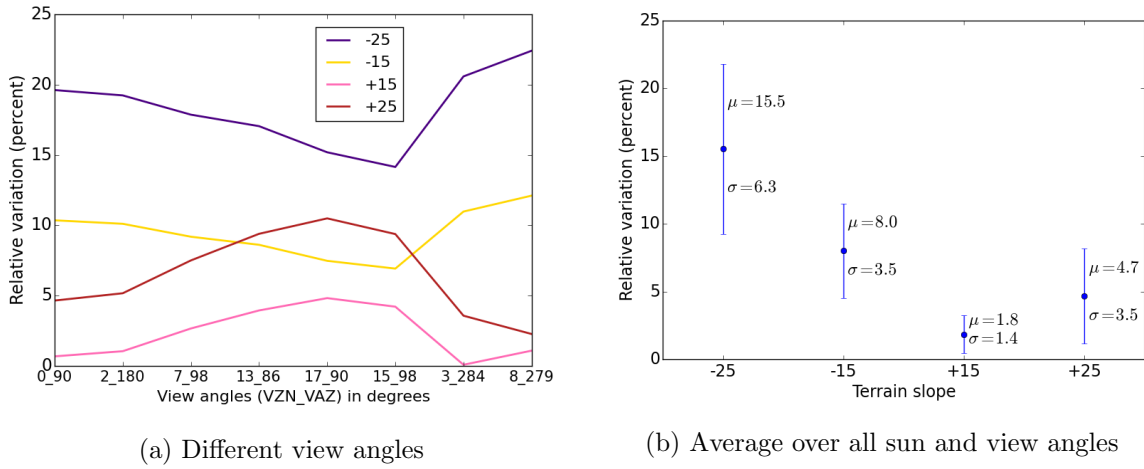
Figure 5.42: Relative differences for terrain slope BRDFs with respect to a flat terrain BRDF for $\lambda = 0.866\mu m$. Figure (a) shows the relative differences for a set of view angles (SZN=35°, SAZ=127°). Figure (b) shows the relative differences averaged over different sun and view angles

Since it is difficult to measure the terrain slope BRF, an attempt is made to approximate its effect by cosine functions. The principal and cross plane BRFs for a terrain with no slope is used as the reference. The relative reflectance of the reference to itself is unity, as shown in Figure 5.43 (green line). Relative reflectance of two terrain slopes ($-15°and-25°$) to the reference are shown in solid lines and their approximation with two simple cosine functions (cos(slope) , cos(slope)*cos(view angle)) are shown in dotted and dashed lines. Neither functions can approximate the effect of the terrain slopes as the dotted or dashed lines could not reproduce the pattern similar to the solid line (terrain BRF), and their magnitude also differs. Therefore, adjusting the BRDF for different terrain slopes requires better understanding of the changes in the radiative transfer of the canopy, and it cannot be approximated by simple functions. More

research is needed in this domain and the capability that the DIRSIG provides can be used in the future to derive a semi-empirical model to account for terrain slope or validate a physics based model, if available.

(a) Principal plane
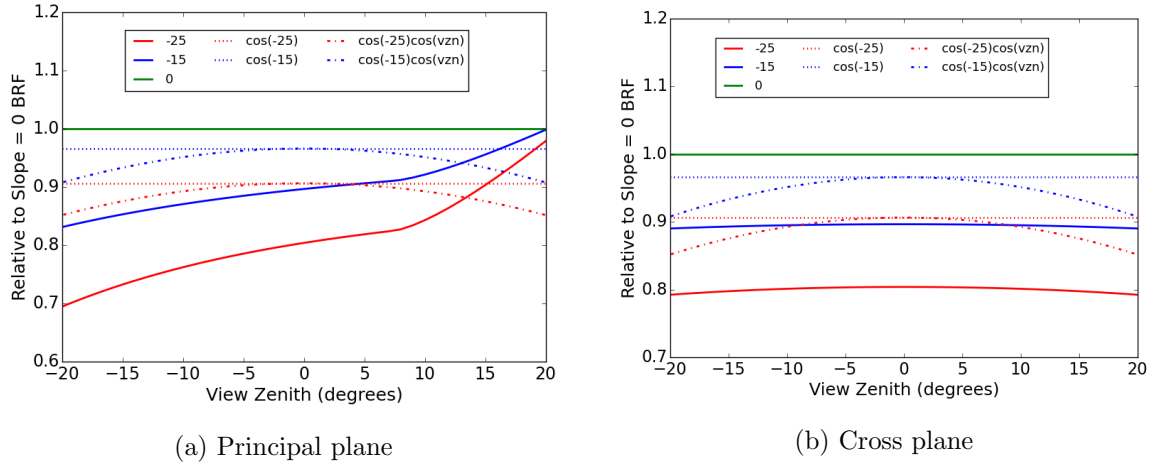
(b) Cross plane

Figure 5.43: Characterizing the principal and cross plane BRF for the terrain slopes modeled in DIRSIG for $\lambda = 0.866\mu m$. The solid line indicates the relative reflectance to the reference (no slope BRF) and dotted line and dashed lines indicate the two cosine functions. The effect of the terrain slopes could not be approximated by simple cosine functions.

5.4.2.2 Random BRDF models

The method used to generate random BRDFs was shown in Section 4.4.2.2. Five sections of the forest (Center, UL, UR, LL, LR) each with an overlap of at least 80% are used (see 5.44) to limit the extent of the RossLi coefficients. The corner section overlaps with the center section by 81% and the two adjacent corner sections overlap by 80%. The large overlap area is required to limit the amount of variability within the generated BRDFs. The advantage of using different sections is that it approximates the textural variation observed in a deciduous canopy. The mean from the center forest section and the deviations of the five forest sections in the PCA transformed space is used to define the mean and STD for the normal distribution. Using the normal distribution, several random samples are chosen in the PCA basis and are transformed back to the original basis to generate the random BRDF models. In this research, 32 random BRDF models are generated for two different terrain slopes ($\pm 25°$).

The relative variation for the 32 random BRDF models for each slope is calculated using Equation 4.16 . Analysis from the MODIS and Landsat-8 data products showed that the natural variability within the Harvard forest is less than 5%. In the case of simulated forest, the relative

variation (when compared with the central section of the forest) for all the randomly generated BRDFs is observed to be less than 4% for both $+25°$ and $-25°$ slopes as shown in Figure 5.45. The standard deviation between the 32 random BRDF models for each slope is found to be less than 3%. Thus, the variations within the random BRDF models are in good agreement with the variations expected in the real forest. The technique used to generate random BRDF models can be very useful to describe the diversities found in real-world deciduous forests.
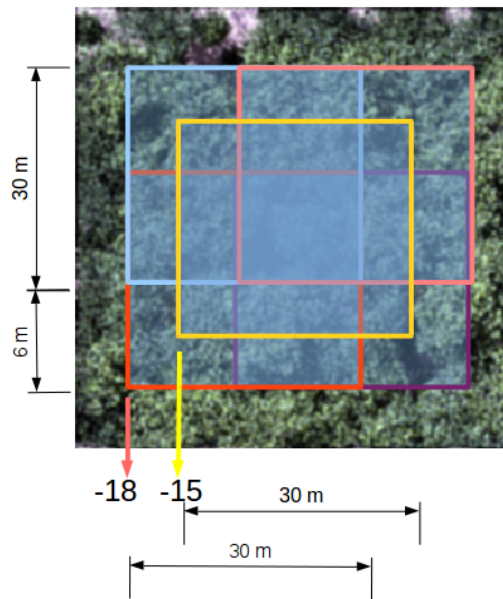
Figure 5.44: Five different sections of the forest (Center, UL, UR, LL, LR) are used to limit the bounds of the RossLi coefficients for generating random BRDFs using PCA technique. The corner section has an overlap of 80% with other corner section and 81% overlap with the center

(a) Slope $= -25°$

(b) Slope $= +25°$
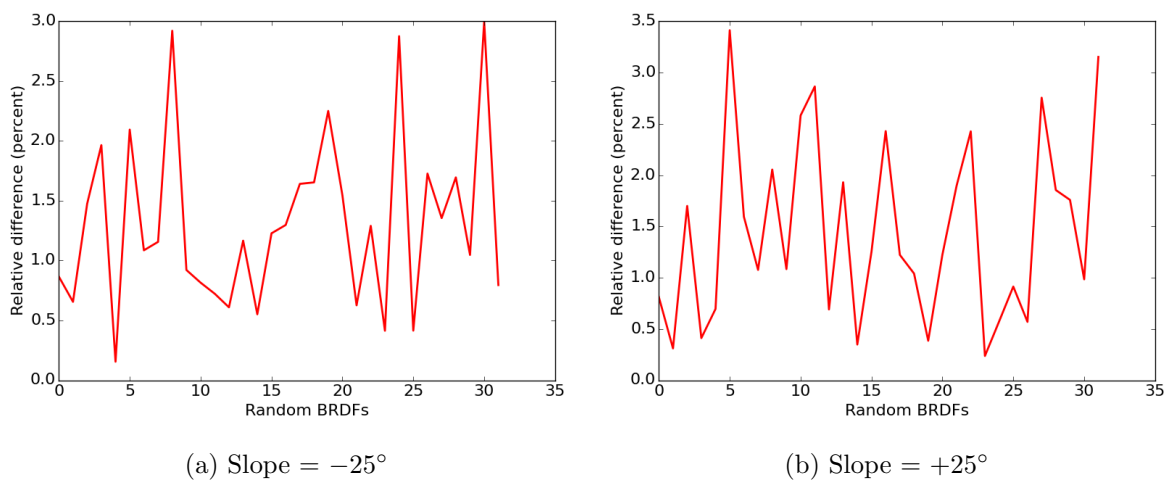
Figure 5.45: Relative variation for the simulated random BRDFs for two terrain conditions in comparison to the central forest section for $\lambda = 0.866\mu m$. The variation for all the randomly generated BRDFs is found to be less than 4% for both cases.

5.5   Modeling forest canopy signal

The methods used in modeling the forest canopy signal was explained in Section 4.5. In this section, the results, analyses and the characterization of the defoliated forest for the various data products are discussed.

5.5.1   Defoliation of Harvard forest

The modeling of defoliated forest in DIRSIG can be useful in approximating the natural process. In this research, nine different levels of defoliated forest BRDFs are generated. Figure 5.46 shows the RGB image for six of the nine defoliated forests. The figure demonstrates the variation expected within a forest for different levels of defoliation. Even at 40% defoliation, the visual image does not show any remarkable differences. This is because, the LAI of the forest even at 40% defoliation is about 3, which in general is high, and secondly, the leaf facets are removed randomly (i.e, entire leaf is not removed). As a result, it is possible to observe similar number of leaves but with smaller leaf sizes. Further, the twigs and branches along with the leaf facets make the image highly cluttered and fail to show any visual differences.

To verify the effect of defoliation, the LAI can be measured for each defoliated forests. The LAIs observed in Figure 5.47 are computed from the mean LAI for the entire forest extent (30m x 30m). The actual LAI within the simulated forest varies as much as the mean, and is dependent on the type of tree species, its distribution and orientations, as is the case with the real forest. In this research, the distribution and the orientation of the trees in the simulated forest are based on the random sampling method (see 4.2.2), and the tree species are modeled to have an LAI similar to the actual Harvard forest's tree species. Variability exists within the 30m scene extent, but the sensor's IFOV is of the same order (20m for MSI and 30m for OLI), so the mean LAI is a good indicator of the forest's LAI.

(a) 0% (LAI = 5)   (b) 20% (LAI = 4)   (c) 40% (LAI = 2.9)

(d) 55% (LAI = 2.1)   (e) 70% (LAI = 1.3)   (f) 85% (LAI = 0.7)
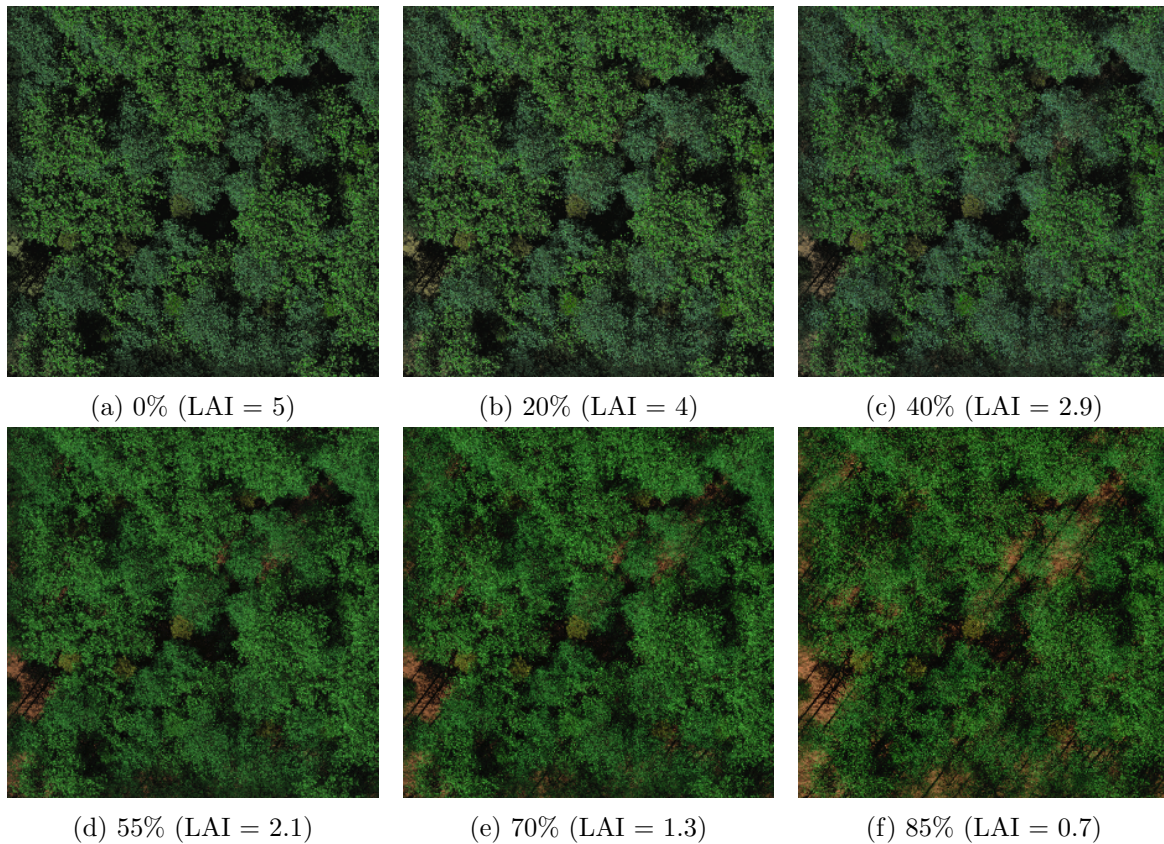
Figure 5.46: Visual (RGB) image of the defoliated forests modeled in DIRSIG. The images are simulated from a nadir view when the sun is located at 35° zenith and 127° azimuth. The reduction of leaf clutter in the images are visible when the defoliation levels are high, but are less apparent for lower levels of defoliation.

Figure 5.47a shows the trend in the LAI for the different levels of defoliation. It is very clear that the LAI reduces with an increase in defoliation and it varies linearly as expected. The relative variation in LAI can be estimated by taking the ratio of the absolute difference in LAI (defoliated vs forest without defoliation) to the LAI of the forest without defoliation (see Equation 5.3). Figure 5.47b indicates that the relative change in the LAI matches one-to-one with the level of defoliation, but with a very small deviation when the levels of defoliation are high. This is negligible ($\approx 2\%$ at 85% defoliation) and does not indicate any error in the modeling of the defoliated forests.

$$\text{Relative variation in LAI} \; = \; 100 \; \frac{\left| \, LAI_i \; - \; LAI_{(i=0\%)} \, \right|}{LAI_{(i=0\%)}}$$

$$\forall \quad i \quad \in \quad \{\text{defoliation levels}: 0\%, 10\%, 20\%, 25\%, 30\%, 40\%, 55\%, 70\%, 85\%\}$$

(5.3)

(a) LAI vs Defoliation

(b) Relative variation in LAI
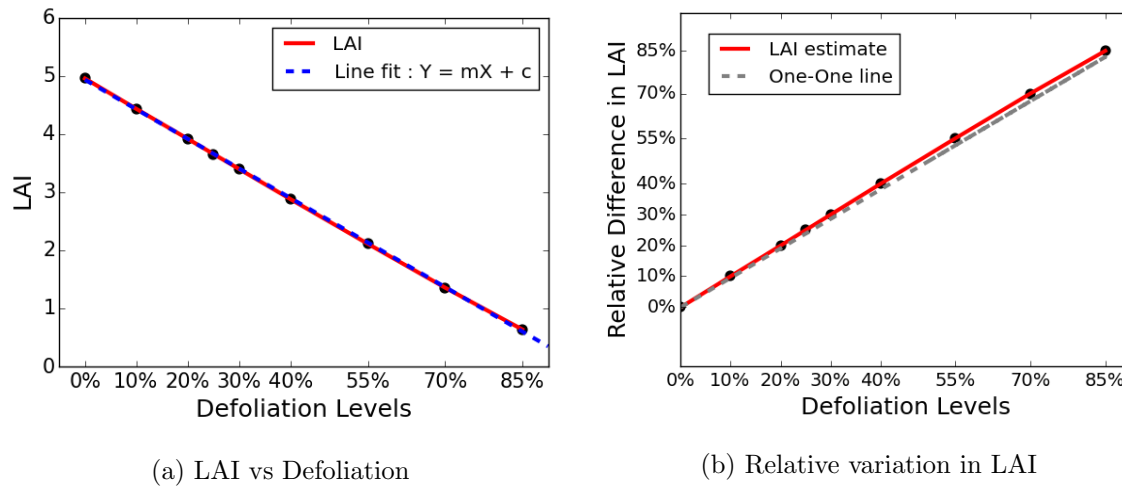
Figure 5.47: Relative variation in LAI for different levels of defoliation. The LAI varies one-to-one with the defoliation, but show a small but negligible deviation when the level of defoliations are high.

The principal and the cross plane BRFs for the different levels of defoliation are shown in Figure 5.48 for the red and NIR spectral bands. The reflectance can vary anywhere between 5% to 60% in the NIR band and even higher in the red band. The hot-spot effect is observed in the principal plane at the back-scatter direction for all the defoliated levels. The red reflectance of the canopy increases, while the NIR reflectance reduces, with the corresponding increase in the level of defoliation. This is expected, since the increase in the level of defoliation reduces the leaf area and the greenness in the canopy. It is due to the greenness of the canopy (leaf optical properties) that the reflectance is high in the NIR band and low in the red band. A similar trend is observed in the cross plane direction and the rate of increase in the reflectance is high for the corresponding increase in the LAI (or reduction in the level of defoliation). This indicates that the reflectance varies non-linearly with LAI (or defoliation).

(a) principal plane (RED)   (b) principal plane (NIR)

(c) cross plane (RED)   (d) cross plane (NIR)
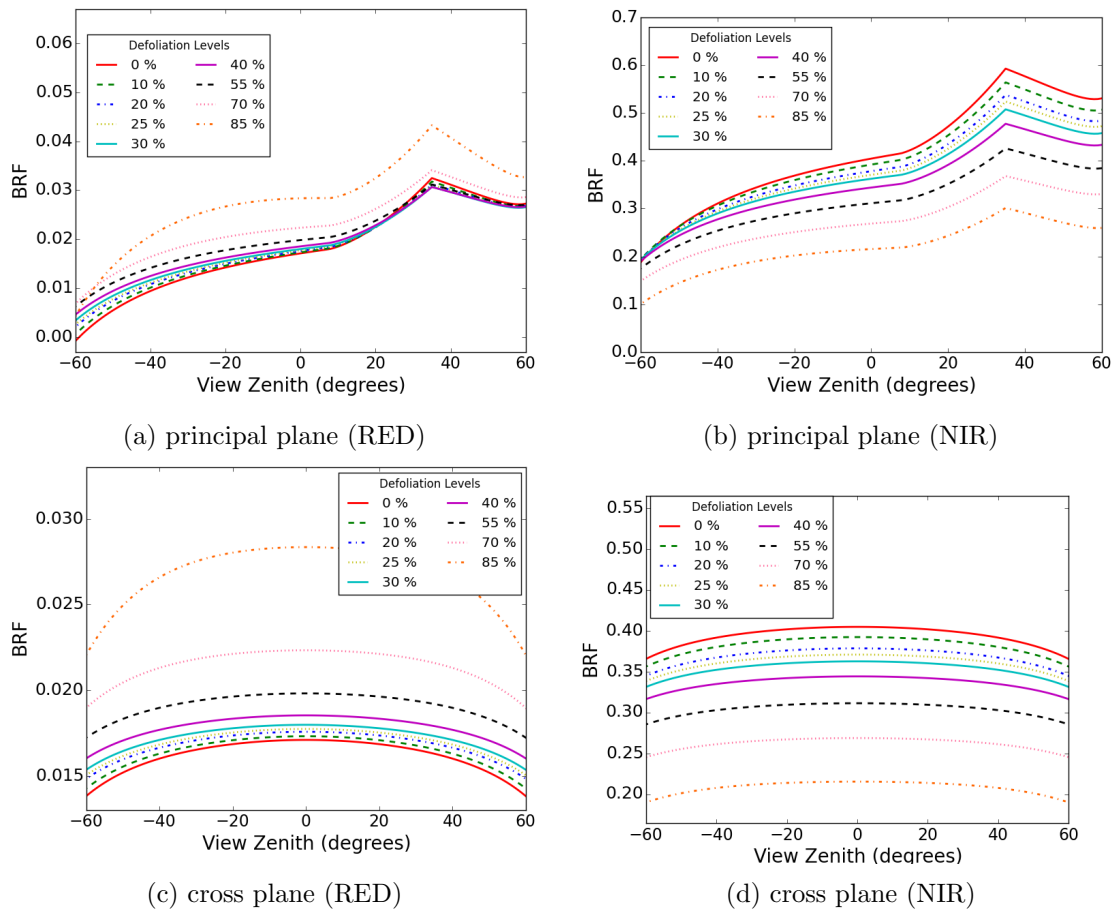
Figure 5.48: Principal and Cross plane BRFs for different levels of defoliation. The hot-spot effect is observed in the principal plane at the back-scatter direction for all the defoliated levels.

5.5.2 Signal modeling for different data products

Section 4.5.2 discussed the methods used in modeling the defoliated forest as a signal for different data products. The curve fit for modeling the signal is performed using all the 9 defoliated forests.

5.5.2.1 Radiance and reflectance products

The curve fit for at-sensor radiance and the TOA reflectance in the red and NIR spectral bands are shown in Figure 5.49. The TOA responses for the red spectral band shows that the trend increases with increase in defoliation. The relative variation increases in the red and NIR spectral bands due to the increase in the red reflectance and reduction in the NIR reflectance. Similar trend is observed for all the different data products, as expected.

(a) TOA RED

(b) TOA NIR

(c) TOA REFL RED

(d) TOA REFL NIR
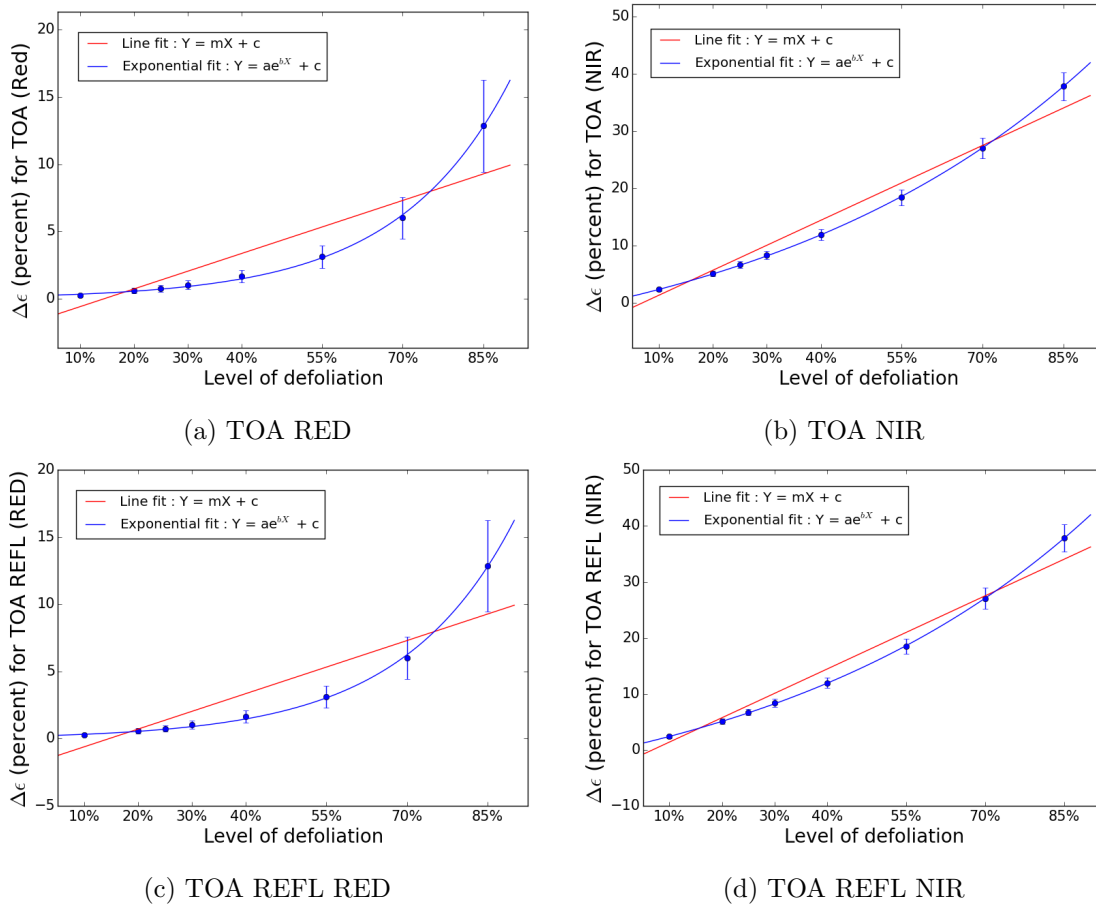
Figure 5.49: Curve-fit for defoliation vs relative variation for TOA radiance and reflectance in the red and NIR spectral bands. The TOA radiance and reflectance products shows an exponential trend.

The STD (error bar) increases with increase in the level of defoliation for both the bands, but are very high for the red spectral band. The observed STD is due to the different simulation combination in view angles, sun angles, RSRs and visibility conditions. The scattering due to the atmosphere is very high in the red spectral band compared to the NIR band, which causes a high variability in the red band. Further, the reflectance of the canopy is very small in the red band, therefore, a small change in the reflectance can introduce a large variation. The variation in the spectral response of the two sensors may also contribute to the STD, but the variations in the visibility condition and the BRDF of the canopy are the major contributors in the red band. This is evident from Figure 5.50 where, the STD is small for the surface reflectance products compared to the TOA products. This is expected since their atmospheric attenuations are compensated. Between the two ELM methods, the ideal compensation technique show smaller STD than the typical ELM compensation method. The STD in the ELM method (ideal) is very similar to the STD observed in the surface reflectance products generated from the BRDF of

the canopy (see Figure 5.51). The STD observed in the surface reflectance products are mainly due to the variation in the sun angles, i.e, the BRDF effect of the forest due to the illumination differences. The STD is observed to be very small in all the NIR band products due to the low atmospheric attenuation and high surface reflectance of the canopy.

(a) ELM-ideal RED   (b) ELM-typical RED

(c) ELM-ideal NIR   (d) ELM-typical NIR
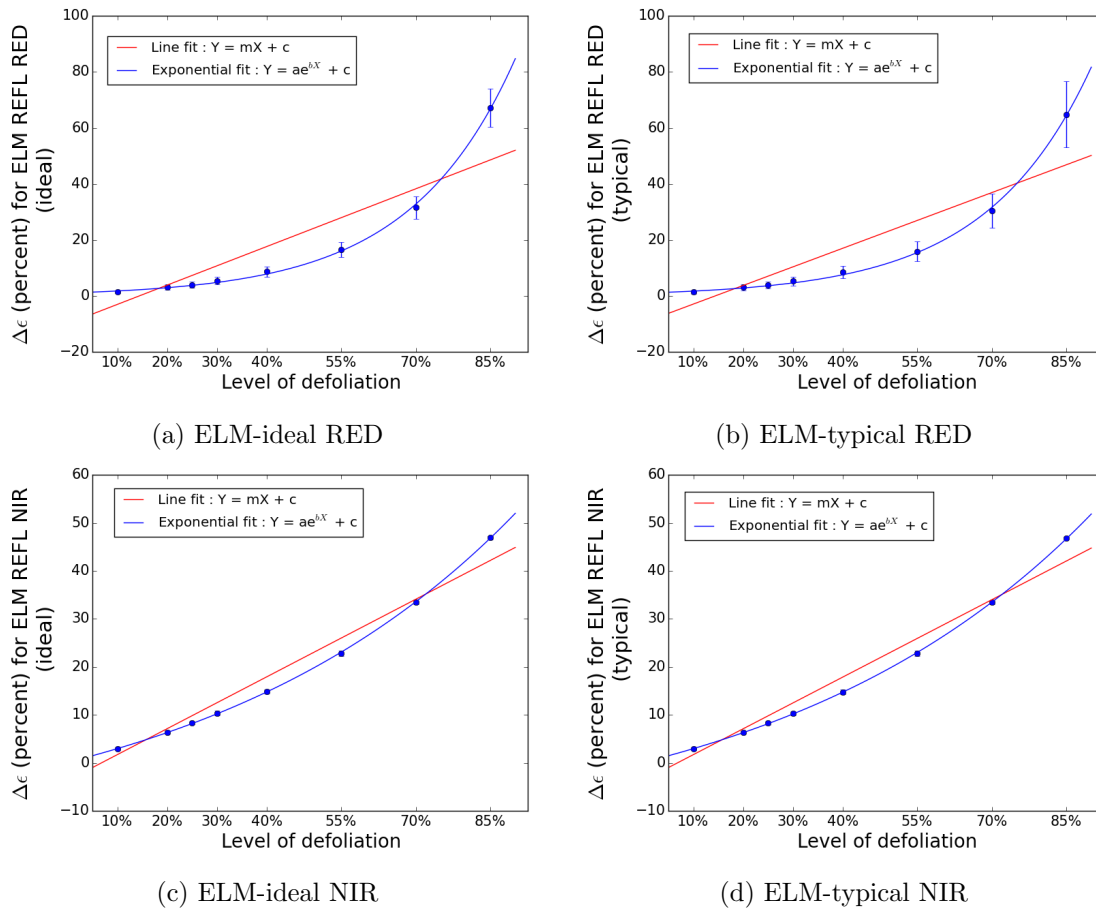
Figure 5.50: Curve-fit for defoliation vs relative variation of the two ELM compensated reflectance products in the red and NIR spectral bands. Similar to TOA reflectance and raidance, the relative variation shows an exponential trend

The relative variations for the BRF products are estimated by calculating the mean reflectance from the different sun and view angle combination (see Table 4.5) as described in Section 4.5.2.3. The results for the red and NIR spectral band are shown in Figures 5.51a and 5.51b. In total, 50 observations (10 sun angles x 5 view angles) are used to estimate the mean and STD for each level of defoliation (see Equation 4.21). The relative variation is also computed using another independent set of sun and view angles (see Table 5.11), as shown in Figures 5.51c, 5.51d. In this case, 72 samples are used to estimate the mean and STD (see Equation 4.21). The two results though generated from different sun and view angles, are consistent in their estimation

of the mean relative variation. This shows that the forest signal can be accurately characterized using the relative variation for the canopy BRDF product.

(a) BRF RED (50 samples)

(b) BRF NIR (50 samples)

(c) BRF RED (72 samples)

(d) BRF NIR (72 samples)
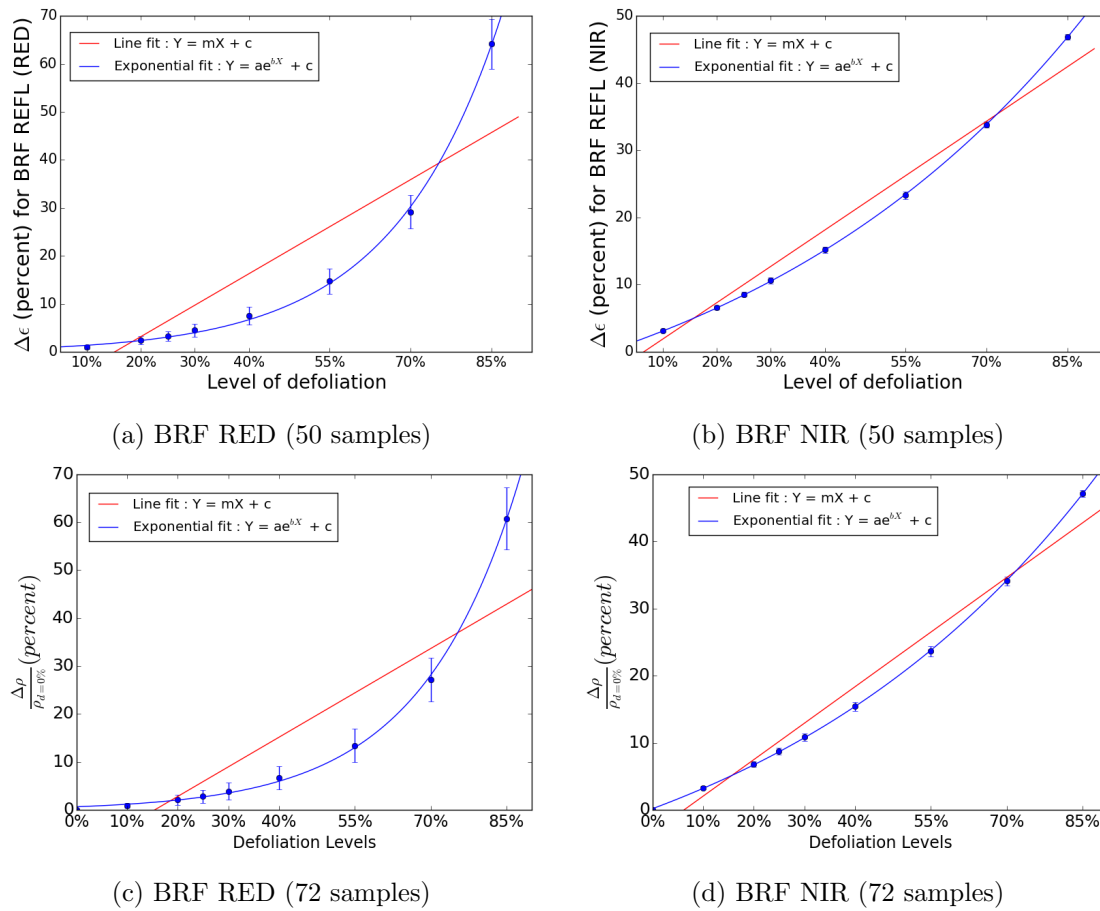
Figure 5.51: Curve-fit for defoliation vs relative variation of the BRF products in the red and NIR spectral bands. Figure (a) and (b) used 50 sun and view angle combinations to estimate the mean relative variation, while Figure (c) and (d) used 72 different samples. The exponential trend and the mean relative variations are similar in both the cases exhibiting the consistency with the forest signal estimation.

5.5.2.2   NDVI products

The radiance, TOA reflectance, and the surface reflectance (ELM-ideal, ELM-typical, BRF) products in the red and the NIR spectral bands are used to estimate the NDVI of the forest canopy. The NDVI indicates the greenness of the vegetation, in this case, the density of the canopy. As discussed earlier, reduction in the leaf area reduces the NIR reflectance and increases the red reflectance. Hence, NDVI is positively correlated with the LAI (negatively correlated with the defoliations) as shown in Figure 5.52. The NDVI in Figure 5.52 is determined from the BRDF models for each of the defoliated forests. The spectral reflectance (from the BRDF

coefficients) for a specific sun angle (SZN=35°, SAZ=127°), nadir view angle, and for the central wavelength of the red and NIR spectral bands of the OLI sensor ($\lambda = 0.866\mu m, 0.654\mu m$), are used to calculate the NDVI for each defoliated forests. An exponential function of the form ($NDVI = a - exp(-b\sqrt{LAI})$), is used to fit the measured LAI with its NDVI, but it is not based on any mathematical relationship between the two variables. Analyzing or an attempt at the theoretical formulation for these two variables would be beyond the scope of this research. In general, the NDVI is assumed to vary linearly with LAI for sparse canopies with low LAI (ranging from 1 to 3) or piecewise linear (Wang et al., 2005), but from the figure, it is observed that the relationship between the NDVI and LAI is better approximated by an exponential function over the full range of canopy densities.
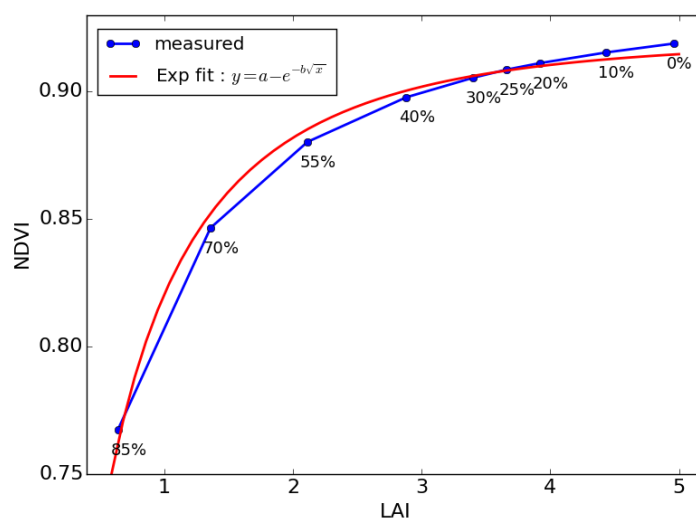
Figure 5.52: NDVI as a function of LAI for simulated Harvard forest canopy. An exponential function of the form $NDVI = a - exp(-b\sqrt{LAI})$ can be useful to model a functional relationship between the two variables.

The relative variation in NDVI for the five data products are shown in Figure 5.53. The trend for the NDVI products are similar to the trend observed for the red and NIR products. As in the case with the radiance and reflectance products, the NDVI products generated from the TOA radiance and reflectance data shows higher STD compared to the NDVI from the surface reflectance data. This is due to the variability in the atmospheric (visibility) conditions which are compensated in the ELM products. This is also observed in the STD between the ELM-ideal and ELM-typical data products. The NDVI from the ELM-ideal data shows better consistency and is closer to the NDVI calculated directly from the canopy BRDF. This indicates that a simple ELM atmospheric compensation technique, accounting for the path length variation in

transmission, has the potential to provide an accurate surface reflectance product. The relative variation for different levels of defoliation is shown in Table 5.12. The mean values are used to relate the signal with the corresponding changes in the forest for factor sensitivity analysis. Note that the STD for the different levels of defoliation for all the NDVI products is small ($< 10\%$ of the mean) which indicates that the relative variations can be useful to estimate the effects of the factors with very low uncertainty.

Table 5.12: Relative variation of the NDVI products for the different levels of defoliation. The values are used to relate the signal with the corresponding changes in the forest for factor sensitivity analysis.

Level of defoliation	NDVI products									
	TOA RAD		TOA REFL		ELM-typical		ELM-ideal		BRF	
	μ	(σ)	μ	(σ)	μ	(σ)	μ	(σ)	μ	(σ)
0%	0	(0)	0	(0)	0	(0)	0	(0)	0	(0)
10%	1.30	(0.16)	0.85	(0.10)	0.38	(0.04)	0.37	(0.03)	0.37	(0.01)
20%	2.89	(0.30)	1.89	(0.18)	0.87	(0.08)	0.84	(0.04)	0.82	(0.03)
25%	3.80	(0.38)	2.50	(0.22)	1.14	(0.10)	1.10	(0.04)	1.100	(0.03)
30%	4.86	(0.47)	3.19	(0.27)	1.50	(0.12)	1.45	(0.05)	1.44	(0.04)
40%	7.26	(0.65)	4.78	(0.37)	2.33	(0.17)	2.26	(0.07)	2.27	(0.05)
55%	12.18	(0.94)	8.06	(0.53)	4.25	(0.26)	4.14	(0.08)	4.17	(0.07)
70%	20.16	(1.26)	13.47	(0.69)	7.97	(0.38)	7.80	(0.10)	7.85	(0.10)
85%	34.07	(1.27)	23.14	(0.60)	16.63	(0.50)	16.39	(0.27)	16.51	(0.21)

In all the data products discussed above, an exponential trend is observed for an increase in the level of defoliation. The exponential function uses 5 redundant observations to estimate the 3 parameters from a total of 9 observations. One of the observations (defoliation = 25%) is removed from the fit to estimate the residual error between the measurement and the fit. The fit residuals for most of the data products are shown in Table 5.13. The residuals are high for the red spectral band compared to the NIR band and this causes an increase in the fit residuals for the NDVI products. The large error in the red band is expected due to the low reflectance of the canopy. The exponential curve fit for all the data products indicate that the level of defoliation can be accurately predicted from the relative variation (change detection) to within 10% uncertainty for the NDVI based products and to less than 2% for the NIR products.

(a) TOA RAD NDVI

(b) TOA REFL NDVI

(c) ELM-ideal NDVI

(d) ELM-typical NDVI

(e) BRF NDVI
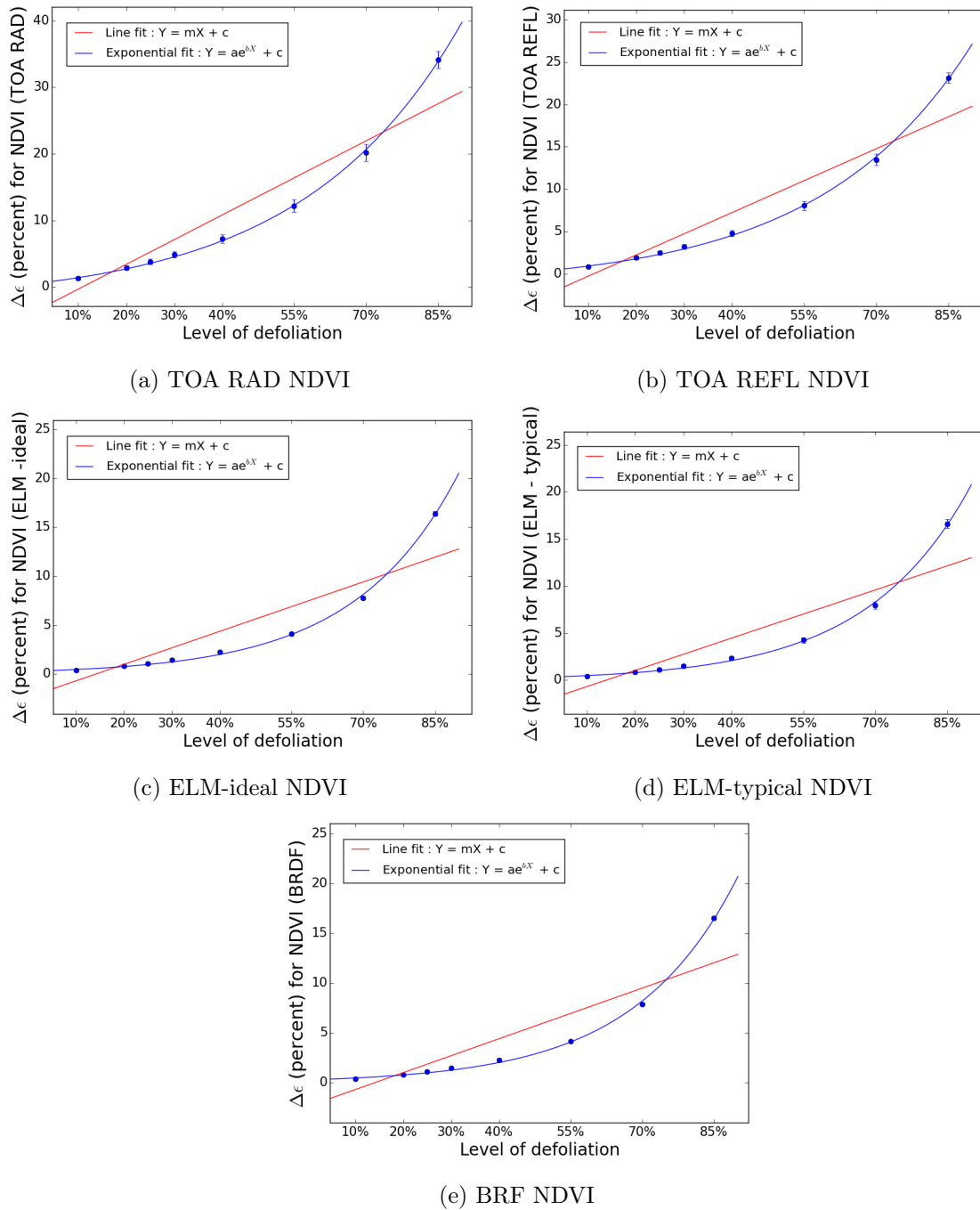
Figure 5.53: Curve-fit for defoliation vs relative variation for the NDVI products. Figure (a),(b),(c) and (d) are NDVI products calculated from TOA radiance, TOA reflectance, ELM-ideal reflectance and BRF data. All the NDVI products are accurately characterized by the corresponding exponential function.

Figure 5.54 shows the NDVI calculated for the different data products as a function of defoliation for the moderate defoliation levels (LAI : 3-5). The trend is very similar across all the data products, however, their slopes are not. It is clearly evident from this plot that the NDVI varies depending on the type and level of compensation methods. For radiance data, the predicted

Table 5.13: Fit residuals between model and measured relative variation for 25% defoliation signal level. The red band shows consistently higher residuals than NIR

Data products	RED relative error	NIR relative error	NDVI relative error
TOA Rad	6.1%	1.5%	6.9%
TOA Refl	6.1%	1.5%	7.4%
ELM-typical Refl	6.3%	1.5%	10.9%
ELM-ideal Refl	6.2%	1.5%	10.9%
BRF Refl	5.3%	0.9%	10.6%

NDVI of the forest is in the range of (0.55-0.65), whereas, the TOA reflectance products which are compensated for the cosine effect of the solar zenith angle, show an improvement over the radiance data (0.7 - 0.75). The atmospherically compensated data using both the ELM methods match very closely to the surface reflectance data (dashed black line), and the ELM-ideal (magenta) method outperforms the ELM-typical method as expected. It is important to note that the applications that use NDVI can underestimate the greenness of the vegetation depending on the compensation uncertainty and the type of data products used to estimate the NDVI. This is well recognized by the remote sensing community where consistency in the method of generation of the NDVI and its relative variation within that consistent methodology are used to study the canopy conditions.
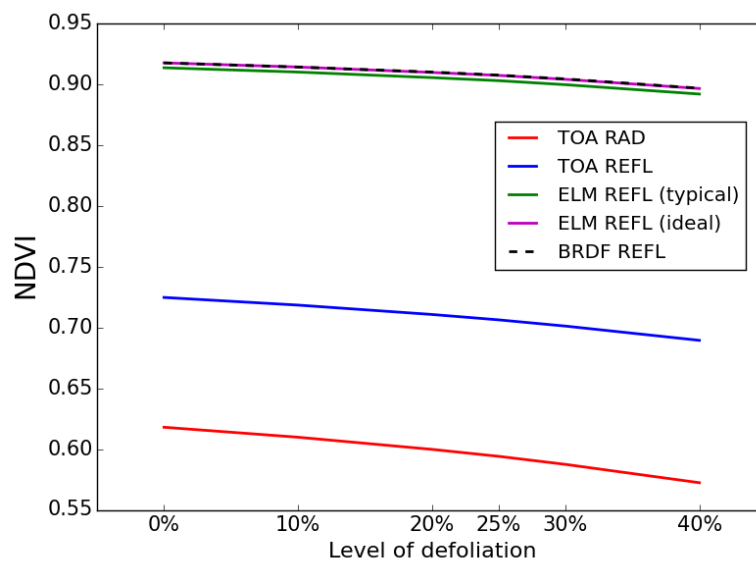
Figure 5.54: NDVI calculated for the different data products as a function of defoliation. The NDVI for the TOA corrected products are lower than the NDVI for the surface reflectance products. This indicates that an accurate NDVI value can be estimated only when compensated for atmospheric attenuations.

From these analyses, we conclude that the forest reflectance varies with LAI or the level of defoliation. The trend is similar and consistent across all the data products evaluated in this section. Further, the relative variation can be accurately characterized with an exponential function and can be directly used as a metric to estimate the effects of other factors. It is to be noted that any application that predict or estimate the biophysical parameters from NDVI needs to ensure that the data are well calibrated and compensated for the effects of environmental and sensor factors.

5.6 Factors' effect analysis

The methods used to analyze the effects of the sensor and environmental factors were discussed in Section 4.6. In this section, the sensitivity of the factors are analyzed one-at-a-time, i.e, the effects of the factor is studied by varying one factor while keeping all the other factors at a constant level. This type of analysis cannot be performed with the real-world dataset as the factors cannot be controlled independently. In the previous section, the characterization of the forest signal using relative variation for different levels of defoliation was discussed. Although the signal was characterized for all the 9 defoliation levels (up to 85% defoliation), the factors are analyzed only for the 6 signal levels (0% to 40%). This is because, the primary focus of this research is to estimate the effects of the factors for a typical deciduous canopy (LAI between 3 to 5). Secondly, the effects are estimated independent of the signal level and the characterization of defoliation as a signal is used only to represent the effects in terms of change detection units. Hence, the signal levels in itself should not affect the estimated effect, unless there exists a relation between the signal level and the effects, which can be sufficiently inferred from the 6 signal levels. Furthermore, the characterization of defoliation shows a linear trend when the levels of defoliation are less than 40%. The effects for the four factors: RSR, across-track, visibility, and solar zenith angle are described in this section.

5.6.1 RSR effects

The two sensors (OLI and MSI) have different spectral responses (RSRs), especially in the red spectral band which also affects the derived products such as the NDVI. Therefore, the analysis for RSR effects are performed using the NDVI products.

The effects are estimated by simulating the TOA radiance for the two different RSRs under varying view angles, sun angles, and visibility conditions. The relative variation in the response is estimated by observing the differences under the same view, sun and visibility conditions between the two sensors and hence, no other effects are included. This is equivalent to a scenario where the OLI and the MSI sensors are imaging the same target on the ground (same BRDF), at the same time (same SZN angles and visibility conditions), and at the same view angle. So, the only difference in their observed radiance is due to the differences in the RSRs. Similar simulations can also be performed at different sun angle, view angle and visibility condition, which is equivalent to an observation for a different day.

In this case, 10 sun angles, 5 view angles and 4 visibility conditions (200 observations) for each signal levels are used to estimate the mean and the STD of the relative variations ($\Delta\epsilon$) for all the NDVI products, as shown in Figure 5.55. The NDVI product that is generated from the BRF of the canopy is not shown, as it based on the actual surface reflectance and not from any sensor observations (i.e, no RSR effect). The high mean with a relatively low STD in Figure 5.55 provides more confidence in the estimation of the RSR effect. The STD is due to the variability in the observation conditions and are high for the TOA products (RAD and REFL) as they are not compensated for the atmospheric attenuations. Between the two ELM methods, the ideal compensation method shows much smaller STD and this indicates that the variability (STD) in the effect estimation is mainly due to the changes in the visibility conditions. In each of the NDVI products, the STD is approximately equal across the 6 signal levels, indicating that the signal level does not affect the variability in the estimation of the RSR effect. The high STD in the TOA and ELM-typical products makes it difficult to ascertain the trend between the RSR effect and the level of defoliation. However, when the variability is reduced using the ideal atmospheric compensation method, an increasing trend in the effect with an increase in defoliation is apparent. This indicates that the effect of RSR can be high when the canopy reflectance is low for well calibrated and compensated products.

(a) TOA RAD NDVI

(b) TOA REFL NDVI

(c) ELM-typical NDVI

(d) ELM-ideal NDVI
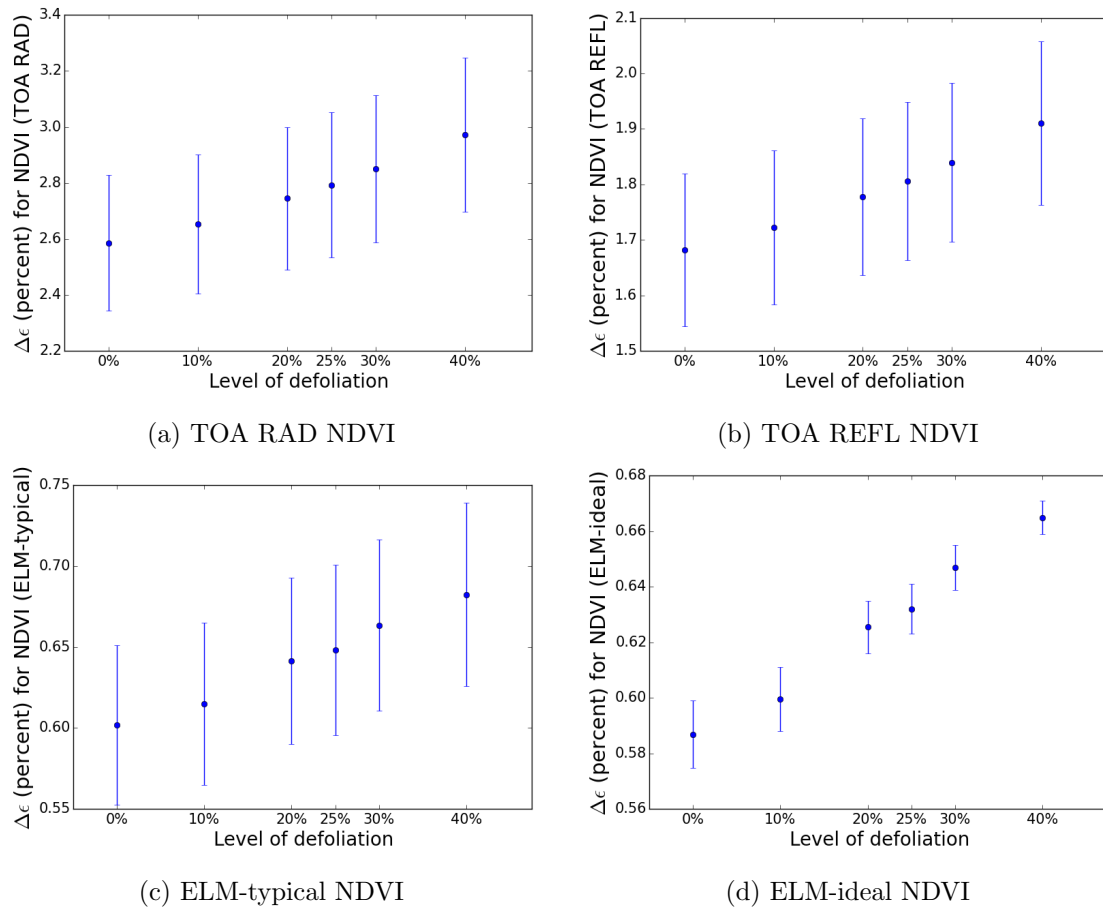
Figure 5.55: Effect of RSR for different NDVI products

In some applications, the effects due to the RSR are compensated using SBAF techniques as discussed in Section 4.6.1. The SBAF methods can be based on the TOA reflectance or the canopy reflectance (BRF). In either case, the at-sensor responses are compensated by the corresponding SBAF values. The SBAF values for the TOA REFL and BRF methods are calculated to be 1.07 and 1.05 respectively.

As in the previous case, the effects are estimated between the OLI response and the SBAF compensated MSI response. The results for both the methods are shown in Figures 5.56, 5.57, 5.59, and 5.60 for the four NDVI products. The SBAF using the TOA REFL method performs better (small $\Delta\epsilon$ and STD) than the BRF method for the TOA radiance and reflectance NDVI products (see Figures 5.56 and 5.57). This is not surprising as the SBAF value was estimated from the TOA reflectance and are likely to correct the TOA products better than the surface reflectance products. Similarly, it is observed from Figures 5.59 and 5.60 that the performance of SBAF using BRF performs better for the surface reflectance products.

For example, if the application uses only TOA products to derive all its information, then SBAF TOA-REFL may provide the best compensation for the RSR effect, but if the atmospheric compensated products are used, then the SBAF correction based on the surface reflectance of the target is more useful. Thus, it may be better to choose the appropriate SBAF techniques depending on the type of products used in the application.

For either SBAF methods, it is seen from Figures 5.56 and 5.57, that the products with no compensation (TOA radiance) has the highest effect and it reduces when the SZN angle is compensated (TOA REFL). The effect reduces for additional compensation (ELM-typical, ELM-ideal) and is lowest for the ideal compensated product (ELM-ideal). This indicates that the effect of RSR is dependent on the accuracy of the compensation techniques.
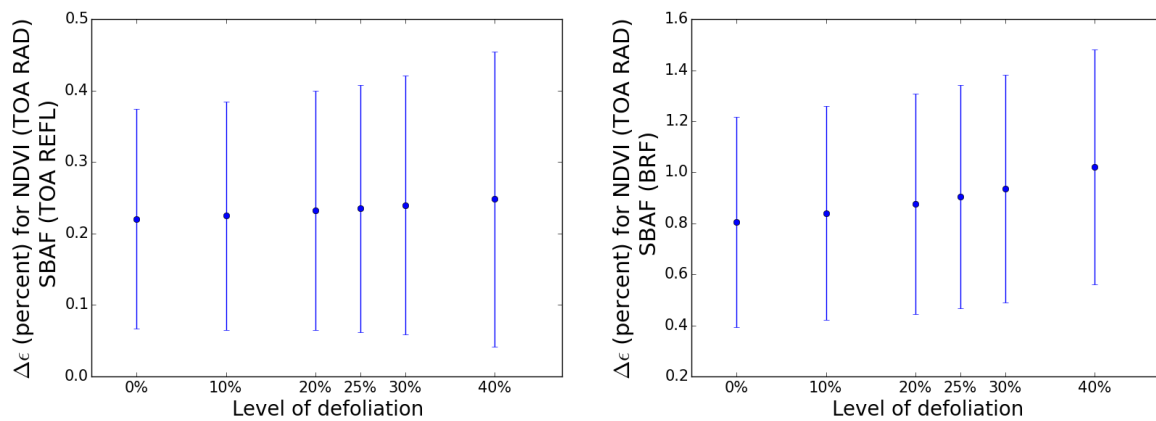
(a) SBAF using TOA REFL   (b) SBAF using BRF

Figure 5.56: Effect of RSR after the SBAF compensation for TOA RAD NDVI product
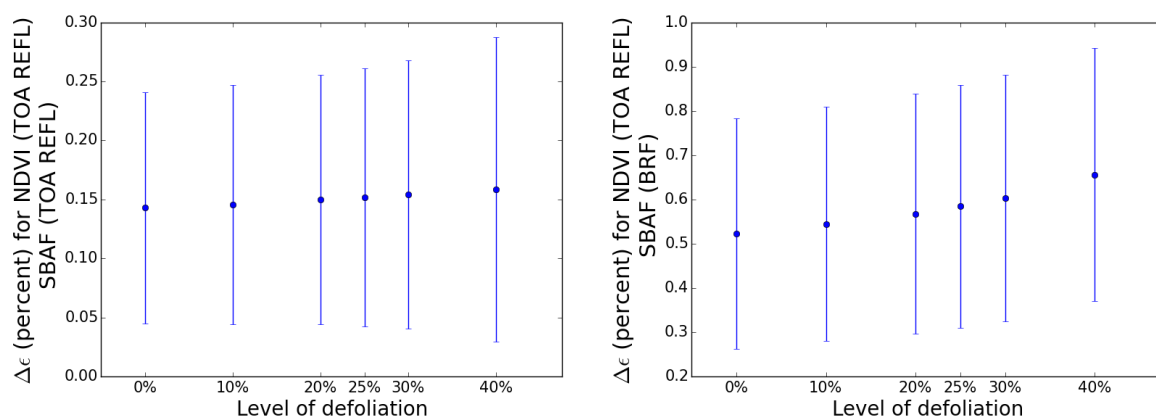
(a) SBAF using TOA REFL   (b) SBAF using BRF

Figure 5.57: Effect of RSR after the SBAF compensation for TOA REFL NDVI product

Unlike the RSR effects in the uncompensated data, no specific trend is observed between the effect and the signal level in the compensated data. This is because, the residual effect after compensation is very small, so that their within-variance (due to visibility, sun and view) is as high as the variation across the signal level. This indicates that the residual effect after compensation is likely to be independent of the signal level, i.e, the SBAF compensation technique performs equally well across different signal levels.

The effect of RSR for the different NDVI products are shown in Table 5.14. The mean relative variation is computed by averaging the effect over all the signal levels. The mean relative variation for each NDVI product is then converted to NED units by interpolating the relative variation to its corresponding level of defoliation based on the signal characterization (see Table 5.12). This is illustrated in Figure 5.58, where the estimated mean relative variation (green line) is used to find its corresponding defoliation level (blue line) from the characterized defoliation curve (red line) to determine the effect in NED units (14%). It is clearly evident from Table 5.14 that the effect of RSR depends on the level of compensation (solar zenith, atmosphere, RSR). The RSR's effect can be as high as 20% in NED units, i.e, the estimated changes on the ground has an uncertainty of 20% due to the effect of RSR if the data from the two sensors are used directly in change detection applications. However, if the data are compensated using SBAF techniques, then the uncertainty can be reduced to about 1% in the change detection applications. In the case of uncompensated data, the STD varies between 1 - 2% which is mainly due to the visibility conditions.
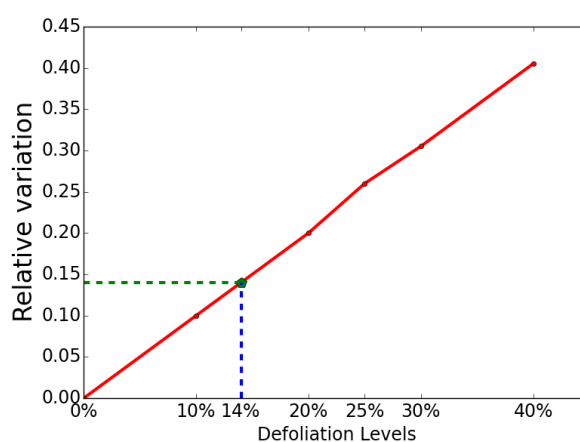
Figure 5.58: Illustration of effect estimation in NED units.

(a) SBAF using TOA REFL

(b) SBAF using BRF
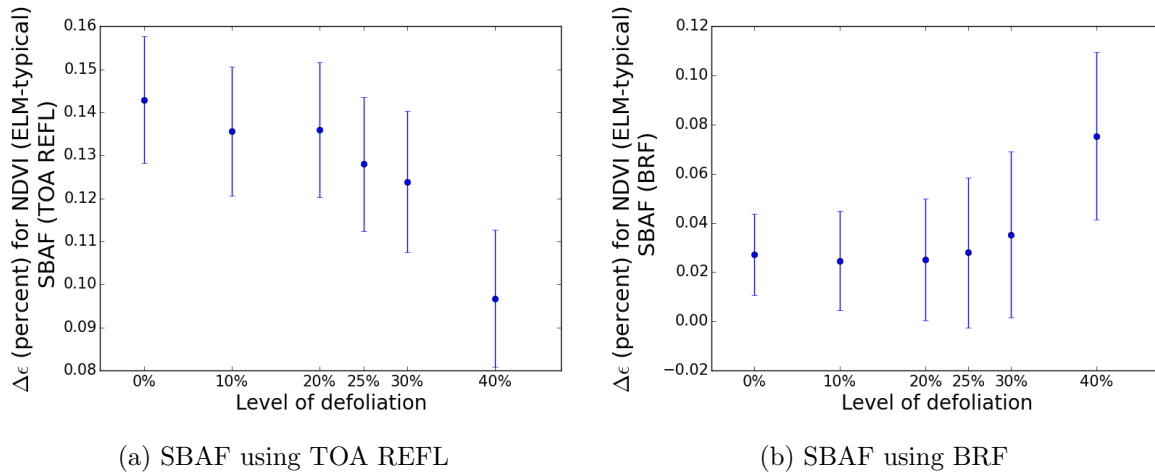
Figure 5.59: Effect of RSR after the SBAF compensation for ELM-typical NDVI product

(a) SBAF using TOA REFL

(b) SBAF using BRF
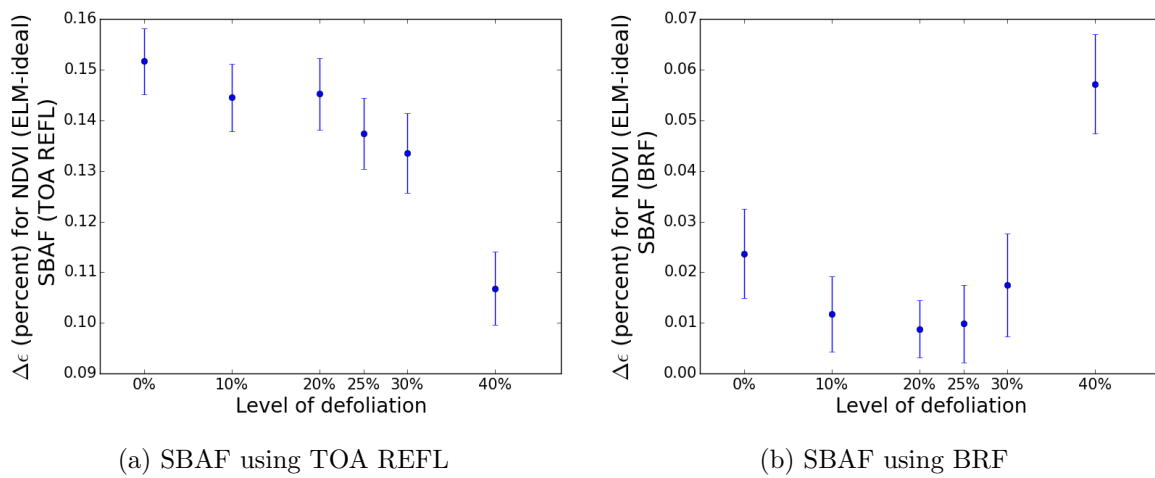
Figure 5.60: Effect of RSR after the SBAF compensation for ELM-ideal NDVI product

Table 5.14: Effect of RSR for NDVI products before and after SBAF compensation. The effects are shown in NED units, i.e, the percentage implies that the effect due to RSR is equivalent to an actual level of defoliation.

Data products	uncompensated Mean (STD)		SBAF (TOA REFL) Mean (STD)		SBAF (BRF) Mean (STD)	
NDVI TOA Rad	19%	(2.2%)	2%	(1.3%)	7%	(3.4%)
NDVI TOA Refl	19%	(1.9%)	2%	(1.3%)	7%	(3.2%)
NDVI ELM-typical Refl	15%	(1.5%)	3%	(0.6%)	1%	(0.9%)
NDVI ELM-ideal Refl	15%	(0.8%)	3%	(0.5%)	1%	(0.5%)

Thus, the effect analysis indicates that the differences due to RSR can introduce as much change as would be observed when there is a 20% defoliation in the deciduous forest. This is based on the assumption that the data are acquired in an ideal, coincident imaging conditions which are

highly unlikely. But, when compensated using SBAF techniques, the uncertainty due to RSR can be reduced significantly.

5.6.2   X-Track effects

The effects due to the difference in view angles are estimated as discussed in Section 4.6.2. The across-track effects are estimated for two cases. In the first case, the effect is estimated between the two sensors (different RSRs) that view the same target from two different across-track angles. This is a likely scenario, where, the two sensors image the same target on the same day at about the same time ($\approx$ 30 mins difference) and the atmospheric conditions may be assumed to not vary, but the solar angles could be different ($\approx 3°$ to $5°$). Nevertheless, this is a plausible scenario and analysis of this case is useful to understand the expected error in change detection. The relative variation in this case is computed between the sensors, and so the estimated effect is a combination of effects due to the RSR and the view angle.

In the second case, the effect is only due to view angles as it is estimated between two extreme view angles (±12) for the same sensor. This is a scenario that is unlikely between OLI like sensors as the view angle of OLI is less than $\pm7.5°$. It is a possible scenario with MSI sensors when both the Sentinel 2 satellites are operational, but assuming the sun angles, atmosphere and RSR to not change between the two acquisitions may not be a reasonable assumption. Yet, this case is useful to estimate the effects due to view angle differences for similar kind of sensors.

The effects due to the difference in across-track angle for three NDVI products are shown in Figures 5.61, 5.62 and 5.63. The results for the other two NDVI products are shown in Appendix G. The effects for the TOA NDVI products are consistent and therefore, only one of them is shown in Figure 5.61. Similarly, the effects for the ELM compensated NDVI products are also consistent.

It is observed from Figure 5.61 that the relative variations are higher when the two sensors view from opposite sides, i.e, one sensor views the target from the positive across-track angle and the other from the negative across-track angle. When the two sensors view from the same side, the relative variations are smaller as the effective difference in view angle (ΔXT) is reduced. There are some inconsistencies in the effects between the simulated products (TOA radiance, reflectance and ELM) and the canopy BRF products. For example, when the L8 view angle is

$-7.5°$ and the S2 view angle varies from $0°$ to $-12°$ (i.e, increase in ΔXT), both the simulated and BRF NDVI products show an increasing trend as expected. However, when the L8 view angle is $7.5°$, for increase in ΔXT, the simulated products show a decreasing trend. If the same analysis is performed on the SBAF compensated products, an increasing trend is observed (see Appendix G, Figure G.2). Therefore, the change in the trend is attributed to the inherent effect of the RSR while computing the view angle effects.

The mean effect for each NDVI data product is determined by averaging the relative variation across all the signal levels and the view angle combinations. The mean effect and the STD in NED units are shown in Table 5.15. The STD observed in this table is mainly due to the large variation in the effects for the different view angle combinations between the two sensors.
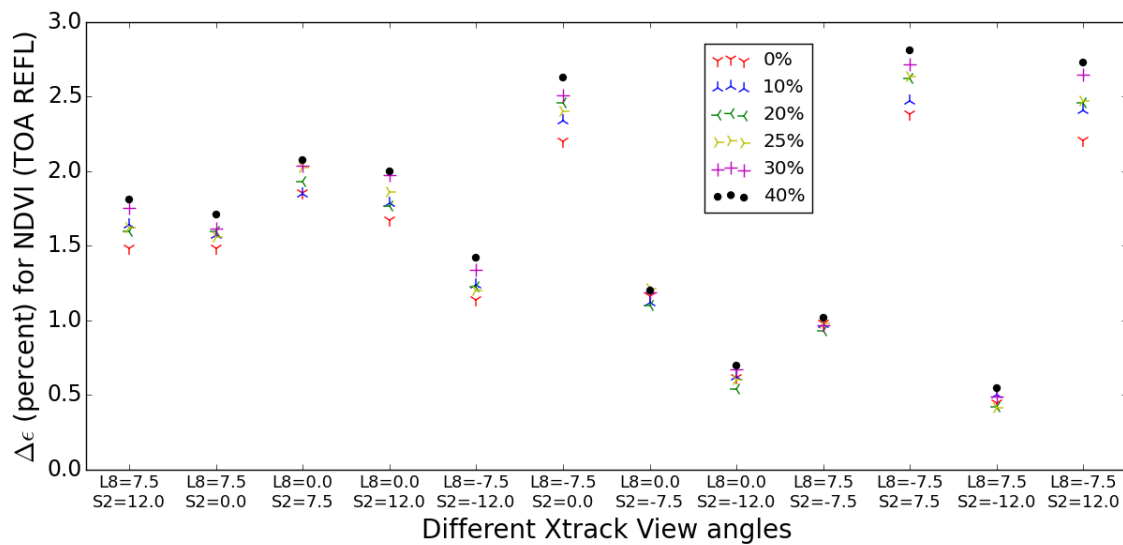
Figure 5.61: Effect of across-track for different view angle combinations for NDVI (TOA REFL) product.
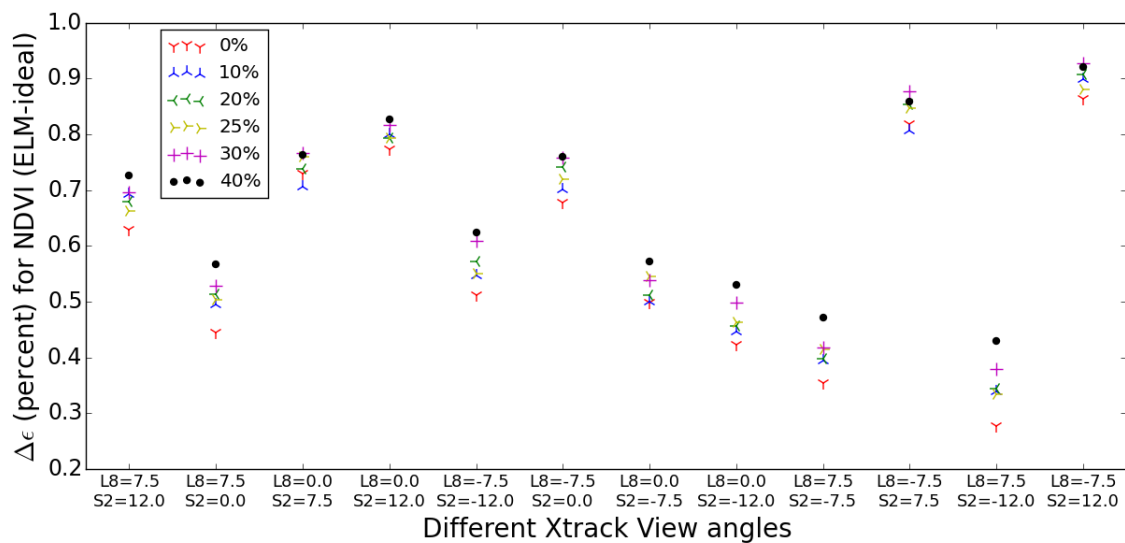
Figure 5.62: Effect of across-track for different view angle combinations for NDVI (ELM-ideal) product.
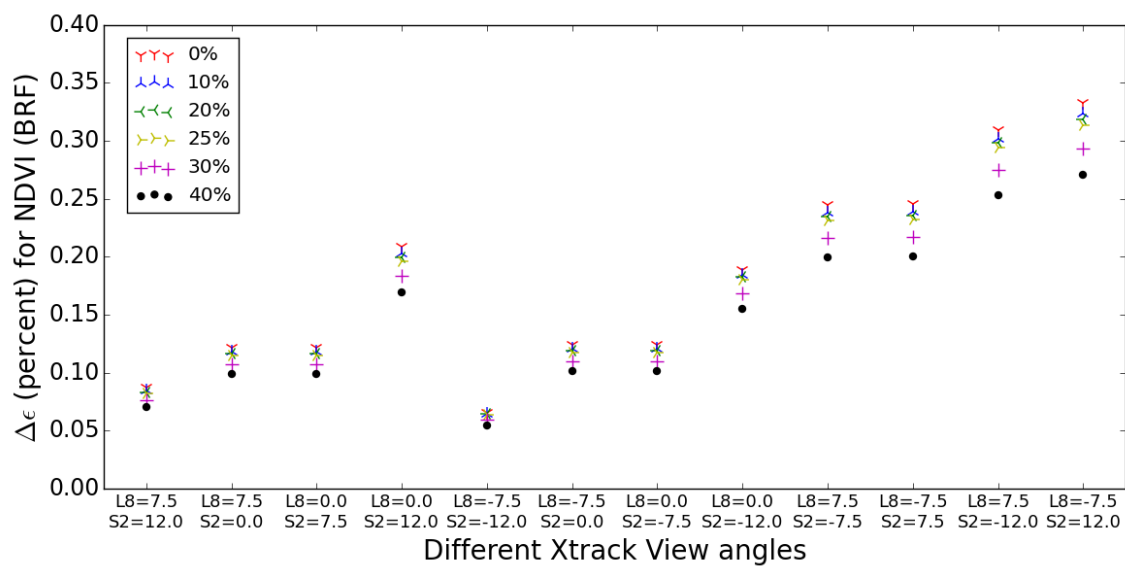
Figure 5.63: Effect of across-track for different view angle combinations for NDVI (BRF) product.

The effects are also estimated at two extreme angles ($\pm 12°$) by evaluating the mean relative response for both the sensors. The mean effects for different signal levels are shown in Figure 5.64. Since effects due to the RSR are excluded in this analysis, the large STD can be attributed to the BRDF of the canopy or the other factors. The high STD for the BRF product (see Figure 5.64d) indicates that the observed variation is mainly due to the BRDF of the canopy. The STD is observed to be independent of the signal level, and as in the other case, there are no apparent trends between the effect and the signal levels for any of the NDVI products. The

Table 5.15: Effect of across-track angle for NDVI products before and after SBAF compensation. The effects are shown in NED units.

Data products	uncompensated		SBAF (TOA REFL)		SBAF (BRF)	
	Mean	(STD)	Mean	(STD)	Mean	(STD)
NDVI TOA Rad	17%	(8.8%)	8%	(5.7%)	11%	(8.2%)
NDVI TOA Refl	17%	(8.6%)	8%	(5.7%)	10%	(8.1%)
NDVI ELM-typical Refl	28%	(21.5%)	28%	(21.3%)	28%	(21.9%)
NDVI ELM-ideal Refl	16%	(5.0%)	5%	(3.8%)	4%	(2.7%)
NDVI BRF Refl	5%	(2.8%)	5%	(2.8%)	5%	(2.8%)

mean effect for each data product is determined by averaging the effect across all the signal levels. The mean effect at the extreme angles for each data product in NED units is shown in Table 5.16

The ELM-typical products show larger effects than any other products (see Table 5.16) mainly due to the error in the atmospheric compensation technique. In the ELM-ideal method, the panels are used to correct for every view angles independently. Hence the estimate of red and NIR reflectance is very close to the canopy BRF. In the case of ELM-typical method, the reflectance panel observations in the nadir view angle are used to compensate for all the view angles. As a result, in the back-scatter direction ($-12°$), the red reflectance is over-estimated due to the difference in path-transmission and increase in the upwelled radiance. In the NIR band, the difference in transmission between nadir and non-nadir view is very small, and so the reflectance estimated in NIR is unaffected. The over-estimation of red reflectance results in the under-estimation of the NDVI score. In the forward scatter direction ($+12°$), the NIR reflectance is unaffected as before, but the red reflectance and the NDVI changes only marginally. Thus, a large difference in the NDVI between the two view angles result in a large effect for the ELM-typical products.

(a) TOA REFL

(b) ELM-typical

(c) ELM-ideal

(d) BRF
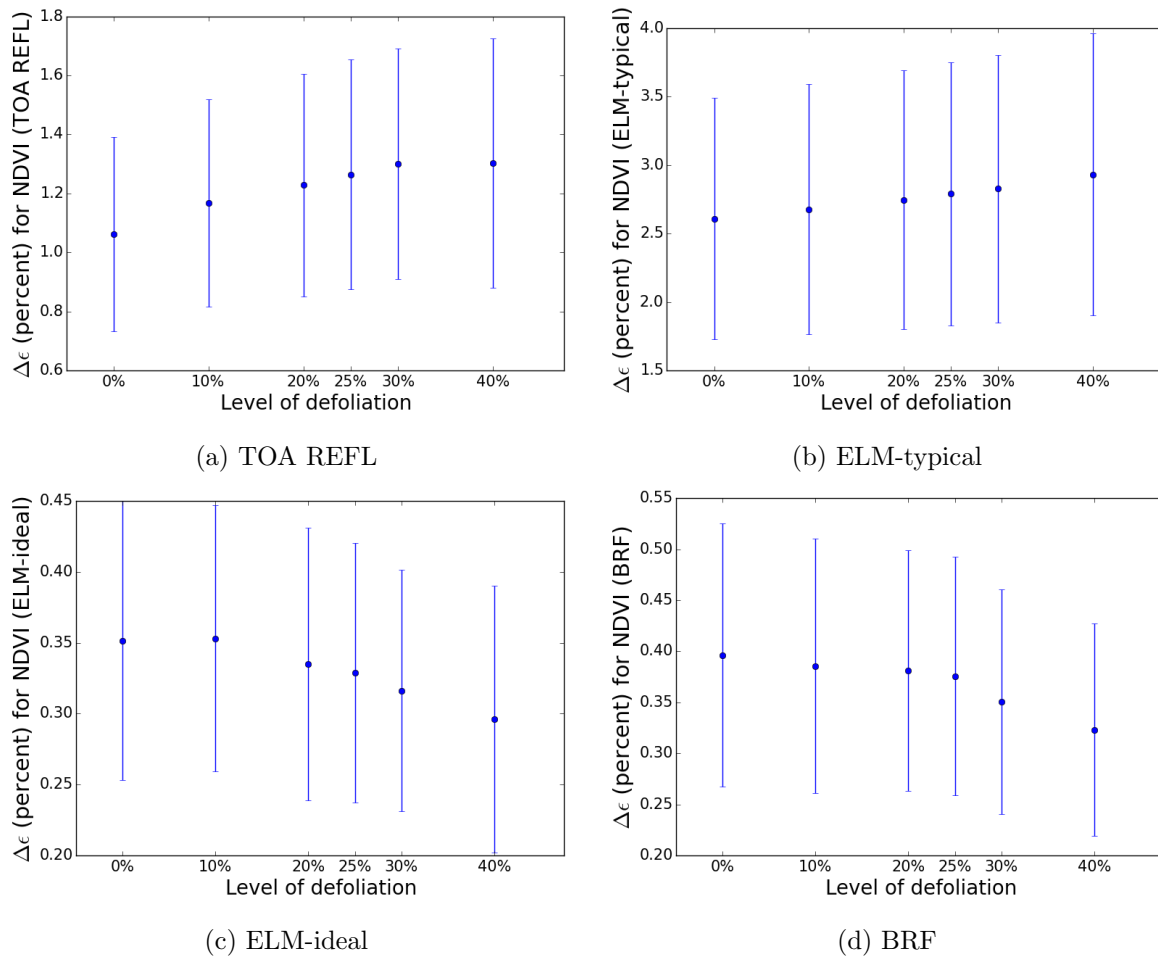
Figure 5.64: Effect of across-track angle for different NDVI products when the view angles are $(\pm 12°)$.

Table 5.16: Effect of two extreme view angles $(\pm 12°)$ for NDVI products before and after SBAF compensation. The effects are shown in NED units. Since the effect estimate calculation is independent of the sensor, the SBAF compensation does not show any improvements.

Data products	uncompensated Mean	(STD)	SBAF (TOA REFL) Mean	(STD)	SBAF (BRF) Mean	(STD)
NDVI TOA Rad	14%	(4.6%)	14%	(4.7%)	14%	(4.8%)
NDVI TOA Refl	14%	(4.5%)	14%	(4.7%)	14%	(4.7%)
NDVI ELM-typical Refl	40%	(21.7%)	40%	(22.1%)	40%	(22.2%)
NDVI ELM-ideal Refl	9%	(2.6%)	9%	(3.8%)	9%	(2.7%)
NDVI BRF Refl	10%	(3.2%)	10%	(3.2%)	10%	(3.2%)

5.6.3 Visibility effects

The effects due to the differences in atmospheric conditions during image acquisition between the two sensors can be estimated by simulating the MODTRAN atmosphere for different visibility

conditions. The simulations and the methods used to estimate the effects were discussed in Section 4.6.3. The simulation in this case uses two sensors (OLI, MSI) to view the target at the same view and illumination angles, but under different atmospheric conditions. More often than not, the sensors due to their orbital parameters, views the same target on two different days. However, the sun and the view angles may not be the same between the two acquisitions. The sun angle is dependent on the time-lag between the two sensors' acquisitions. For example, if the two images are acquired anywhere between 3 to 5 days, then the sun angle between the two acquisitions can differ by about 5° in SZN angle. The constancy in view angle is dependent on the geographic location of the target. Nevertheless, the conditions in this case are more likely than any other scenarios discussed earlier. The relative variation in this case is computed between the sensors for different visibility conditions, and so the estimated effect is a combination of the effect due to the RSR and visibility differences.

The analysis is performed for both the radiance products (red and NIR bands) and the NDVI products. The radiance products are included as they are directly affected by the changes in the visibility conditions. The NDVI product generated using BRF of the canopy is not analyzed since the BRF is independent of the atmospheric conditions. The relative variations for all the products are shown in Figures 5.65 to 5.70. The relative variation is higher for the TOA radiance product in the red band (see Figure 5.65) than the NIR band (Figure 5.66). This is due to the higher atmospheric scattering in the red band, which leads to a large difference in the transmission between the visibility conditions. The effect due to visibility is observed to be signal dependent in the NIR band but not in the red band.

The effect increases with increase in the visibility differences (ΔVis). For example, when the visibility of L8 is at 20 km and the visibility for S2 varies from 20 km to 10 km, their relative variation increases for increase in ΔVis (from 0 km to 10 km). This is expected as the visibility is directly proportional to transmission and so a large ΔVis will cause a large difference in the transmissions. Similarly, for the same ΔVis, the relative variation is high when the visibility is low. This is because, the optical depth is exponentially related to visibility (not linear).
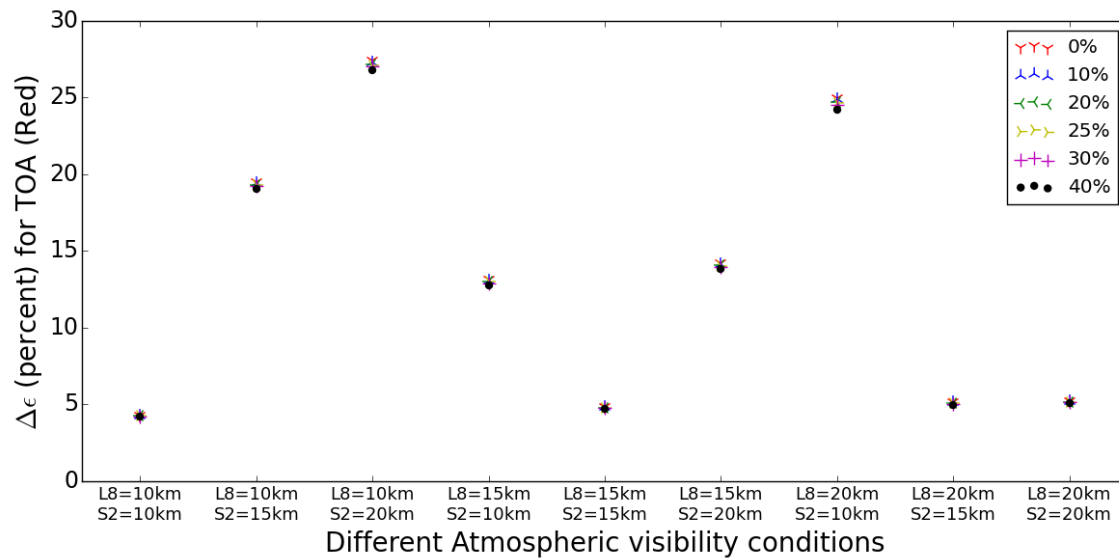
Figure 5.65: Effect of atmospheric differences for different visibility combinations for TOA radiance product (RED).
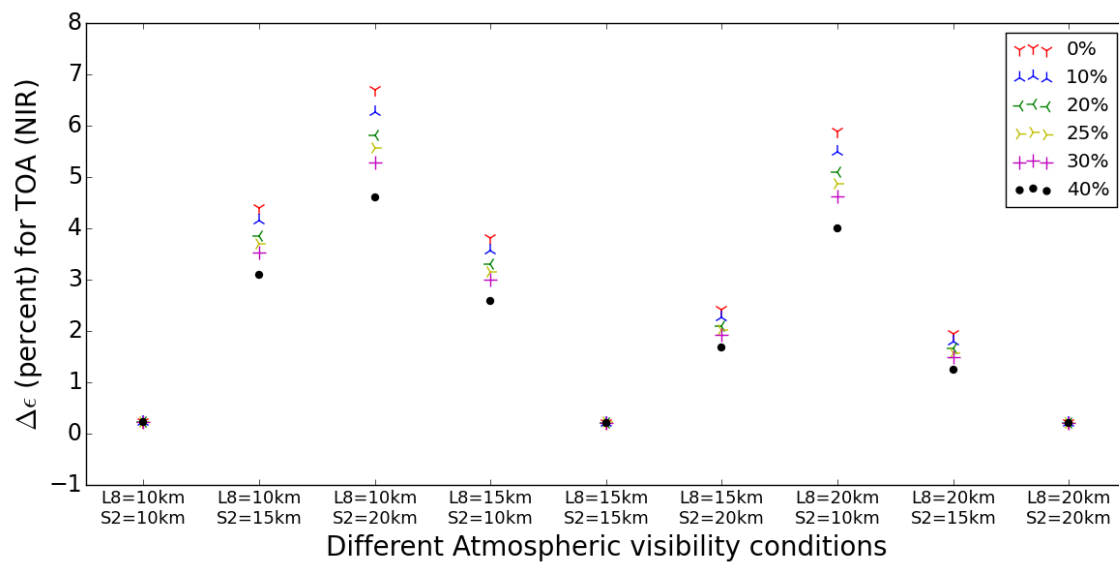
Figure 5.66: Effect of atmospheric differences for different visibility combinations for TOA radiance product (NIR).
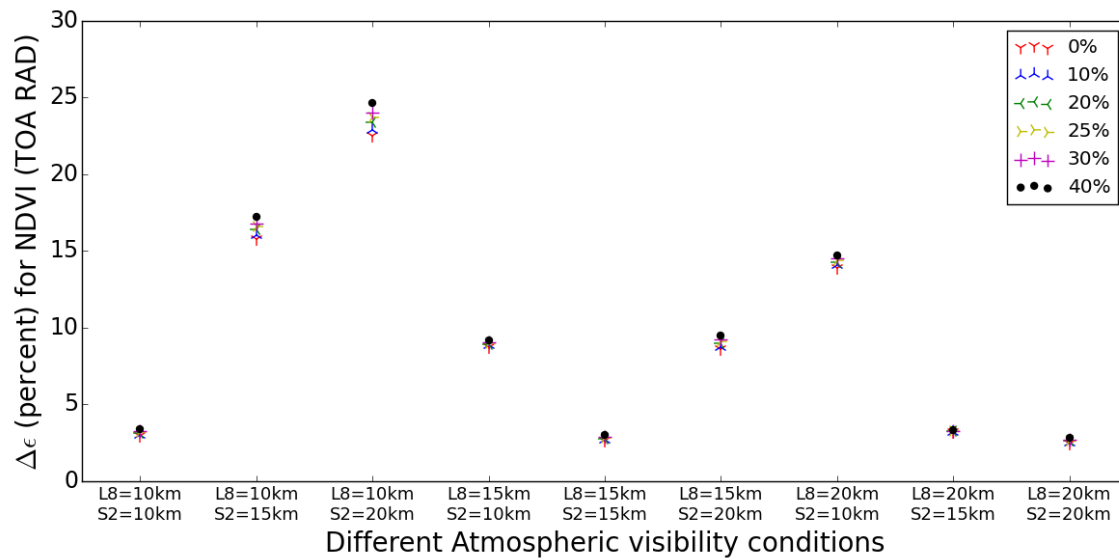
Figure 5.67: Effect of atmospheric differences for different visibility combinations for NDVI product (TOA RAD).
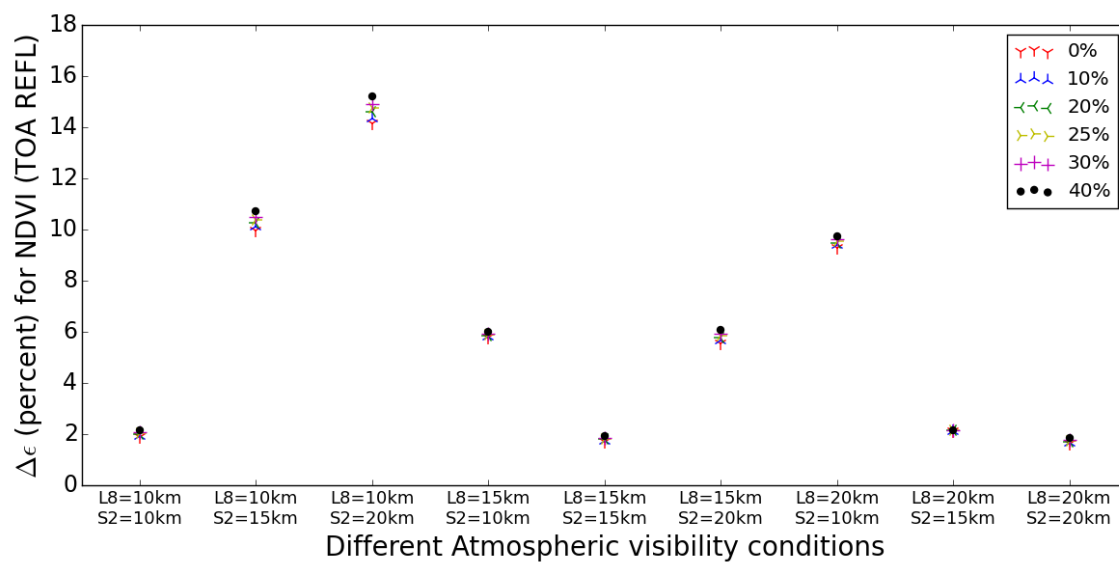
Figure 5.68: Effect of atmospheric differences for different visibility combinations for NDVI product (TOA REFL).
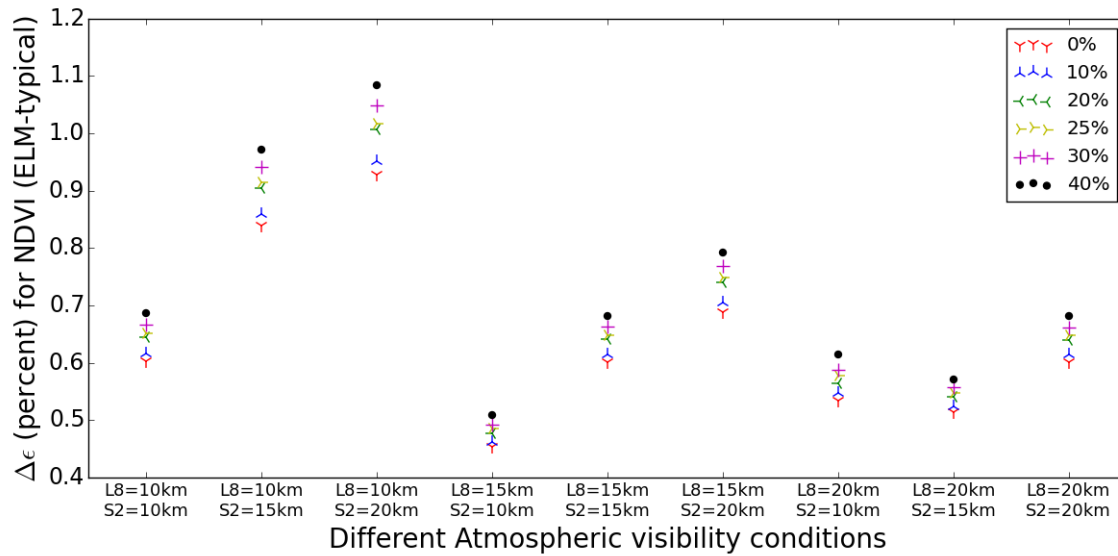
Figure 5.69: Effect of atmospheric differences for different visibility combinations for NDVI product (ELM-typical).
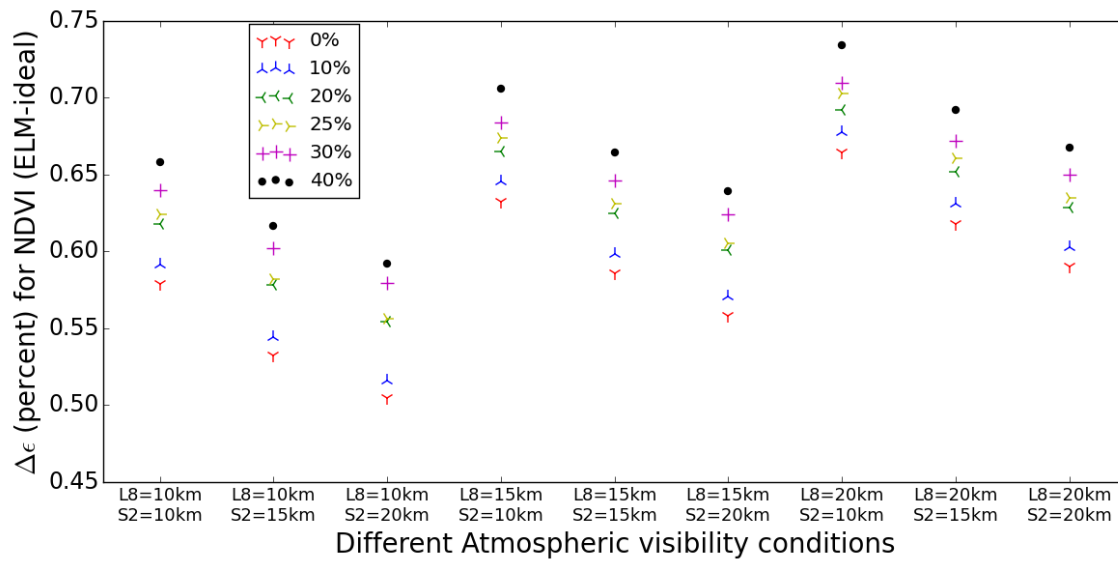
Figure 5.70: Effect of atmospheric differences for different visibility combinations for NDVI product (ELM-ideal).

Among the NDVI products, TOA radiance and reflectance products show relative variation as high as 20% for certain cases. Although this includes the effect due to RSR, its contribution is less than 3%. In the TOA products, the variation due to ΔVis is significantly higher than the variation due to the signal levels. For the ELM compensated products, the variation due to ΔVis is relatively low. The residual effect in the ELM compensated products is mainly due to the effect of the RSR.

The effect due to the RSR can be reduced by SBAF compensation techniques. The relative variation for the SBAF compensated products are shown in Figures 5.71 to 5.74. The results for both the SBAF techniques (BRF and TOA REFL) are consistent, and any improvements of the TOA REFL method over BRF method for the TOA products are only marginal. Hence, only one of the SBAF method is discussed in this section. The result for SBAF compensated TOA NIR product is not provided as it did not differ much from the TOA NIR product shown in Figure 5.66. This is because, the RSRs for the two sensors are similar in the NIR band, and its SBAF value is close to unity. As expected, the relative variations for the TOA products are high for SBAF compensated products (see Figures 5.71 and 5.72). In the case of SBAF adjusted and ELM compensated products, the relative variations are very small and are consistent with the results observed in Section 5.6.1 for the RSR effects (see Table 5.14).
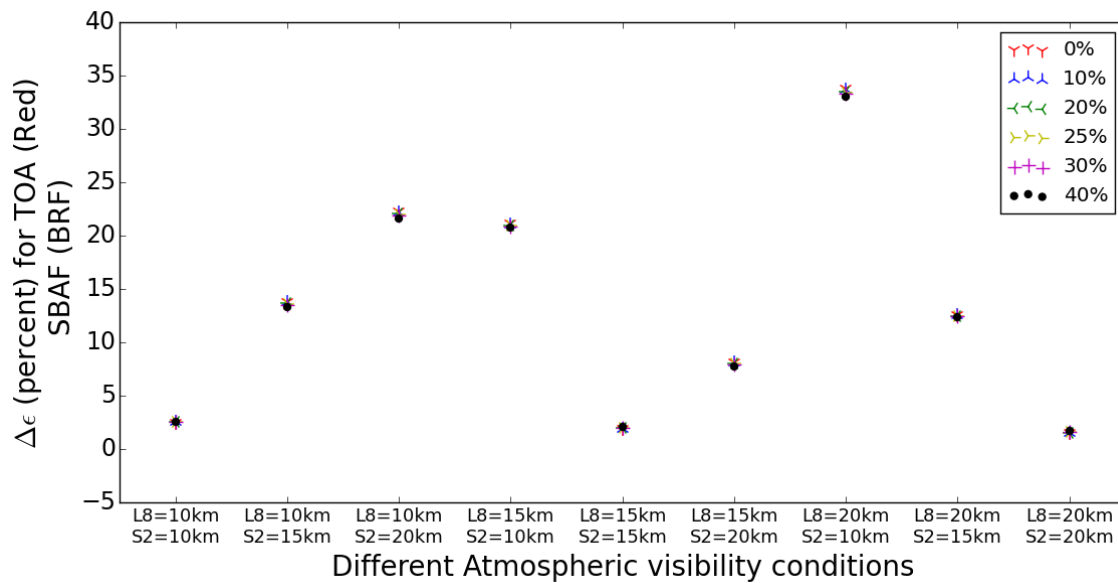
Figure 5.71: Effect of atmospheric differences for different visibility combinations for SBAF (BRF) compensated TOA radiance product (RED).
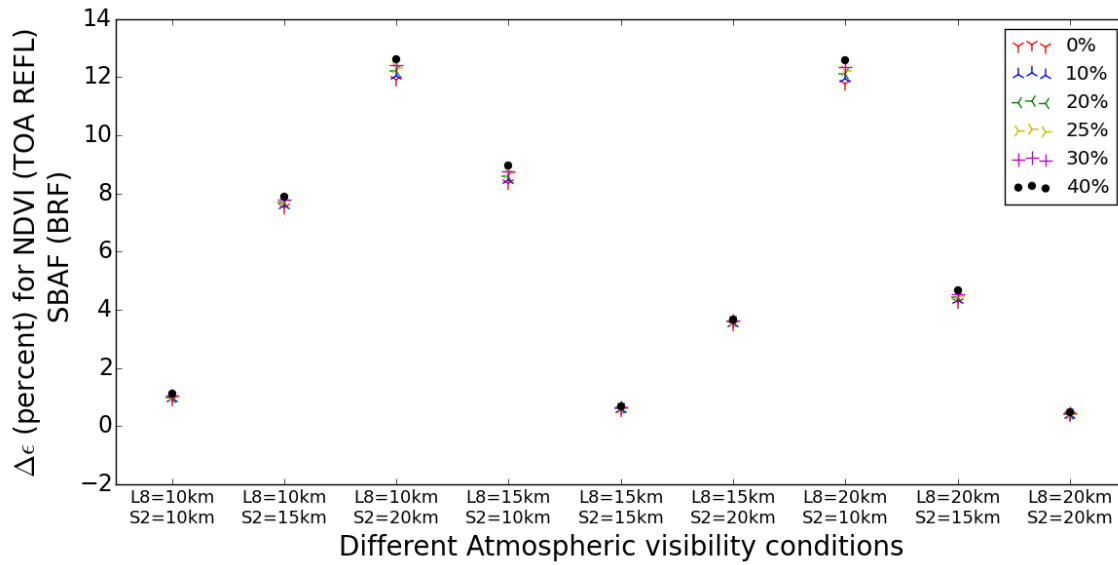
Figure 5.72: Effect of atmospheric differences for different visibility combinations for SBAF (BRF) compensated NDVI product (TOA REFL).

The mean effect for each NDVI product is determined by averaging the relative variation across all the signal levels and the visibility combinations. The mean effect and the STD in NED units are shown in Table 5.17. The STD observed in this table is mainly due to the variation in the visibility combinations. As discussed earlier, the ELM-typical method does not account for off-nadir path length differences, which leads to high NED compared to the ELM-ideal compensated products.
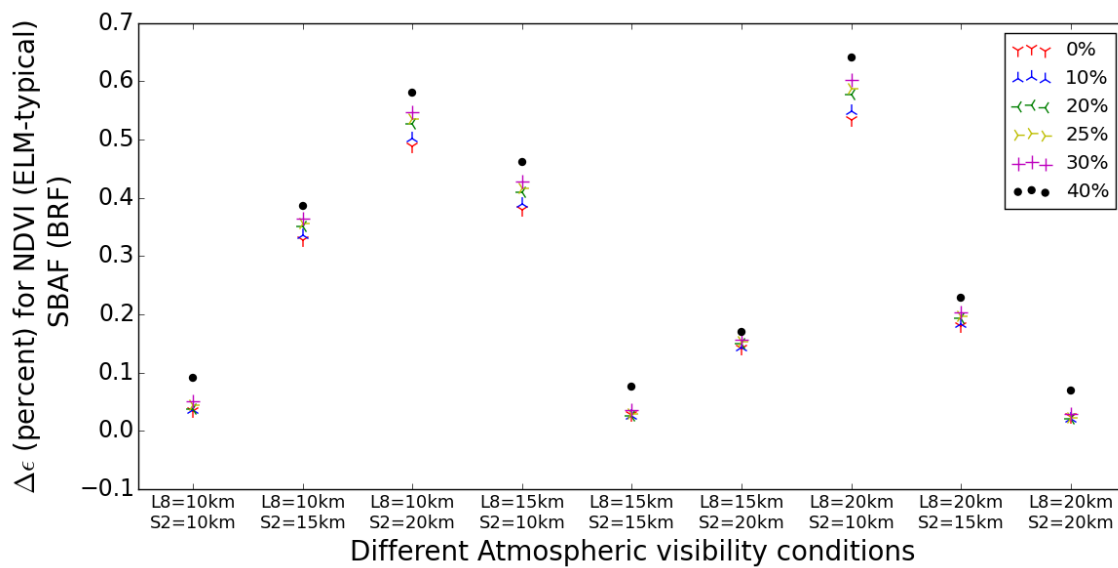
Figure 5.73: Effect of atmospheric differences for different visibility combinations for SBAF (BRF) compensated NDVI product (ELM-ideal).
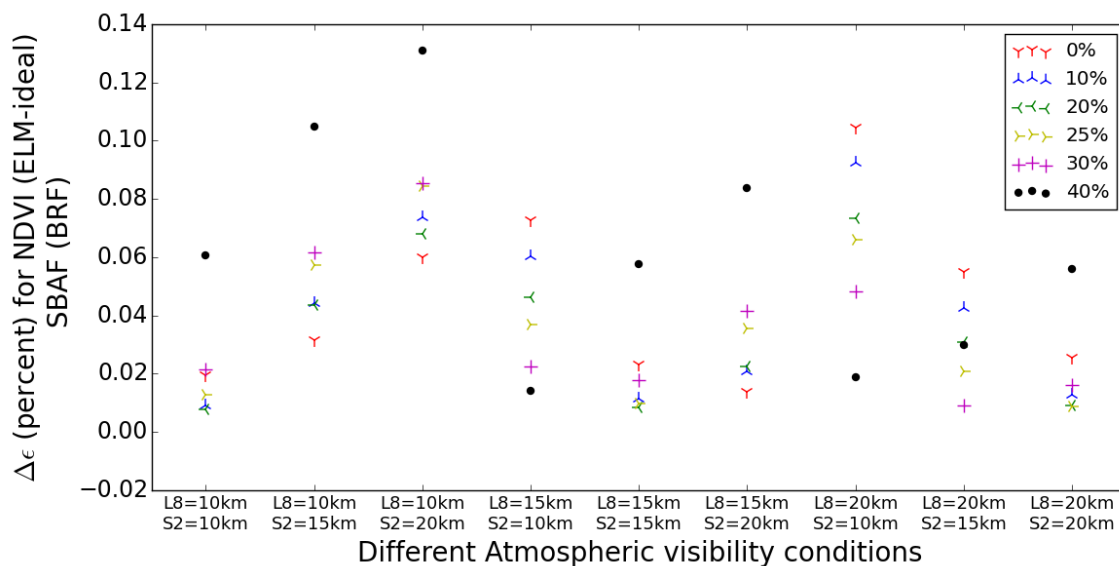
Figure 5.74: Effect of atmospheric differences for different visibility combinations for SBAF (BRF) compensated NDVI product (ELM-typical).

Table 5.17: Effect of visibility for NDVI products before and after SBAF compensation. The effects are shown in NED units.

Data products	uncompensated Mean (STD)		SBAF (TOA REFL) Mean (STD)		SBAF (BRF) Mean (STD)	
NDVI TOA Rad	40%	(39%)	40%	(40%)	40%	(39%)
NDVI TOA Refl	40%	(38%)	40%	(39%)	40%	(38%)
NDVI ELM-typical Refl	16%	(10%)	8%	(9%)	7%	(9.7%)
NDVI ELM-ideal Refl	15%	(1.4%)	4%	(1.3%)	1%	(0.9%)

Thus, in a situation where both the sun and the view geometries are the same, and the RSR and atmospheric effects compensated, then the residual effect is equivalent to 1% in NED units. This indicates that the changes on the ground can be ascertained to within 1% uncertainty. In practice, it is difficult to achieve an ideal compensation for the atmosphere, and the sun and view geometries are less likely to be similar between the two acquisitions. However, if the two effects are compensated (RSR and visibility), the remaining effects (SZN and across-track) are dependent on the BRDF of the target. In such a case, the OLI and the MSI observations can be used to model the BRDF of a target at a high spatial resolution.

5.6.4 SZN effects

The methods used to analyze the effect of SZN was discussed in Section 4.6.4. The analysis is performed for two cases; the first case is based on the expected difference in the SZN angles between the two sensors, and in the other case, the simulated sensor response for a fixed difference in the SZN angle is used to analyze its effect.

The OLI sensor has a revisit period of 16 days and the MSI sensor revisits every 10 days. If both the Sentinel-2 sensors are operational, the revisit period from any of the MSI sensor drops to 5 days. The OLI and the MSI sensors may image the target on the same day or could be off by several days (upto 10 days). The geographic location of the target and the orbital parameters of the two satellites can be useful to predict their cross-over dates. But, operational issues in acquiring the target or the presence of clouds over the target may nullify the possibility of using the data for change detection studies. Hence, it is important to consider a wide range of days for the two sensors to cross-over the same target.

The two sensors imaging the same target on different days can introduce changes in the illumination and view angles, and atmospheric conditions apart from sensor differences (RSR). The RSR effects can be compensated to first order if SBAF corrections are applied, or if the two sensors' RSR are similar (two MSI sensors). Therefore, we need to assume the unlikely scenario where the view angles and the atmospheric conditions are the same for two different days.

Figure 5.75 shows the difference in the SZN angle (ΔSZN) between the OLI and the MSI sensors for an entire year, over the Harvard forest (Lat: $42°$, Lon : $-72°$). The absolute difference in the SZN angle is shown for different cross-over days (from 0 to 20 days). For example, when the two sensors image the same target 5 days apart, the expected ΔSZN for Harvard forest is anywhere from $2°$ to $4.5°$ but, if they image on the same day, then it is about $3°$. The ΔSZN ranges from $0°$ to about $8°$ when the two sensors image 20 days apart. Since the ΔSZN is dependent on the geographic position, a similar analysis is performed for a location at the equator. Figure 5.76 shows the result for the Amazon forest (Lat: $0°$, Lon : $-72°$). The ΔSZN ranges from $5°$ to $7°$ and is about the same, irrespective of whether the two sensors image on the same day or 5 days apart. The ΔSZN is higher for the Amazon forest than the Harvard forest, which indicate that the ΔSZN reduces as we move away from the equator.

(a) 0 to 5 days

(b) 6 to 20 days
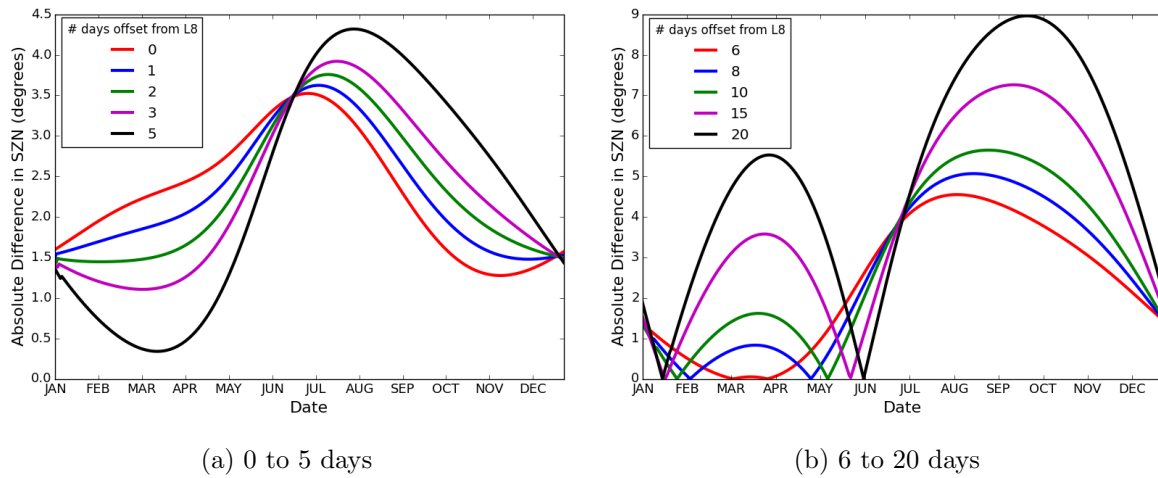
Figure 5.75: Difference in SZN angles between L8 and S2 over Harvard forest

(a) 0 to 5 days

(b) 6 to 20 days
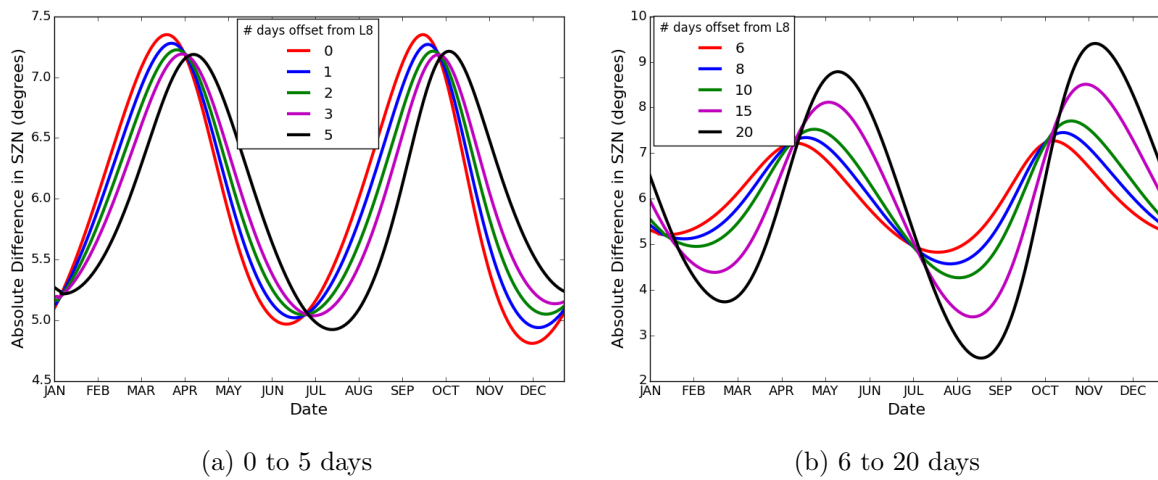
Figure 5.76: Difference in SZN angles between L8 and S2 over Amazon forest

A plot of ΔSZN between the two sensors and its corresponding relative variation at different SZN angles, estimated using cosine effect of SZN angles (see Section 4.6.4), is shown in Figure 5.77. In the case of a deciduous forest canopy, $3°$ to $7°$ in ΔSZN induces about 3-10% in relative variation. The relative variation estimated in this case is only due to the difference in the SZN angles and does not include any other factor's effect or its interaction.
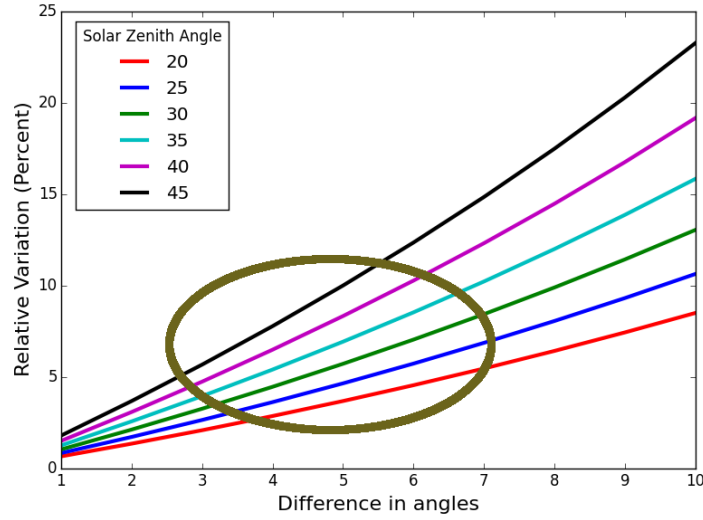
Figure 5.77: Expected variation as a function of difference in SZN angles. In the case of deciduous forest canopy, for an expected difference of 3° to 7° in the SZN, about 3% to 10% in relative variation is observed (ellipse).

The effect due to ΔSZN is analyzed using the simulated data, similar to the analysis performed for other factors. The effects are estimated for a 5° ΔSZN as it varies between 2° to 7° for the Harvard forest. The relative variation as shown in Figures 5.78 and 5.79, is estimated for each sensor independently (no RSR effect), and assumes that the atmospheric conditions and the view angles are the same.

Comparing Figures 5.78a and 5.78b, it is observed that the product compensated for SZN (TOA REFL) shows only marginal improvement compared to the uncompensated product (TOA RAD). This is because, the cosine correction $cos(\theta_s)$ is negated by taking the ratio of the spectral bands in the NDVI calculations. The large STD observed in the TOA products are due to the variations in the visibility conditions. The relative variation for all the NDVI products are consistent and so, the ELM-ideal case is shown in Figure 5.79. The high relative variation when ΔSZN is 6° (SZN1 is 27° and SZN2 is 33°) indicate that the relative variation increases with increase in ΔSZN.

The mean effect for each product is determined by averaging the relative variation across all the signal levels. The mean effect and the STD in NED units are shown in Table 5.18. Thus, in a situation when two sensors with similar RSRs image the target at the same view angle and atmospheric conditions, then the effect due to SZN is about 10% for a ΔSZN of 5°. If the data are also compensated for the atmospheric attenuations, then the effect varies between 4-8% depending on the accuracy of the compensation technique. Similar analysis is performed

for ΔSZN=10° and its results are summarized in Table 5.19. It is seen that the effect almost doubles for twice the difference in SZN angles.

(a) NDVI (TOA RAD)

(b) NDVI (TOA REFL)

(c) NDVI (ELM-ideal)

(d) NDVI (BRF)
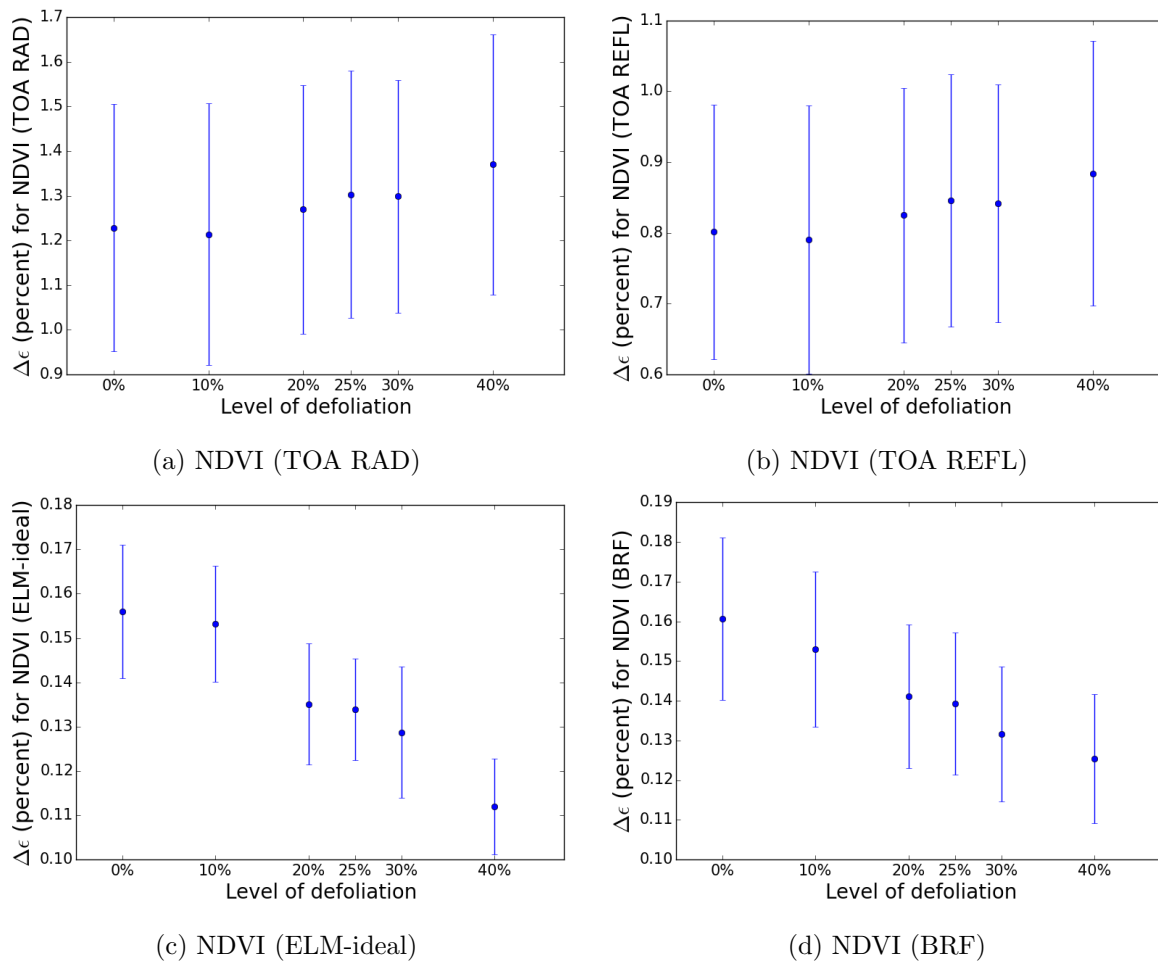
Figure 5.78: Effect of SZN for different NDVI products for 5° difference in SZN angles.
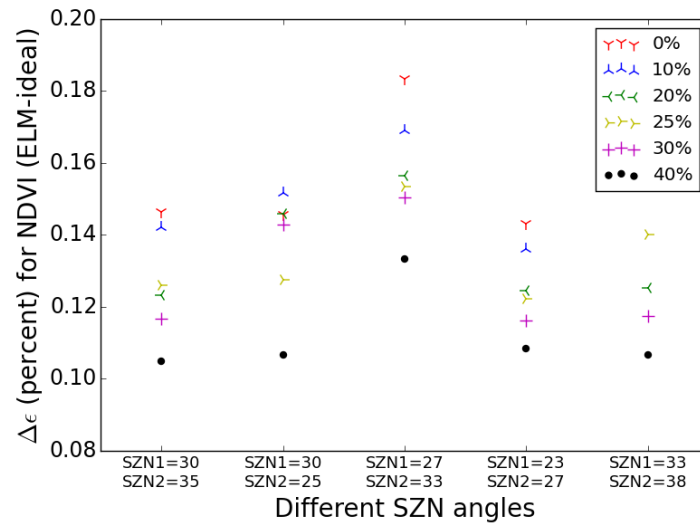
Figure 5.79: Relative variation for $\approx 5°$ difference in SZN angles

Table 5.18: Effect for 5° difference in SZN angles for NDVI products before and after SBAF compensation. The effects are shown in NED units.

Data products	uncompensated Mean (STD)		SBAF (TOA REFL) Mean (STD)		SBAF (BRF) Mean (STD)	
NDVI TOA Rad	10%	(5%)	10%	(5%)	10%	(5%)
NDVI TOA Refl	10%	(5%)	10%	(5%)	10%	(5%)
NDVI ELM-typical Refl	8%	(7%)	8%	(7%)	8%	(7%)
NDVI ELM-ideal Refl	4%	(1.4%)	4%	(1.4%)	4%	(1.4%)
NDVI BRF	4%	(1.6%)	4%	(1.6%)	4%	(1.6%)

Table 5.19: Effect for 10° difference in SZN angles for NDVI products before and after SBAF compensation. The effects are shown in NED units.

Data products	uncompensated Mean (STD)		SBAF (TOA REFL) Mean (STD)		SBAF (BRF) Mean (STD)	
NDVI TOA Rad	19%	(8%)	19%	(8%)	19%	(8%)
NDVI TOA Refl	19%	(8%)	19%	(8%)	19%	(8%)
NDVI ELM-typical Refl	11%	(7%)	12%	(7.5%)	12%	(7.5%)
NDVI ELM-ideal Refl	8%	(1.9%)	8%	(2%)	8%	(2%)
NDVI BRF	8%	(2.2%)	8%	(2.2%)	8%	(2.2%)

In an ideal condition with no other effects (NDVI BRF), the effect in SZN is due to the BRDF of the canopy and is about 4% when ΔSZN=5°. This indicates that the changes on the ground can be ascertained to within 4% uncertainty for an ideal compensated products. In practice, it is difficult to achieve an ideal acquisition and compensation, and so the expected uncertainty in change detection studies due to SZN will be higher than 4%.

5.7 Simulated vs Real data

5.7.1 Reflectance comparison

The real and the simulated data over the Harvard forest are compared using the Landsat 8 scenes acquired on July 11th and August 3rd of 2015. The sun angles and the view angles were different between these two acquisitions and are shown in Table 5.20. A graphical representation of the view angles of the OLI sensor on these two dates, while imaging the Harvard forest, is shown in Figure 5.80. The STK tool is used to estimate the view azimuth angles by simulating the Landsat 8 orbit using the TLEs for the two scenes. It can be inferred from the figure that the scene acquired on July 11th views the forest in the back-scatter direction, while the other scene views from the forward-scatter (or near nadir) direction. The azimuth angles are almost equal except for a rotation of 180° between them. The SZN angles differ by about 5° between the two acquisitions. Note that the two scenes are acquired by the same sensor, i.e., no difference in the equatorial crossing time. Therefore, the difference in SZN angles for scenes that are 23 days apart is smaller than the difference observed in Figure 5.75 for two different sensors.

(a) July 11, 2015   (b) Aug 3, 2015
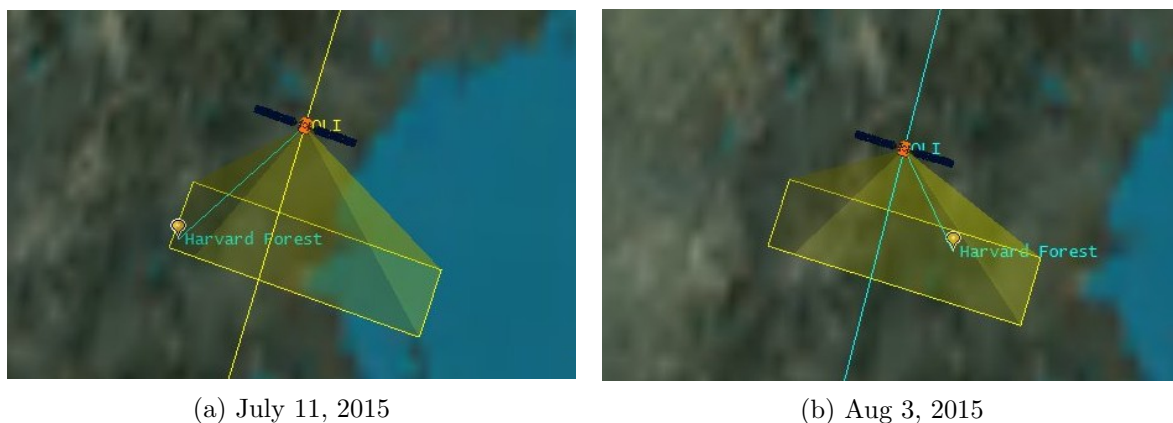
Figure 5.80: Graphical representation of Landsat 8 view angles over Harvard forest for two different dates. The OLI's view azimuth angles are estimated for the two acquisition dates by simulating the actual position and view angle of the Landsat 8 orbit using the STK tool (STK, 2016).

The view azimuth and the illumination angles (SAZ,SZN) of the Haravard forest, estimated from the OLI surface reflectance products, are used as the input parameters for the RossLi BRDF model. The MODIS BRDF product uses Ross (Thick) and Li (Sparse) model with geometric parameters HB and BR set to 2 and 1 respectively. The model coefficients from the BRDF product and the geometric parameters are used to estimate the reflectance for the VZN angle,

Table 5.20: Illumination and view geometry of the Landsat 8 scenes acquired over the Harvard forest.

Date	WRS Path / Row	SZN	SAZ	VZN	VAZ
			angles in degrees		
July 11, 2015	12 / 31	26.42	131.63	-7.18	284
August 3, 2015	13 / 30	31.16	139.15	2.38	104

ranging from $-45°$ to $+45°$. Similarly, the RossLi models generated for the simulated Harvard forest, as discussed in Section 5.4 (HB = 2.5 , BR = 1.9), are used to estimate the corresponding reflectance. The plot of the surface reflectance and the two BRDFs are shown in Figures 5.81 and 5.82 for July and August in the red and NIR wavelengthns. The measured and the derived reflectance for both the dates are shown in Table 5.21. The relative errors (see Equation 4.17) for the MODIS and the simulated BRFs are estimated using the OLI reflectance measurement as reference. The relative error estimated for the MODIS product in the NIR and the red wavelength is about 6% and 30% respectively. Although the error reduces in the NIR to about 2% for the other scene, the error in the red reflectance increases to 56%. Large errors observed in the MODIS product, especially in the red band, can be attributed to the following reasons. The MODIS BRDF product is generated by fitting atmospherically compensated observations from different sun and view geometries over a 16-day period to the RossLi BRDF model. Any residual errors in the atmospheric compensation method can affect the coefficients, especially in the red band. The 16-day observations assume that the ground does not change during this time, which may not be a valid assumption for the deciduous canopy. Further, the differences in the RSRs of the two sensors and the large GSD (500m) of the MODIS pixels can also contribute to errors.

The geometry of the virtual Harvard forest is different than the real Harvard forest, but the reflectance measured by the OLI sensor on July 11th matches the simulated BRF to within 2% in both the NIR and the red spectral bands. This suggest that a real forests' spectral characteristics can be approximated using a simulated forest canopy. The difference in reflectance is comparatively high for the reflectance observed on Aug 3rd for both the red and NIR spectral bands. This is expected, as the simulated BRDF is static and does not change with the changes on the ground. It is important to remember that an exact match of the Harvard forest is never the goal, but a plausible model that can represent a deciduous forest canopy. The consistency of the simulated BRDF in comparison to the shape of the MODIS BRDF and its

spectral similarity to the real data (OLI) clearly demonstrates that a plausible deciduous forest canopy can be modeled using the DIRSIG model.

(a) RED

(b) NIR
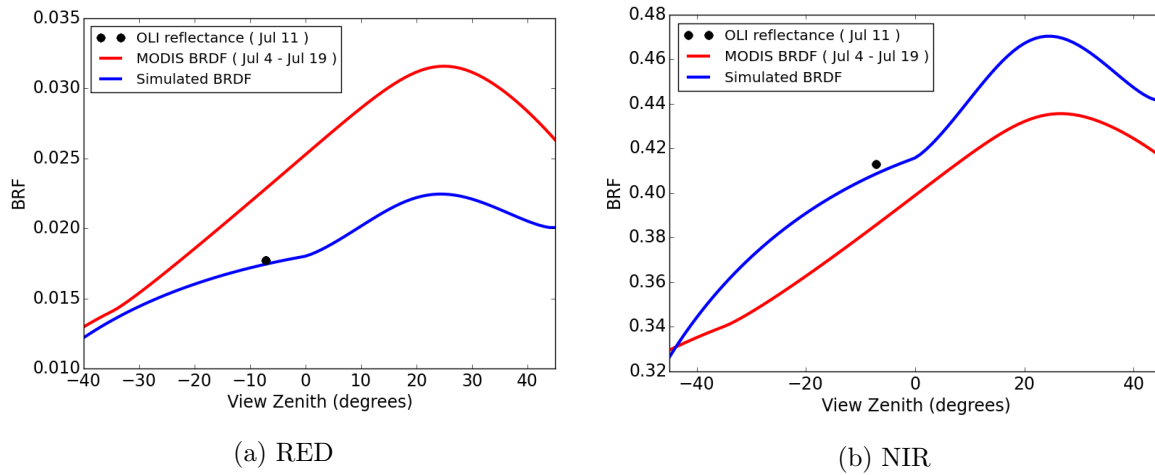
Figure 5.81: Comparison of real and simulated canopy reflectance of the Harvard forest in the red and the NIR spectral bands. The OLI scene was acquired on July 11, 2015 while the MODIS BRDF product was generated based on the 16-day acquisition from July 4 through July 19, 2015.

(a) RED

(b) NIR
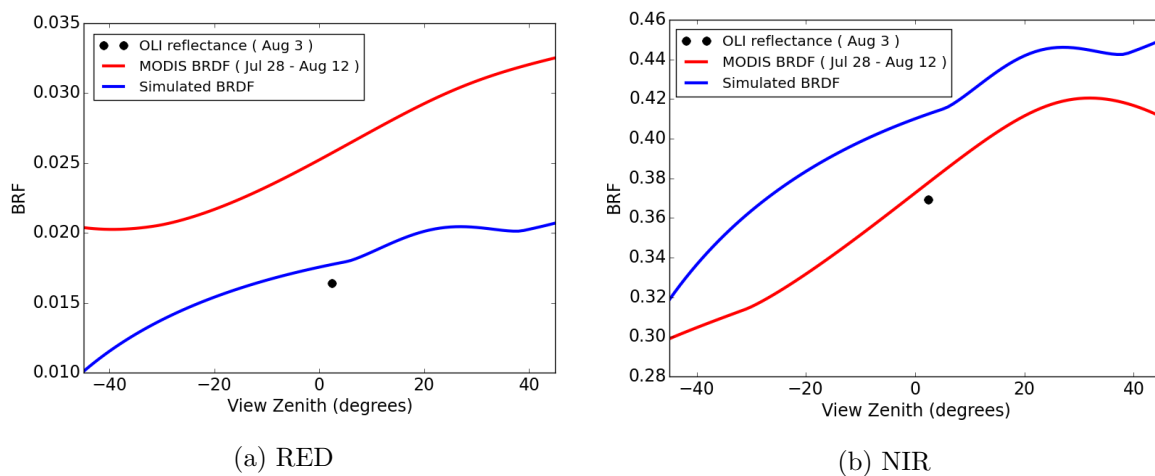
Figure 5.82: Comparison of real and simulated canopy reflectance of the Harvard forest in the red and the NIR spectral bands. The OLI scene was acquired on August 3, 2015 while the MODIS BRDF product was generated based on the 16-day acquisition from July 28 through August 12, 2015.

5.7.2 Effect analysis

The comparison of the real and the simulated data in the previous section concluded that the simulated data can be useful in representing the real-world conditions accurately. This

Table 5.21: Comparison of real and simulated BRF reflectance

Date	Product type	reflectance factor OLI	MODIS	Simulated	error relative to OLI (%) MODIS-OLI	Simulated-OLI
July 11, 2015	RED	0.018	0.023	0.017	29.3	1.4
	NIR	0.413	0.386	0.409	6.6	1
	NDVI	0.9178	0.8880	0.9181	3.2	0.03
Aug 3, 2015	RED	0.016	0.026	0.018	56.2	8
	NIR	0.369	0.377	0.412	2.0	11.6
	NDVI	0.9150	0.8726	0.9176	4.6	0.29

conclusion supports the analysis and provides more confidence in the results derived from the factors' sensitivity studies. In this section, the contribution of the factors are analyzed using the real and the simulated data. The data from the OLI sensor are used to estimate the actual changes, while the simulated data are used to estimate the effects due to the sensor and the environmental factors.

In the real data, the estimated change is due to the actual changes on the ground and the effects of different factors. Because the same sensor acquired both the scenes, the effect due to RSR differences are eliminated, but the effects due to visibility, sun and view angles are still present in the real data. It is not possible to separate the real changes in the forest from the combined effect (real and factor effects), but it can be inferred using the simulated data. In this study, we are interested in the contribution of the factors in comparison to the observed change in the real data. This is determined by evaluating the effect of the factors, as shown in the previous section, using the data simulated for the same view and illumination conditions as that of the real data.

The ROIs chosen for the Loyalsock state forest and the Harvard forest are shown in Figure 5.83. The ROIs of the Loyalsock forest are cloud-free, but clouds in the vicinity may introduce errors in the atmospheric compensated products. The view angles, sun angles and the acquisition dates for the Harvard and the Loyalsock forests are shown in Table 5.22. The OLI surface reflectance products are converted to NDVI products and the relative changes (NDVI) estimated for the Loyalsock and the Harvard forests are shown in Figure 5.84. The plot shows the mean relative variation and the relative variation of each ROI (red dots) for the two forests. The STD for the forests indicate that the observed changes varies depending on the location of the forest. The large change, observed in the Loyalsock forest within a short duration (ΔT=7 days), indicates

that the actual changes may be over-estimated due to the effect of other factors. Although an accurate estimate of the factors' effects is not possible without the actual BRDF model of the Loyalsock forest, the effects can be inferred from the analysis of the Harvard forest.

(a) Harvard Forest
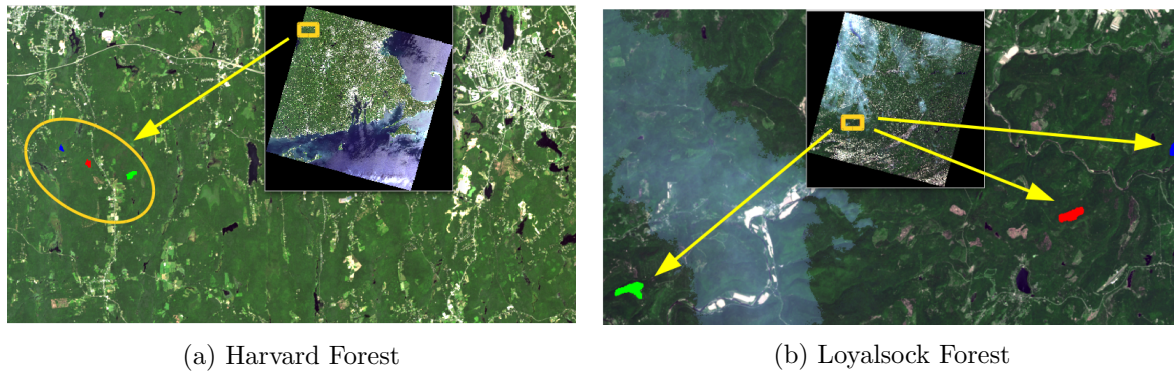
(b) Loyalsock Forest

Figure 5.83: ROIs from the Harvard and the Loyalsock forests. The ROIs are cloud-free but the atmospheric compensated products for Loyalsock forest may be cloud contaminated.

Table 5.22: Illumination and view geometry of the Landsat 8 scenes acquired over the Harvard and the Loyalsock forests.

Forest name	Date	WRS Path / Row	angles in degrees			
			SZN	SAZ	VZN	VAZ
Harvard forest	July 11, 2015	12 / 31	26.42	131.63	-7.18	284
	August 3, 2015	13 / 30	31.16	139.15	2.38	104
Loyalsock forest	May 29, 2015	15 / 31	25.75	135.55	-4.63	283
	June 5, 2015	16 / 31	25.13	133.86	5.33	103
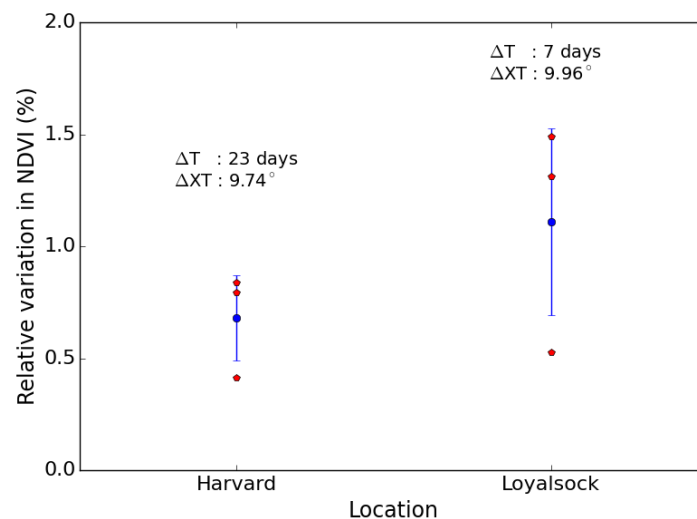
Figure 5.84: Relative variation (NDVI) between the two acquisition dates for the Loyalsock and the Harvard forests. The red dots indicate the relative variations for the ROIs.

The Harvard BRDF model is used as the scene in the DIRSIG tool to simulate the OLI sensor reaching radiance for the same sun and view geometries as observed by the real data over the Harvard forest (see Table 5.22). For a lack of better knowledge of the visibility conditions, the relative variations are estimated for the two cases: 1) assuming the visibility conditions of both the scenes are equal, (Vis(1)=20 km, Vis(2)=20 km); 2) scene acquired on August 3rd is assumed to be 15 km while the reference is 20 km (Vis(1)=20 km, Vis(2)=15 km).

In the case of equal visibility conditions, the estimated effect is due to the differences in the view and the sun angles between the two acquisitions. In the other case, the relative variation is a combined effect of the three factors (visibility, sun and view angle differences). The relative variations for the different data products (using simulated data) are shown in Appendix G (see Table G.1).

Although the effect of factors cannot be separated from the changes in the canopy using the real data, the contribution of the factors' effect in relation to the measured changes can be estimated by taking the ratio of the simulated and real data effects as shown in Equation5.4.

$$
\begin{aligned}
RV_{real} &= 100 \left| \frac{DATE(1)_{real} - DATE(2)_{real}}{DATE(1)_{real}} \right| \\
RV_{DIRSIG} &= 100 \left| \frac{DATE(1)_{DIRSIG} - DATE(2)_{DIRSIG}}{DATE(1)_{DIRSIG}} \right| \\
RATIO &= \frac{RV_{DIRSIG}}{RV_{real}}
\end{aligned}
\tag{5.4}
$$

where,
$DATE(1)_{real}, DATE(1)_{DIRSIG}$ are the responses from the real OLI and DIRSIG simulated data for 7/11/2015
$DATE(2)_{real}, DATE(2)_{DIRSIG}$ are the responses from the real OLI and DIRSIG simulated data for 8/3/2015
RV_{real} is the relative variation for the real data
RV_{DIRSIG} is the relative variation for the DIRSIG simulated data
$RATIO$ is the ratio of the simulated to real response

Note that the response in this equation can be sensor reaching radiance, TOA reflectance, surface reflectance, or NDVI generated from these data products. The ratio, thus determined will indicate the over-estimation error that are likely to occur due to effect of sensor and

environmental factors. For example, let us assume the relative change measured between two real scene equals 10%. For the same sun and view geometry, similar atmospheric conditions, and RSRs, let us assume the relative change from the simulated data is 8%. This indicates that 80% ($\frac{8}{10}$) of the measured change in the real data are likely to come from the uncompensated factors and the actual change in the canopy is only 2%, i.e, the change in the forest are over-estimated by 5 times using the real data. Similar analysis conducted using the real and simulated data for different data products are shown in Table 5.23. Note that when the value is equal to 100 in the table, then the inference is that the change observed is mainly due to the factors without any actual change in the canopy. Values greater than 100 in the table indicate that the estimated effect of the factors are higher than the measured change in the real data, which can be possible under two situations: 1) when the effect of the factors are large and positive, while the actual change in the canopy is relatively less and negative, which leads to a net effect that is positive for the real data, but smaller than the effect estimated for the factors or 2) possibility of large errors in the estimation of the factors' effects due to inaccurate assumption of the sensor and environment conditions. For example, atmospheric conditions used for the simulated data may be different compared to the real data, which may lead to over-estimation of the effects and causes the contribution ratio to increase beyond 100.

Table 5.23: Contribution of the factors in comparison to the actual change measured from the real data over Harvard forest. (*) Values greater than 100 indicate that the estimated change is higher than the measured change

Product type	Contribution relative to the changes on the ground (%)	
	Vis(1) = 20km Vis(2) = 15km	Vis(1) = 20km Vis(2) = 20km
TOA RED	174*	95
TOA NIR	47	61
TOA REFL RED	236*	95
TOA REFL NIR	24	45
ELM-ideal RED	198*	223*
ELM-ideal NIR	47	53
NDVI (TOA RAD)	331*	18
NDVI (TOA REFL)	332*	18
NDVI (ELM-ideal)	57	64

In the case of unequal visibility conditions, for the red band, the estimated effects are much

higher than the measured effects. This suggests that the visibility condition of the second scene is likely to be higher than 15 km. The effects in the red band are high even when the visibilities are equal, which indicate that the low canopy reflectance in the red band makes it difficult to ascertain the changes accurately. Further, the red band scatters considerably higher than the NIR band. So, any error in the atmospheric compensation of the real data may result in the wrong estimation of the red reflectance, which may significantly impact the relative change measured in the red band, leading to values greater than 100.

Similarity between the real and simulated data for the at-sensor radiance in the red spectral band (95%) suggests that the real atmospheric and sensor conditions are very well approximated in the simulation as the canopy reflectance are very low (and does not change significantly) in the red spectral band and the radiance reaching the sensor is primarily due to the upwelled radiance. This provides more confidence in the results estimated for the same visibility conditions between the two dates. Assuming that the two scenes were acquired under the same visibility conditions, in the NIR band, it is observed that about 50% of the measured change can be attributed to the factors. Similarly, in the case of NDVI products that are atmospherically compensated, more than 60% of the contribution are likely to be caused by the factors. If the NDVI products are not atmospherically compensated, for similar atmospheric condition between the two scenes, the observed effect is low, suggesting that the atmospheric perturbation reduces the perceived effect due to the sun and the view angle differences. This suggests that the effect of the factors tend to become more apparent when the atmosphere is compensated.

From these analysis, we can conclude the following: 1) the simulated data appears to be a reasonable surrogate for real data, and should be understood as indicative of a deciduous forest canopy but not an exact match to a specific forest, and 2) the factors can contribute as much as 50% of the observed change in the real data, and therefore, the change detection analysis from the real data are like to over-estimate the actual changes in the canopy.

5.8 Factor screening experiments

This section discusses the results from the factor screening experiments for five different forest sections (center, UL, UR, LL, LR) as shown in Figure 5.44 and for the randomly generated forest BRDFs as discussed in section 4.4.2.2. The factorial experiment is repeated six times

(center, UL, LL, LR, UR, random) to evaluate the consistency in identifying the significant factors. In the case of the five forest sections, all the runs in an experiment use the same reference BRDF, i.e., of the 64 runs in the experiment, 32 of them use the same $+25°$ slope BRDF while the remaining 32 runs use the same $-25°$ slope BRDF. The five experiments differ in their BRDF depending upon the section of the forest used. For the random BRDF factorial experiment, all the experimental runs use randomly generated BRDFs, i.e, the BRDF for 32 runs for each slope are different from each other. Each factorial experiment is independently analyzed for six spectral bands (Blue, Green, Red, NIR, SWIR-1, SWIR-2) with sensor-reaching radiance as their response variable. The ANOVA table and the interaction plots for the center and random forests are shown in Appendix F. Tables 5.24 and 5.25 shows the p-value and the main effects from the factorial experiment for all the six experiments in the NIR spectral band after model reduction. The effects in the factorial experiment corresponds to an average change in radiance ($Wm^{-2}sr^{-1}\mu m^{-1}$) from one level to another. The main effects for all the other spectral bands are shown in Appendix F.

Table 5.24: The p-value from the ANOVA table for all the 6 experiments are summarized for the NIR spectral band. The "*" indicates the effects of the factor that are insignificant in relation to the other factors when $\alpha = 0.05$

Factors	LR	LL	UR	UL	Center	Random
Solar Zenith	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Visibility	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Terrain slope	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
X-Track	0.278*	0.004	< 0.001	0.035	0.07 *	0.036
Solar Azimuth	0.001	0.297*	0.118*	0.323*	0.344*	0.326*
RSR	0.25 *	0.57 *	0.512*	0.603*	0.691*	0.639*
Along Track	0.562*	0.562*	0.568*	0.831*	0.873*	0.639*

From the seven factors in the factorial experiment, only four factors were identified as significant. For the case of the NIR spectral band, the results from the main effects table indicate that the solar zenith angle (SZN) is the most contributing factor followed by visibility, terrain slope and across track (X-Track). The factor effects for SZN, visibility and terrain slope are consistent across all the six factorial experiments. For the center forest, the X-track was marginally insignificant whereas, it was found to be a significant factor for all the other sections except the lower right (LR) forest. For the LR section, solar azimuth was found to be significant while X-Track was insignificant. No other experiment showed that the solar azimuth factor was significant and hence it is assumed to be insignificant. The X-Track factor although not

Table 5.25: The main effects in the NIR spectral band for all the six factorial experiments. The "*" indicates the effects of the factor that are insignificant in relation to the other factors. The values are equivalent to the slope of the regressors for the two-level factorial experiment and large value indicates that the factors are significant.

Factors	LR	LL	UR	UL	Center	Random
Solar Zenith	−84.84	−83.0	−81.37	−76.12	−81.33	−81.07
Visibility	−22.46	−20.35	−20.4	−17.48	−19.79	−19.77
Terrain slope	4.19	8.83	6.29	10.52	8.01	8.38
X-Track	0.22*	1.25	1.5	0.99	1.07	1.10
Solar Azimuth	−0.71	0.42*	0.61*	0.45*	0.55*	0.51*
RSR	0.23*	0.23*	0.25*	0.23*	0.23*	0.24*
Along Track	−0.12*	−0.23*	−0.21*	−0.09*	−0.09*	−0.01*

as significant as the other 3 factors and is marginally insignificant in the center forest factorial experiment, can interact with the other factor effects. The interaction of X-Track with SZN and terrain slope is found to be significant as shown in the ANOVA table(see Appendix F). Since the X-Track interactions are significant, by the principle of hierarchy, the main effects for the X-Track are considered significant during model reduction. Similar analysis for the red spectral band is shown in Table 5.26. The results clearly indicate that the main effects are consistent across all the six factorial experiments. In this case, the relative spectral response factor (RSR) is found to be significant along with the other four significant factors from the NIR results. This is because the RSR for the OLI and MSI sensors have differences in their shape (and a slight shift in the central wavelength) for the red spectral band as shown in Appendix E.2. However, in the case of the NIR spectral band, the RSR shape and bandwidth for the OLI and MSI sensors are very similar and hence the RSR was found to be an insignificant factor in the factorial experiments. In comparison to the NIR results, visibility is found to be the most contributing factor rather than SZN for the red spectral band. This is expected since the atmospheric attenuation is higher in the visible region (blue, green, red) than in the NIR or SWIR spectral bands. The significant factors and their order of significance from the main effects for all the spectral bands are shown in Table 5.27.

It can be easily inferred from the table that the four factors (SZN, visibility, terrain slope, X-Track) are found to be significant across all the spectral bands. Further, the RSR factor is found to be significant for those spectral bands whose RSR shape differs between the two sensors. The solar azimuth factor, though found to be significant for certain spectral bands (blue, SWIR-1, SWIR-2), has an extremely small effect compared to all the other factors. For

Table 5.26: The main effects in the red spectral band for all the six factorial experiments. The ''*'' indicates the effects of the factor that are insignificant in relation to the other factors.

Factors	LR	LL	UR	UL	Center	Random
Visibility	−103.04	−102.96	−103.25	−103.22	−103.15	−103.14
Solar Zenith	−38.52	−38.54	−38.73	−38.83	−38.71	−38.68
RSR	−2.28	−2.29	−2.28	−2.30	−2.3	−2.30
Terrain slope	0.75	0.73	0.83	0.75	0.80	0.76
X-Track	−0.25	−0.25	−0.24	−0.25	−0.26	−0.26
Solar Azimuth	−0.04*	−0.04*	−0.04*	−0.04*	−0.04*	−0.04*
Along Track	0.01*	0.01*	0.01*	0.01*	0.01*	0.02*

Table 5.27: Significant factors identified using fractional factorial analysis for different spectral bands. The factors are ordered by the magnitude of their effects.

Blue	Green	Red	NIR	SWIR-1	SWIR-2
Visibility	Visibility	Visibility	SZN	SZN	SZN
SZN	SZN	SZN	Visibility	Visibility	Visibility
RSR	Slope	RSR	Slope	Slope	Slope
X-Track	RSR	Slope	X-Track	X-Track	X-Track
Slope	X-Track	X-Track		SAZ	RSR
SAZ					SAZ

all the spectral bands, the along track factor effect is found to be insignificant (see Table 5.26 and Appendix F). This is expected since the along track angles are very small ($\pm 2°$) for these sensors. Thus, from the factor screening experimental analysis, it is concluded that five factors (SZN, visibility, terrain slope, RSR, and X-Track) are important and these are modeled in the regression analysis.

5.9 Regression analysis

Section 4.9 discussed the approach used to build the polynomial regression model. In this section, the results and the analysis of the regression model are discussed. The five significant factors, identified from the factor screening experiment, are used to construct the model. This research primarily focuses on the two sensors (OLI and MSI), and therefore, the RSR of these two sensors define the two levels of the factor. The other four factors are continuous variables, and so 5 distinct levels are chosen to represent their range. The levels for each factor is chosen such that they span the range of conditions expected from a typical remote sensing observation.

The levels for the 5 factors are shown in Table 5.28. For some factors, the levels are unequally distributed as their variations are typically non-linear. For example, visibility factor has a non-linear effect on the radiance reaching the sensor and its impact is more pronounced at low levels.

As explained in Section 4.9, the levels for each factor are transformed to reduce the complexity of the model. It is beyond the scope of this research to find the most appropriate transformation for each factor, however attempts were made to find a reasonable transformation. The main effects plot for the different transformations are shown in Appendix H (see Figures H.1, H.2, H.3, and H.4). The function that appears to linearly fit the data are used as the functional variables and are shown in Table 5.28. For lack of better functions for the terrain slope and across-track factors, the angles are converted from degrees to radians. For the terrain slope, the angles in radians exhibit a linear trend for the red response and a non-linear trend for the NIR response. This can lead to a regression model with higher order polynomials in the terrain slope for the NIR response. In the case of X-Track, a non-linear trend is observed for all the three functional variables and therefore, higher order polynomials in the regression model are likely. The result of the transformation functions for the logarithmic response variable is not shown as it is similar to the regular variable. The transformed levels are used to estimate the orthogonal polynomial contrasts for each factor, which are shown in Appendix H (see Tables H.1, H.2, H.3, and H.4).

Table 5.28: Factors' levels and their functional variables used in the regression analysis.

Factor name	Transformation function	Levels
RSR		OLI , MSI
X-Track	radians	$-15°$, $-7.5°$, $0°$, $7.5°$, $15°$
Visibility	ln(vis)	5 , 7 , 10 , 15 , 25
Terrain slope	radians	$-25°$, $-15°$, $0°$, $15°$, $25°$
SZN	cos(SZN)	$1°$, $15°$, $30°$, $40°$, $55°$

The regression model is built using 4 factors with 5 levels each and a factor with 2 levels, which results in 1250 experimental runs $(2*5^4)$ for each response variable in a full factorial design. The number of terms in the regression model can be estimated from the multinomial theorem. The 5^4 model requires 69 terms for the fourth-order polynomial model, as discussed in Section 4.9. The number of terms increase by 35 additional terms due to the RSR factor, which contributes

1 main effect term (linear) and 34 interaction terms (2-f.i : $3 * \binom{4}{1} = 12$, 3-f.i : $1 * \binom{4}{2} = 6$, 4-f.i : $2 * \binom{4}{2} + 1 * \binom{4}{3} = 12 + 4$). In total, there are 105 terms in the regression model including the bias term, and the degree of the polynomial is 4.

A combination of full factorial and fractional factorial design is constructed such that the design is full factorial for the factor with 2 levels (RSR) and fractional design for the factors with 5 levels (4 factors). Each fraction consists of 250 observations of which 125 observations are simulated with the OLI RSR and the remaining with the MSI RSR. One of the fractions is used for validation, while the remaining fractions are used for training the model using a K-fold modeling technique with K = 4. The training set is equally divided into 4 folds, one for each fraction. Of the 4 subsamples, one of them is used for testing the model, and the remaining 3 subsamples are used as training data. This is repeated such that each of the 4 subsamples are used exactly once as the testing data to build 4 model sets. Each model set uses 750 observations for training and 250 observations for testing the model. The four training sets are used to construct four regression models using the forward stepwise selection method for all the 105 terms. The RMS for each model is calculated from their corresponding test set. The RMS at each step in the forward selection method for all the four models are shown in Figure 5.85. The RMS of the four models are similar, and although not shown here, the order of the terms in the four regression models were mostly consistent. The mean RMS of the four models are compared with the system noise estimated for the OLI sensor. For the regular response variables (Figures 5.85a and 5.85c), the system noise model coefficients published by Morfitt et al. (2015) were used to estimate the noise for an average response, whereas they were estimated indirectly for the logarithmic response variables (Figures 5.85b and 5.85d). First, a large number of normally distributed random numbers are generated using the average response as mean, and the noise estimated for the regular response variables as STD of the normal distribution. Then, the system noise for the log transformed response variable is estimated from the STD of the log transformed random numbers.

(a) RMS : RED

(b) RMS : ln(RED)

(c) RMS : NIR

(d) RMS : ln(NIR)
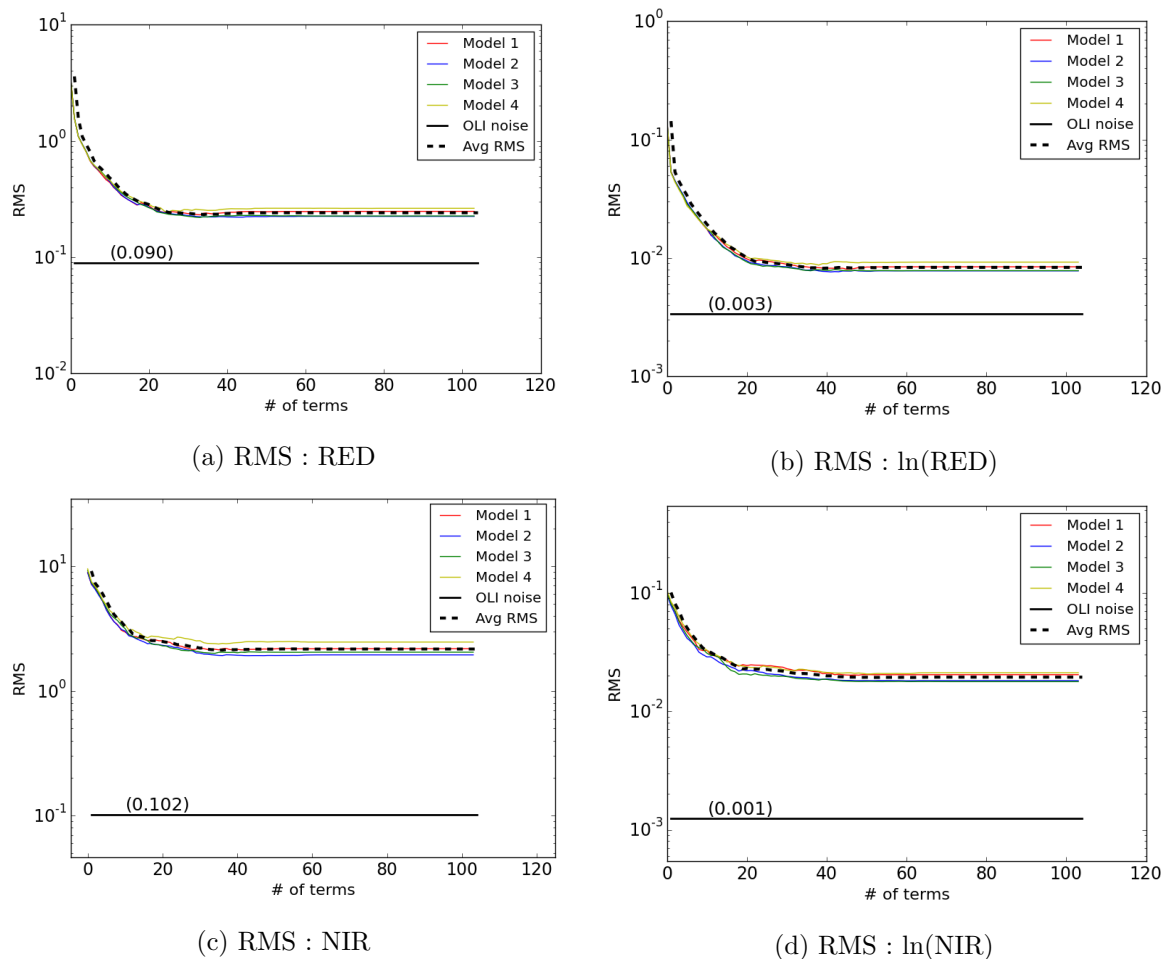
Figure 5.85: RMS error in the stepwise regression model for all the response variables. Note that the RMS for all the models that includes all the terms are still higher than the system noise level (Avg RMS - black solid line) of the OLI sensor.

The cross validation error (blue solid line) and the standard error (error-bar) computed for the regular and logarithmic response variables are shown for in Figures 5.86 and 5.87. Figures 5.86b and 5.86d show the zoom-in view of the CV error plot for the red and the NIR responses. The blue and red dashed lines indicate the location of the minima and the optimal number of terms selected from the 4-fold CV technique respectively. Comparing the regular and the logarithmic response variables, it is clearly evident from the figure that the CV method for the regular response variable selected fewer number of terms for both the red and the NIR response variables. This suggests that the untransformed response variables are a better choice than the log transformed response variables, and hence, the regression model is constructed only for the regular response variable. The red and the NIR response variables require 26 and 30 terms in the model respectively. The terms chosen for the red and the NIR response variables are different and it supports the fact that the sensor reaching responses are spectrally dependent.
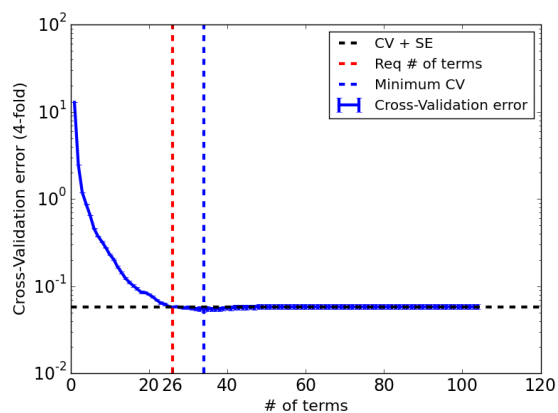
(a) CV error : RED
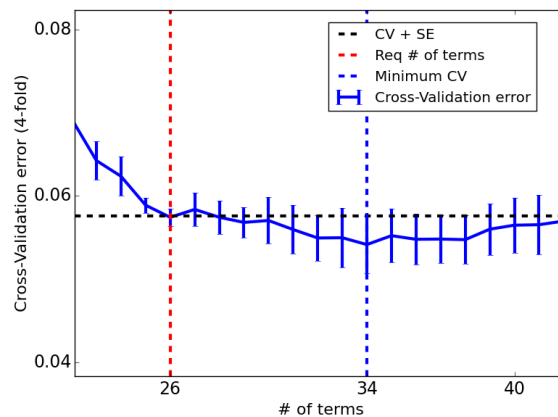
(b) Zoom-in : RED
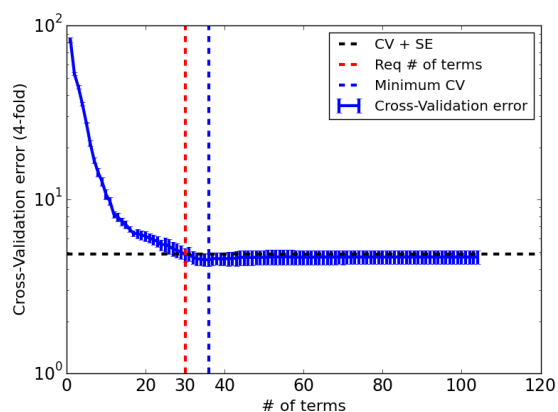
(c) CV error : NIR
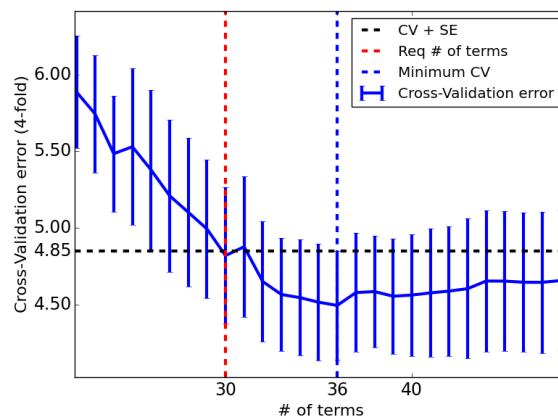
(d) Zoom-in : NIR

Figure 5.86: Cross validation error in the stepwise regression model for the two response variables. The number of terms selected using the CV approach for the red and NIR spectral bands are 26 and 30 terms respectively.

(a) CV error : ln(RED)
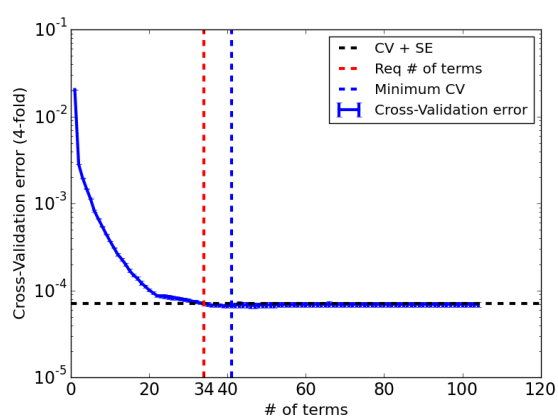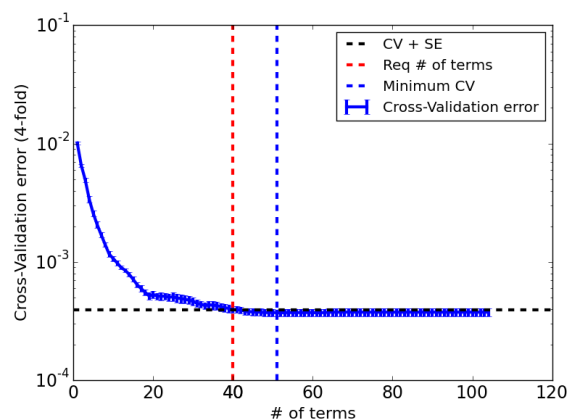
(b) CV error : ln(NIR)

Figure 5.87: Cross validation error in the stepwise regression model for the two log response variables

All the observations in the training data (1000 observations) are used to construct the polynomial regression model using the forward selection method. The regression model is evaluated using the validation dataset (250 observations) and the summary statistics are shown in Table 5.29. A small increase in the number of terms after employing the hierarchy principle indicates that the constructed model using the forward selection is well-structured. The model's predictive RMS is approximately 3 and 24 times higher than the estimated noise of the OLI sensor for the red and the NIR responses respectively. This suggests that the red response can be predicted very well using the regression model, but the errors in the NIR response are somewhat larger. This is not unexpected, as the reflectance in the NIR bands can vary to a larger extent than in the red wavelengths for forest canopies. Further, the range and the variations in the NIR at-sensor responses are also much higher compared to the red responses. Although the models are not accurate to predict within the system noise levels of the OLI sensor, the predictions are comparable to the ETM+ sensor (see Table 3.2), which is still being used by the remote sensing community. More importantly, the primary objective of this research is to identify the relative significance of the factors and not to generate an accurate model of the radiance reaching the sensor, therefore, the observed error in the model is inconsequential.

Table 5.29: Summary statistics for the regression model

	RED	NIR
Num of terms (CV)	26	30
Num of terms (Hierarchy)	28	38
R^2	0.9989	0.9941
$adj R^2$	0.9989	0.9938
RMS	0.24	2.11
System noise	0.09	0.102

The relative significance of the factors shown in Tables 5.30, 5.31, and 5.32 for the red and the NIR responses are determined from the ratio of the Sum of Squares (SS) of each factor to the SS explained by the model. The contribution of each term is grouped by the order of the polynomial for each factor to help assess the contributions due to the higher order terms. Comparison of the first order and higher order terms for the red response variable reveals that more than 96% of the contribution is mainly due to the linear terms in the model. This suggests that a linear factorial experiment such as two-level factorial experiments can adequately explain the relative significance of the factors. The second order term contributes about 3%, whereas, the interaction terms contribute only about half-a-percent, but higher than the residuals (0.1%),

indicating that the second order and the interaction terms are necessary in modeling the response variable to reduce the error, but not directly useful in assessing the relative significance. The X-Track factor is observed to contribute more than the RSR factor, but their relative order of importance reverses when only the first order effects are considered, suggesting that the X-Track and the RSR are likely to produce similar effect for the red response. The contributions due to the terrain slope are about the same as the residuals in the model, indicating that the terrain slope may not be a critical factor for modeling. The low contribution of the terrain slope is expected, as the variations in the reflectance are very small (1 - 2%) in the red wavelengths, but are significantly higher in the case of NIR wavelength (see Figures 5.42). Large contribution due to the visibility factor implies that the atmospheric attenuation needs to be compensated for reducing the overall uncertainty in the biophysical estimation.

Table 5.30: Relative significance of the factors for the red response variable

Factors name	1st order	2nd order	3rd order	4th order	Interaction terms	All terms
			Contribution (%)			
VIS	75.71	2.41	0.02	0.01	0.05	78.20
SZN	19.72	0.37	0.10	0.01	0.19	20.39
X-Track	0.42	0.09	< 0.01	< 0.01	0.18	0.70
RSR	0.60				0.01	0.61
SLOPE	0.09				< 0.01	0.10
Total	96.55	2.87	0.12	0.03	0.43	100

Table 5.31: Relative significance of the factors for the NIR response variable

Factors name	1st order	2nd order	3rd order	4th order	Interaction terms	All terms
			Contribution (%)			
SZN	88.05	1.35	0.65	0.05	1.11	91.21
VIS	4.55	0.04			0.72	5.31
SLOPE	1.27	1.01	< 0.01	0.02	0.24	2.54
X-Track	0.29	0.26	< 0.01	0.01	0.36	0.93
RSR	< 0.01				< 0.01	0.01
Total	94.17	2.66	0.66	0.08	2.43	100

Similar to the red response variable, the first order contributes approximately 95% of the explained variation in the model for the NIR response variable. Unlike the red response, the

Table 5.32: Relative significance of the interaction terms for the NIR and red response variables

| | Contribution (%) | |
Terms	Red	NIR
2-f.i	0.43	2.38
3-f.i	-	0.05
Residuals	0.10	0.59

interaction terms contribute as high as the higher order terms. The contributions due to the 3-factor interaction (f.i) terms are statistically insignificant and can be ignored. As in the case with the red response, the addition of higher order and interaction terms are useful for modeling the response variable but does not provide any additional information in deciding their relative significance. The contribution of RSR is insignificant due to the similar shape and spectral wavelength of the two RSRs. The contribution of the visibility factor in the NIR response is significantly smaller than the red response due to higher atmospheric transmission in the NIR wavelengths. The higher order terms for the terrain slope and X-Track factor can be attributed to the inaccuracy in the transformation functions for these independent variables. The contribution due to the solar zenith angle is observed to be 20 times higher than the next contributor, implying that the SZN angle effect should be compensated for reducing the uncertainty in the derived products.

The main and interaction effects' plots are shown in Figures 5.88 - 5.93. The slope of the main effects curve for the red response variable shows that the SZN, visibility and the RSR factors have steeper slope than the other two factors, indicating their relative importance. In the case of the NIR response variable. the slope of the RSR factor is flat, indicating its insignificance. The 2-f.i interaction plots between any two factors for the red response variable shows parallel lines across all the levels for most of the 2-f.i plots, suggesting that the 2-f.i terms are not significant. For the NIR response variable, few 2-f.i terms show non-parallel lines indicating their significance, and the 2-f.i plots that show near-parallel lines were selected lower in the order during stepwise regression, revealing the consistency in the forward selection method. The results observed in the interaction plots are consistent to the results explained in Tables 5.30 and 5.31.
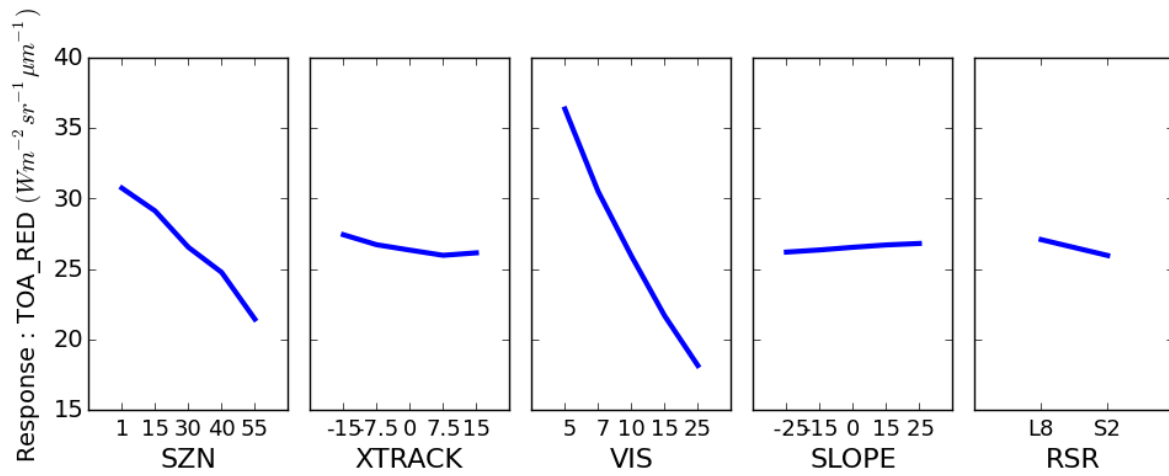
Figure 5.88: Main effects plot for the red response variable. The X-axis corresponds to different levels of the factors and the Y-axis represent the response variable in $Wm^{-2}sr^{-1}\mu m^{-1}$.
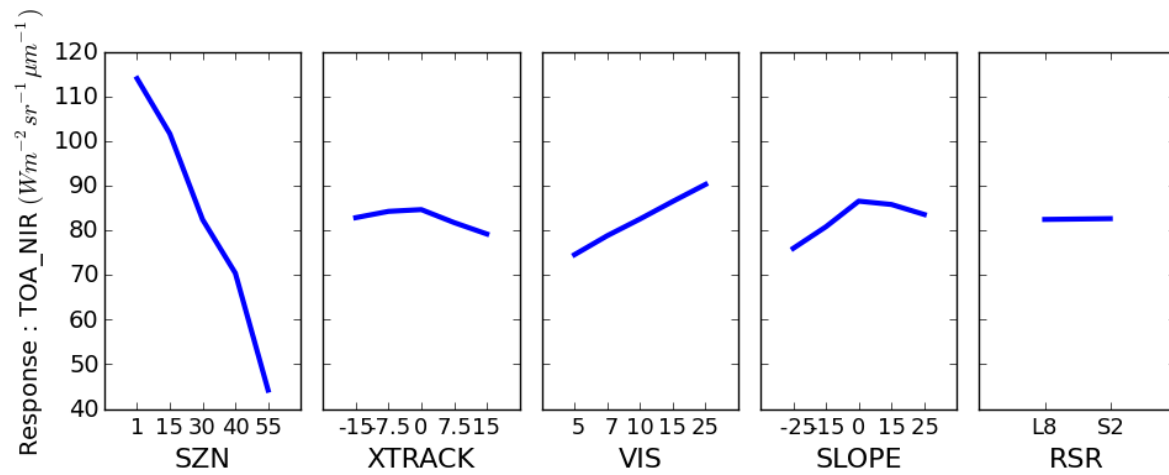
Figure 5.89: Main effects plot for the NIR response variable. The X-axis corresponds to different levels of the factors and the Y-axis represent the response variable in $Wm^{-2}sr^{-1}\mu m^{-1}$.
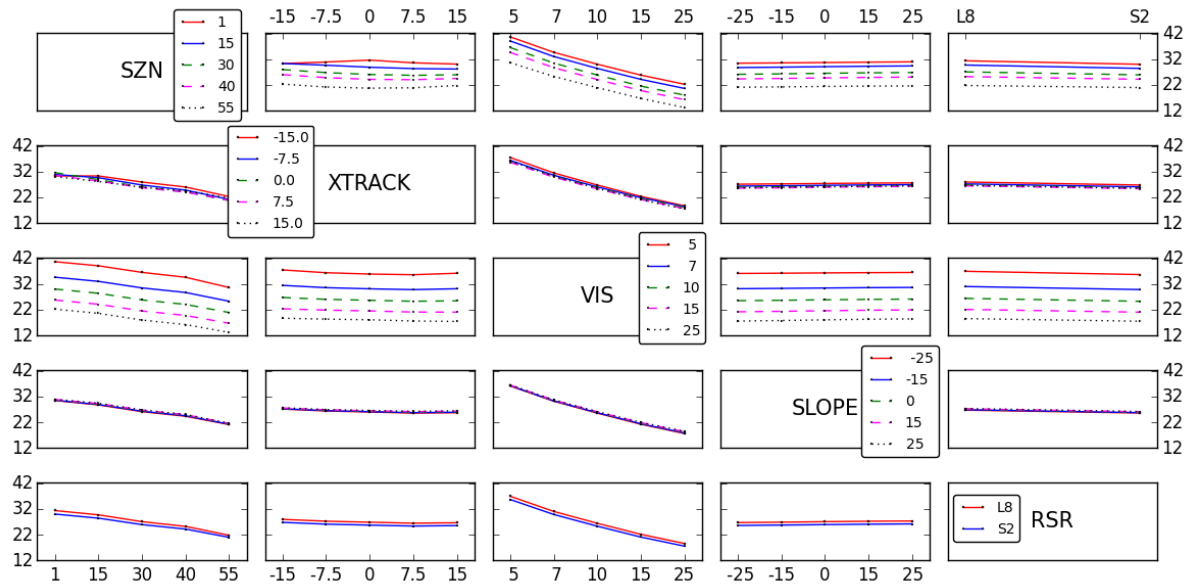
Figure 5.90: Two factor interaction effects plot for the red response variable. The X-axis corresponds to different levels of the factors and the Y-axis represent the response variable in $Wm^{-2}sr^{-1}\mu m^{-1}$.
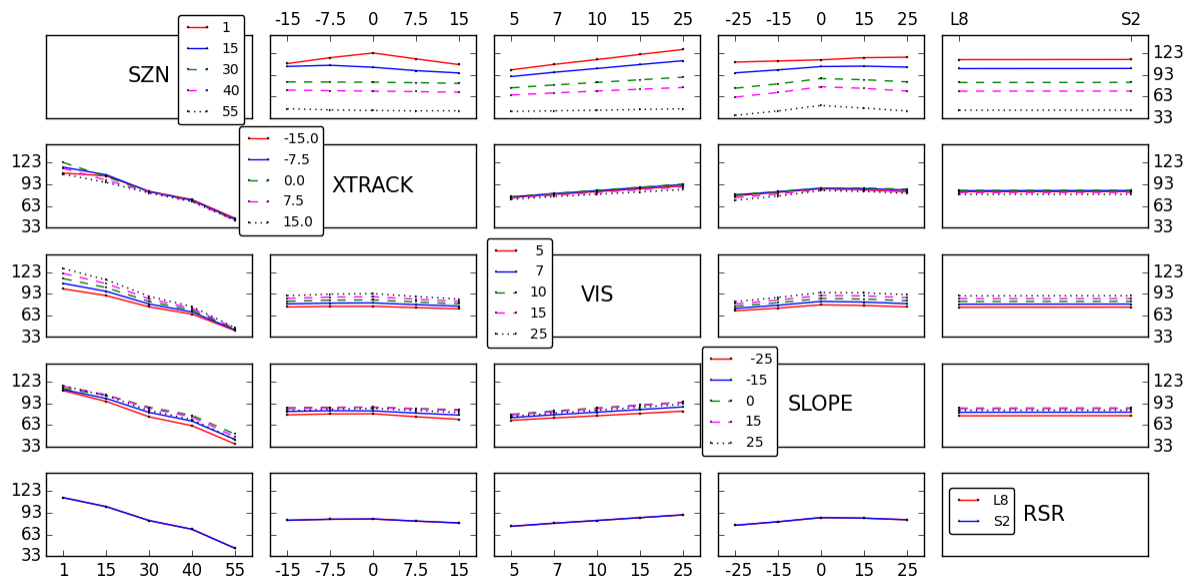
Figure 5.91: Two factor interaction effects plot for the NIR response variable. The X-axis corresponds to different levels of the factors and the Y-axis represent the response variable in $Wm^{-2}sr^{-1}\mu m^{-1}$.
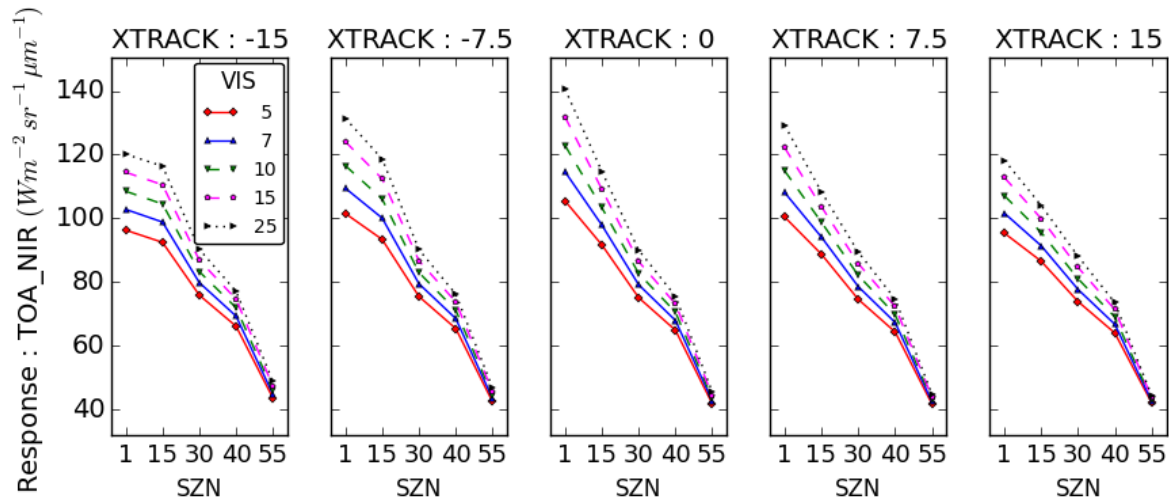
Figure 5.92: Three factor interaction effects plot (visibility, SZN, X-Track) for the NIR response variable. The X-axis corresponds to different levels of the factors and the Y-axis represent the response variable in $Wm^{-2}sr^{-1}\mu m^{-1}$.
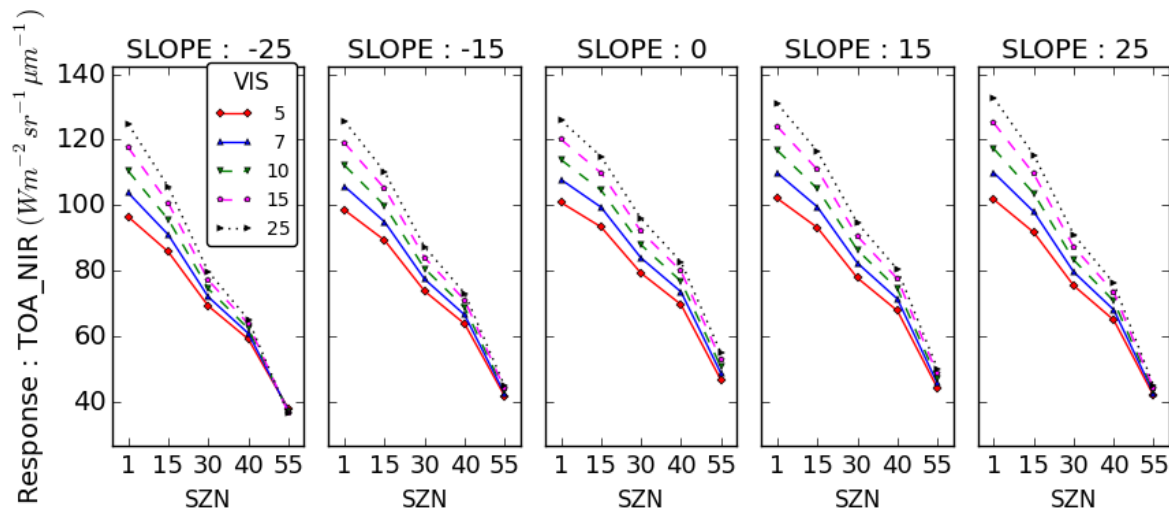
Figure 5.93: Three factor interaction effects plot (visibility, SZN, slope) for the NIR response variable. The X-axis corresponds to different levels of the factors and the Y-axis represent the response variable in $Wm^{-2}sr^{-1}\mu m^{-1}$

.

In summary, the regression analysis indicate that the visibility and the solar zenith factors are the primary contributors to the variation observed in the sensor reaching radiance, and the effect of other factors are comparatively small. Furthermore, the higher order and the interaction terms are found to be statistically significant but their contributions have no impact on the factor's relative order of significance.

5.10   Summary

This chapter presented the results and findings of the validations of the canopy radiative transfer in DIRSIG, BRDF modeling of the forest canopy, characterization of the forest defoliation, factor sensitivity studies and factorial experiments. The qualitative and quantitative comparison of DIRSIG simulations with the existing radiative transfer validated the ray tracing model of DIRSIG for forest canopy studies. This was followed by 3D modeling of the Harvard forest canopy in DIRSIG using the tree modeling tools and ground collected optical properties. The sampling and model fitting of BRDF measurements to the canopy BRDF model, and its sensitivity analysis revealed that the RossLi canopy BRDF model is extremely useful in modeling the deciduous forest canopy. The characterization of defoliation showed the prospect of modeling the biophysical parameters in a simulation environment. Finally, the effects of the factors were analyzed using a sensitivity study based on the analysis of one-factor at a time and using design of experiment techniques. The analysis showed that the results from the two techniques were consistent and the sensitivity analysis can be extremely useful in describing the effect as changes on the ground. The following chapter details the conclusions drawn from this study and how it can be extended in the future.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

6.1.1 DIRSIG validation

An integral part of this research effort is dependent on the ability of the DIRSIG tool to simulate the real-world conditions accurately. In this research, the DIRSIG model is used primarily for estimating the radiative transfer in the canopy and to simulate the responses, as observed by a sensor in space. The validation of DIRSIG in simulating the ground observations accurately for a variety of sensors has been studied extensively in the past (Brown and Goodenough, 2015, Brown and Schott, 2010), but its capability to evaluate the complex interactions within a canopy were not explored. Consequently, the validation for the radiative transfer within the canopy was performed in this research using qualitative and quantitative methods. The qualitative method compares the observed result with the expected outcomes based on the intuitive knowledge of light interaction within the canopy, thus eliminating the possibility of gross errors in the DIRSIG model. The quantitative method makes use of the benchmarked scenes designed by RAMI (RAMI III) to compare against other radiative transfer models. The RAMI III scenes are structurally unrealistic and abstract models, however, at the time of validation, results for more complex realistic RAMI IV scenes were unavailable. Further, the RAMI benchmarks have no "true" solution to measure an error, hence the DIRSIG results were validated for its consistency against the other RT models. The validation results suggest that, the BRF computed using the DIRSIG model is consistent with the results published by RAMI for all the test cases evaluated

in this research. While the test scenes are relatively simple compared to an actual canopy, the radiative transfer process within the DIRSIG model remains the same, irrespective of the complexity of the scene, and thus ensuring its validity for any type of forest scenes. These validations have given enough confidence in the DIRSIG tool to model the complex interactions expected in forest canopies.

6.1.2 Forest canopy model

It has been shown that forest canopies can be modeled in a virtual environment but its realistic representation is very much dependent on the modeling tools such as OnyxTree and PROSPECT. The OnyxTree software has shown immense potential in modeling deciduous trees, but its use is limited due to its inability to model trees based on forest inventory information directly. An accurate estimation of the biophysical parameters and reconstruction of reflectance using PROSPECT inversion model in this work supports its capability, and can be extended to model the optical properties at different stages of leaf growth (e.g. senescence). The measurement of canopy BRDF in this work demonstrates the usefulness of DIRSIG in modeling the complex canopy interactions which are difficult to measure in the real-world. Further, the number of reflectance measurements can be reduced significantly by sampling strategies that depend on the geographic locations and sensor parameters. It was shown that the BRDF measurements can be used to fit to the RossLi BRDF model to within 1% relative error for low solar zenith angles and less than 5% relative error for high solar zenith angles. Although the fit error is high for high solar zenith angles, they are insignificant and negligible, as most remote sensing observations of the forest canopies are performed at low solar zenith angles.

6.1.3 RossLi BRDF model

The modeling of BRDF using the RossLi BRDF model has not only demonstrated its ability to fit the DIRSIG measurements accurately, but also can be considered as one of the effective canopy models to represent the deciduous forests. Koukal and Schneider (2010) compared the accuracy of the different BRDF models and concluded that the RossLi BRDF model for different biomes is as accurate as leading non-linear models such as the RPV model. Although they have been validated using the measured data at coarse resolution, high fidelity measurements in

a simulated environment in this research, helped to evaluate the sensitivity of the model to understand its effectiveness and limitations.

The measurement uncertainty was evaluated using the LOOC method, wherein, a set of measurements are removed randomly, and the model coefficients are estimated based on the DIRSIG measurements for the remaining observations. The differences in the model coefficients and its effect in reflectance were used as metrics to evaulate the sensitivity of the model to the measurements. The summary statistics on the measurement sensitivity metrics suggest that the model is insensitive to the selection of view angles, but is likely to be sensitive for sun angles. However, upon analyzing the effect of the sensitivity in reflectance reveals that the relative error is less than $\frac{1}{5}^{th}$ of a percent, which is extremely small and beyond the measuring capability of many instruments.

For evaluating the model uncertainty and its accuracy, view angles were sampled at a high resolution in the zenith and azimuth directions and the model fit statistics were used to test the validity of the model. The canopy model is found to be accurate for modeling when the measurements are from low view zenith ($< 40°$) and sun zenith angles ($< 50°$). The fit RMS was observed to be less than 3%, indicating an accurate modeling of the measurements simulated using the DIRSIG tool. The RossLi model may be applicable even for larger view zenith angles, but could not be validated due to the limited extent of the forest canopy. Furthermore, the model is observed to be spectrally insensitive in the VIS-NIR-SWIR regions, suggesting that an accurate spectral BRDF can be generated for deciduous canopies using the RossLi BRDF model.

6.1.4 Comparison with real data

Although the DIRSIG tool, forest models, and the canopy models were validated, any concerns on the simulated results cannot be alleviated without comparing it with the real data. In this research, the reflectance from the Landsat's surface reflectance product (generated by USGS) for the known view and illumination angles over the Harvard forest was compared with the corresponding simulated canopy BRDF.

Interestingly, the reflectance from the simulated BRDF matches the surface reflectance of the Harvard forest to within 2% in both the NIR and the red spectral bands. Similar analysis

using a different scene (acquired in August) indicated differences greater than 10%. This is not unusual due to the static nature of the simulated BRDF compared with the changing aspect of the real data. It is important to understand that the goal of the comparison was to ensure a plausible deciduous forest canopy can be modeled and not an exact replica of the real forest. This research has sufficiently proved that complex canopies can be modeled accurately and can even match the spectral characteristics of the real data. The similarity between the real and simulated data demonstrates the validity of the process used in generating the canopy BRDF.

6.1.5   Forest defoliation

The advantage of using a virtual environment comes to the fore when modeling the biophysical parameters is of primary importance. In this research, the modeling of defoliation in the forest was demonstrated by varying the geometric properties of the trees. The process of modeling the 3D geometry and its optical properties to generate the canopy BRDF (see Rengarajan and Schott (2016)), was used to model the defoliation by randomly removing the leaf facets from the trees.

In most applications, the defoliation of the forest is inferred from the change detected on the higher level products such as NDVI data. Therefore, a similar change difference method is used in this research to characterize the defoliation using the simulated sensor data. It was shown that the reflectance in the principal and cross planes vary non-linearly with the LAI and can introduce difference ranging between 5% and 60% in the NIR and red spectral bands.

The characterization of defoliation for the sensor reaching radiance, TOA reflectance, atmospherically compensated, and NDVI products suggest that the change detection metric (relative variation) varies exponentially with the level of defoliation of the canopy. However, it can be approximated to a linear function when the level of defoliation is less than 40%, ie. for forests with LAI varying between 3 to 5, changes in the LAI can be modeled as a linear function. The predictive error using the modeled exponential function showed a relative error of about 6% and 2% in the red and NIR spectral bands respectively, thus indicating the validity of the model. In general, the NDVI is assumed to vary linearly for sparse canopies with low LAI (Wang et al., 2005), but this research has demonstrated that the relationship is better approximated by an exponential function of the form, $\text{NDVI} = a - e^{-b\sqrt{LAI}}$, over the full range of canopy densities. Although the NDVI has been the choice of many researchers for forest

applications, a small rate of change in NDVI with respect to LAI suggests that a metric better than NDVI may need to be considered in the future for less precise instruments or sensors with low SNR.

6.1.6 DOE techniques

In this study, the DOE techniques such as factorial experiments and regression analysis were performed with the intent to explore the important factors and their relative contributions to the sensor reaching radiance. Initially, the two level fractional factorial experiment, which fits a linear model, was conducted to screen for the important factors from seven factors over a range of predefined values. Results from this experiment indicated that the along-track and solar azimuth angles are not significant in comparison to the other five factors, namely visibility, solar zenith, RSR, X-Track and terrain slope. The five factors were then analyzed using a polynomial regression model after transforming the independent variables to appropriate functions. The cross validation (CV) approach and the stepwise regression method were found to be useful in determining the required number and the appropriate terms needed in the regression model. The CV method showed that the red spectral band can be modeled with 28 terms, whereas, the NIR spectral band required 38 terms, as the reflectance of the canopy has a larger variation in the NIR bands. Although the predictive RMSE of the regression model is higher than the system noise levels of the OLI sensor, they are useful and comparable for relatively low SNR systems such as the ETM+ sensor. Findings from the regression analysis revealed that, more than $\frac{3}{4}^{th}$ of the contributions are mainly due to a single factor: visibility conditions for the red spectral band and SZN angles in the case of NIR spectral band. The contributions of the higher order and interaction terms are significantly smaller than the first order terms, indicating that the sensor reaching radiance can be adequately modeled using a linear regression model with interaction terms.

The conclusion derived from the regression based analysis indicates that the visibility conditions and the SZN angles are the most important factors that impacts the sensor reaching radiance in either of the spectral bands.

6.1.7 Effect of factors

The effect of factors cannot be separated in the real data as they are inherent in the process of imaging, but simulated data eliminates such issues and provides the capability to study the effect of a specific factor while keeping everything else constant. The effect can be accurately realized only when the models and the modeling tools are precise and accurate, and this was well established in this research through the validation process of the DIRSIG tool and the simulated canopy BRDFs. To date, there are no better ways to represent the effects in meaningful units that are well understood by foresters. In that regard, we have demonstrated the use of Noise Equivalent Defoliation (NED) as a metric to indicate the effects in terms of the uncertainty in detecting (or estimating) defoliation. Furthermore, the characterization of forest defoliation helped to assess the effect of the factors for different types of data products used by the scientific community. Thus, not only the effects of a specific factor was studied, but how those effects change with additional compensation techniques was also explored.

The sensitivity analysis in this study included four main factors: RSR, across-track, visibility differences, and the changes in the SZN angles expected between two dates and/or two sensors. For time series applications, the effects due to the factors affects the uncertainty in the estimation, i.e, the effect is equivalent to the expected uncertainty in the estimated change. The effects in this study are represented in NED unit, which is defined as the amount of defoliation that is contributed due to the effect of a specific factor. In other words, the effect due to a specific factor is equivalent to the effect that would be observed when the forest defoliates by a certain amount. A summary table showing the impact in defoliation unit (NED) for each factor is shown in Table 6.1.

Table 6.1: Effect of each factor for NDVI products in NED units.

Data products	Visibility	SZN (ΔSZN=5°)	RSR	X-Track (±12°)
NDVI TOA Rad	40%	10%	19%	14%
NDVI TOA Refl	40%	10%	19%	14%
NDVI TOA Rad + SBAF Adj	40%	10%	7%	14%
NDVI TOA Refl + SBAF Adj	40%	10%	7%	14%
NDVI Surf Refl (ELM-Typical)	16%	8%	15%	40%
NDVI Surf Refl (ELM-Ideal)	15%	4%	15%	9%
NDVI Surf Refl + SBAF Adj (ELM-Typ)	7%	8%	1%	40%
NDVI Surf Refl + SBAF Adj (ELM-Ideal)	1%	4%	1%	9%

In the case of RSR differences, as observed between the OLI and MSI sensors, the effect is equivalent to a change observed by the forest that defoliates about 20% (NED $\approx$ 20%). However, this can be significantly reduced to about 1% when RSR compensation techniques such as SBAF are employed for the atmospherically compensated products, thereby suggesting that the effect due to the RSR should be considered when change detection analyses are performed.

The view angle differences are common in the remote sensing observations and can introduce effect as high as 14% in defoliation (NED) when the sensors view from opposing view angles ($\pm$12) for TOA products, but reduces to a NED value of about 9% for an ideal atmospherically compensated product (ELM-ideal). Interestingly, the products compensated using typical ELM technique shows an effect as high as 40% in NED due to the differences in the path transmission and upwelled radiance. In the case of nominal view angles expected between the OLI and the MSI sensors, the effect in NED varies between 5% and 30% depending on the compensation techniques. The analysis in this study clearly suggests that the errors due to the view angle differences cannot be ignored and may affect the change detection studies significantly if the products are not compensated correctly for the atmosphere.

Although Vermote et al. (2016) has shown that the uncertainty of the Landsat's surface reflectance products are less than 5% (equivalent to about 10% in NED), many scientists still rely on the TOA reflectance products for their analysis. This study shows that the effect in the TOA products can be as high as 40% in NED, whereas, the atmospherically compensated products are likely to exhibit an effect equivalent to 1 - 7% in NED depending on the accuracy of the compensation techniques. Thus, using the surface reflectance products currently generated by USGS, the effect due to the atmospheric conditions can be reduced significantly ($\approx$ 10% NED) assuming any effect due to the sensor differences (RSR) are compensated between the scenes. In this research study, only the visibility parameter was chosen to introduce the differences in the atmospheric conditions, but the estimated effects are likely to be the same, if not worse, when other parameters such as aerosol differences, and cloud contaminations are also considered.

The effect of the solar angle differences is typically compensated in the TOA products based on the cosine effect of the solar zenith angles. However, any residual effects due to atmospheric attenuations (skylight) and the BRDF of the canopy can still affect the TOA and the surface reflectance products. Results from this study indicate that the solar zenith angle can introduce a change of about 5 degrees, when the datasets from the two sensors (OLI, MSI) are acquired

3 to 5 days apart. Assuming a 5 degree difference in SZN angles, the effects in the red and the NIR spectral bands can be as high as 30% in NED for the TOA radiance products but are significantly reduced to 10% in NED when compensated for the earth-sun distance and cosine effects. In the case of NDVI products, the effect varies between 4% to 10% depending on whether atmospheric compensation is applied or not. The residual 4% effect in NED is due to the BRDF effect of the canopy and cannot be compensated unless the BRDF of the surface is known. For datasets that were acquired more than 20 days apart, the corresponding change in solar zenith angle is about $10°$ which induces approximately twice the effect($\approx 8\%$ in NED due to BRDF).

The sensitivity analysis indicates that the visibility, solar zenith, RSR, and across-track can affect the sensor reaching radiance significantly. Compensation techniques such as SBAF and ELM methods can reduce the effect depending on the accuracy of the compensation techniques. The ideal and typical ELM methods serve as an indication of what can be expected from the existing and future atmospheric compensation techniques. Similarly, the SBAF techniques can be useful in eliminating the effects due to the differences in RSR, but their performance is dependent on the type of method (BRF vs TOA reflectance) and the type of data products that users are interested in (TOA reflectance vs surface reflectance). Further, comparison with the real data acquired over the Harvard forest by OLI sensor over a 23-day period indicates that more than half the observed changes are due to the effect of the factors.

6.1.8 Summary

The findings of the study bring to light the effect of the sensor and environmental factors for time series analysis, and how it can be reduced significantly with compensation. Although the purpose of the three methods (factorial experiments, regression analysis, and effect analysis) are different, they produced consistent results in evaluating the order of significance, thereby, increasing the confidence in the derived conclusions. The research work recommends that all the factors should be compensated to reduce the uncertainty in the change detection studies even when imaged by the same or similar sensors. Understanding that the effects due to the BRDF of the canopy may not be easily separated, as in the case with the across-track and SZN factors, future research should focus their resources on accurate compensation techniques for atmosphere and RSR differences.

Although the results were shown for the two sensors, the usefulness of this research is not limited to these two sensors, and can be easily extended for other types of sensors. The effects may not be the same, and can vary depending on the differences in the sensor characteristics and environmental conditions, but their order of significance and their relative contributions are highly likely to be the same. In the same light, the outcomes of this research, especially the order of significance, are not confined just to the deciduous forest, but are applicable to other biomes. The actual effects are dependent on the BRDF of the biomes of interest; hence the process established in this study can be replicated for other biomes in the future to estimate the corresponding effects.

6.2 Future work

6.2.1 DIRSIG Validation

In this research, the DIRSIG tool is used for measuring the BRDF of the canopy based on the geodensity radiative transfer algorithms, designed specifically for canopy interactions. Although, the validation of the DIRSIG model was performed with the RAMI III benchmark scenes, the geometric construct of RAMI III scenes are simple and abstract. The test scenarios provided in RAMI IV consists of "actual canopies", which are based on detailed inventories of existing forest and plantation sites. The results for the RAMI IV actual canopies were published recently (Widlowski et al., 2015), and could be used to validate the DIRSIG model.

6.2.2 Analysis for other spectral bands

Most of the analysis were performed only for the red and the NIR spectral bands as they are commonly used in the forest studies. However, for forest and non-forest applications, spectral bands such as SWIR are also used, which necessitates similar analysis for other spectral bands. Furthermore, the effect analysis in this research is primarily based on the NDVI of the forest, but many applications depend on more than one vegetation index. Therefore, the effect of various factors for different spectral bands and other vegetation indices can be explored in the future. The general process outlined in this research for sensitivity analysis can be used directly, except that the sensor reaching radiance would be simulated for different spectral bands.

6.2.3 Improvements to compensation

The factors' effect analysis in this research discussed the residual effect that remain after using SBAF and ELM-based compensation techniques. Although these algorithms are widely prevalent in the research community, a direct comparison and evaluation of the residual effect can be studied using existing and widely prevalent compensation techniques such as the methods used by the Landsat and MODIS communities. This can be accomplished by simulating the sensor reaching radiance for all the spectral bands and then use the published atmospheric compensation algorithms to compensate the simulated data. The comparison of compensated data with the known "ground truth" will indicate the performance and effectiveness of the compensation techniques.

6.2.4 Validation with real data

Comparison of real and simulated data is useful to validate the results derived from the simulated data. In this research, reflectance values from the OLI sensor were compared with the simulated BRDFs and the factors' effects were assessed with the real data, but the effects due to the RSR differences could not be validated for lack of data from the MSI sensor. This problem can be solved in the future when more Sentinel 2 data becomes available, especially after the launch of the Sentinel 2B satellite. This will also increase the possibility of acquiring coincident imaging between the OLI and MSI sensors which can be useful for validation purposes.

6.2.5 Extension to other biomes

The current study focused primarily on the effects of different factors for deciduous canopies, but it can be extended in the future to include other types of biomes such as agriculture, coastal land, etc. For these biomes, their corresponding 3D models may be modeled with appropriate modeling tools (e.g. OnyxTree, Xfrog, etc) and the DIRSIG tool can be used to characterize their spectral properties, similar to the process outlined in this research. The rationale behind the use of RossLi BRDF model for forest may not be valid for other biomes, but this can be validated by comparing the DIRSIG measurements to the RossLi BRDF model, as demonstrated in this research. A study similar to Koukal and Schneider (2010) can also be performed by comparing the simulated data with the existing BRDF models for different

types of biomes. Such comparative studies can help assess the performance of different models and provide appropriate recommendations for the use of biome-specific BRDF models to the scientific community.

Appendix A

Useful Equations

Spherical Coordinate transformation

$$
\begin{aligned}
X &= r\ \sin(\theta)\ \cos(\frac{\pi}{2} - \phi) \\
Y &= r\ \sin(\theta)\ \sin(\frac{\pi}{2} - \phi) \\
Z &= r\ \cos(\theta)
\end{aligned}
\tag{A.1}
$$

where,

θ, ϕ represents zenith and relative azimuth angles

Computation of Sun angles

$$
\begin{aligned}
Local\ Standard\ Time\ Meridian\ (LSTM) &= 15°\ \Delta T_{GMT} \\
B &= \frac{360}{365}(d - 81) \\
Equation\ of\ Time\ (EoT) &= 9.87\ \sin(2B) - 7.53\ \cos(B) - 1.5\sin(B) \\
Time\ Correction\ (TC) &= 4(longitude - LSTM) + EoT \\
Local\ Solar\ Time\ (LST) &= LT + \frac{TC}{60} \\
Hour\ Angle\ (HRA) &= 15°(LST - 12) \\
Declination\ angle\ (\delta) &= 23.45°\ \sin[B] \\
Elevation\ (\alpha) &= \sin^{-1}\ [\sin\delta \sin\psi + \cos\delta \cos\psi \cos(HRA)] \\
Zenith\ (\theta) &= 90°\ -\ \alpha \\
Azimuth\ (\phi) &= \cos^{-1}\ [\frac{\sin\delta \sin\psi - \cos\delta \sin\psi \cos(HRA)}{\cos\alpha}]
\end{aligned}
\tag{A.2}
$$

where

ΔT_{GMT} is the difference of the Local Time (LT) from Greenwich Mean Time (GMT) in hours

d is day of the year

ψ is Latitude

Position and orientation of the camera

$$
\begin{aligned}
X &= R\ sin(\theta_v)\ cos(90 - \phi_v) \\
Y &= R\ sin(\theta_v)\ sin(90 - \phi_v) \\
Z &= R\ cos(\theta_v) \\
r_x &= \theta_v \\
r_y &= 0 \\
r_z &= 180 - \phi_v
\end{aligned}
\tag{A.3}
$$

where

X, Y, Z and r_x, r_y, r_z are the position and orientation angles for camera

R is the altitude of the camera

θ_v, ϕ_v are zenith and azimuth view angles

Appendix B

ANOVA table

Table B.1: ANOVA table for single factor, fixed effects model

Source	SS	df	MS	F_{stat}	p-value
Between	$SS_{treatment}$	(a-1)	$\frac{SS_{treatment}}{(a-1)}$	$\frac{MS_{treatment}}{MS_E}$	
Within	SS_E	a(n-1)	$\frac{SS_E}{a(n-1)}$		
Total	SS_T	na-1			

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Expected Mean Square	F_0
A	SS_A	$a-1$	MS_A	$\sigma + \dfrac{bcn \sum \tau_i^2}{a-1}$	$F_0 = \dfrac{MS_A}{MS_E}$
B	SS_B	$b-1$	MS_B	$\sigma^2 + \dfrac{acn \sum \beta_j^2}{b-1}$	$F_0 = \dfrac{MS_B}{MS_E}$
C	SS_C	$c-1$	MS_C	$\sigma^2 + \dfrac{abn \sum \gamma_k^2}{c-1}$	$F_0 = \dfrac{MS_C}{MS_E}$
AB	SS_{AB}	$(a-1)(b-1)$	MS_{AB}	$\sigma^2 + \dfrac{cn \sum \sum (\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$F_0 = \dfrac{MS_{AB}}{MS_E}$
AC	SS_{AC}	$(a-1)(c-1)$	MS_{AC}	$\sigma^2 + \dfrac{bn \sum \sum (\tau\gamma)_{ik}^2}{(a-1)(c-1)}$	$F_0 = \dfrac{MS_{AC}}{MS_E}$
BC	SS_{BC}	$(b-1)(c-1)$	MS_{BC}	$\sigma^2 + \dfrac{an \sum \sum (\beta\gamma)_{jk}^2}{(b-1)(c-1)}$	$F_0 = \dfrac{MS_{BC}}{MS_E}$
ABC	SS_{ABC}	$(a-1)(b-1)(c-1)$	MS_{ABC}	$\sigma^2 + \dfrac{n \sum \sum \sum (\tau\beta\gamma)_{ijk}^2}{(a-1)(b-1)(c-1)}$	$F_0 = \dfrac{MS_{ABC}}{MS_E}$
Error	SS_E	$abc(n-1)$	MS_E	σ^2	
Total	SS_T	$abcn-1$			

Figure B.1: ANOVA table for three factor, fixed effects model (Montgomery, 2012).

Appendix C

DIRSIG simulation

Table C.1: Configuration for the frame camera used in DIRSIG simulations for BRDF measurement

Parameters	Values
Focal length	1000 mm
Spectral Bandwidth	1 nm
Detector sub-samples	10 , 10 (x,y)
Number of detectors	45 , 45 (x,y)
Pixel size	10 micron
Camera Altitude	100 KM

Appendix D

Fractional factorial design

Table D.1: Design Matrix for the 2^{7-1} design. The generators for the design is G = ABCDEF. The defining relation is given as I = ABCDEFG

Standard Order	A	B	C	D	E	F	G	Standard Order	A	B	C	D	E	F	G
1	-	-	-	-	-	-	+	33	-	-	-	-	-	+	-
2	+	-	-	-	-	-	-	34	+	-	-	-	-	+	+
3	-	+	-	-	-	-	-	35	-	+	-	-	-	+	+
4	+	+	-	-	-	-	+	36	+	+	-	-	-	+	-
5	-	-	+	-	-	-	-	37	-	-	+	-	-	+	+
6	+	-	+	-	-	-	+	38	+	-	+	-	-	+	-
7	-	+	+	-	-	-	+	39	-	+	+	-	-	+	-
8	+	+	+	-	-	-	-	40	+	+	+	-	-	+	+
9	-	-	-	+	-	-	-	41	-	-	-	+	-	+	+
10	+	-	-	+	-	-	+	42	+	-	-	+	-	+	-
11	-	+	-	+	-	-	+	43	-	+	-	+	-	+	-
12	+	+	-	+	-	-	-	44	+	+	-	+	-	+	+
13	-	-	+	+	-	-	+	45	-	-	+	+	-	+	-
14	+	-	+	+	-	-	-	46	+	-	+	+	-	+	+
15	-	+	+	+	-	-	-	47	-	+	+	+	-	+	+
16	+	+	+	+	-	-	+	48	+	+	+	+	-	+	-
17	-	-	-	-	+	-	-	49	-	-	-	-	+	+	+
18	+	-	-	-	+	-	+	50	+	-	-	-	+	+	-
19	-	+	-	-	+	-	+	51	-	+	-	-	+	+	-
20	+	+	-	-	+	-	-	52	+	+	-	-	+	+	+
21	-	-	+	-	+	-	+	53	-	-	+	-	+	+	-
22	+	-	+	-	+	-	-	54	+	-	+	-	+	+	+
23	-	+	+	-	+	-	-	55	-	+	+	-	+	+	+
24	+	+	+	-	+	-	+	56	+	+	+	-	+	+	-
25	-	-	-	+	+	-	+	57	-	-	-	+	+	+	-
26	+	-	-	+	+	-	-	58	+	-	-	+	+	+	+
27	-	+	-	+	+	-	-	59	-	+	-	+	+	+	+
28	+	+	-	+	+	-	+	60	+	+	-	+	+	+	-
29	-	-	+	+	+	-	-	61	-	-	+	+	+	+	+
30	+	-	+	+	+	-	+	62	+	-	+	+	+	+	-
31	-	+	+	+	+	-	+	63	-	+	+	+	+	+	-
32	+	+	+	+	+	-	-	64	+	+	+	+	+	+	+

Appendix E

Spectral BRDF : modeling and measurements

PROSPECT Inversion model
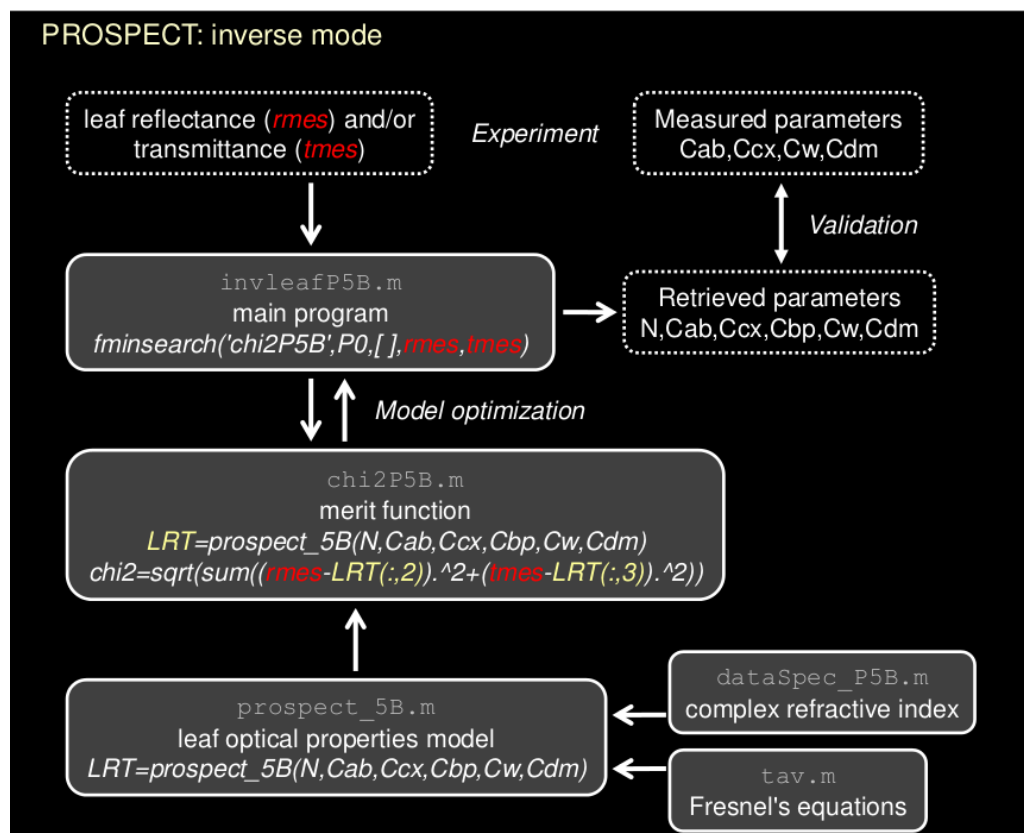
Figure E.1: A flowchart explaining the PROSPECT inversion model (Frederic Baret, 2015).

Spectral samples
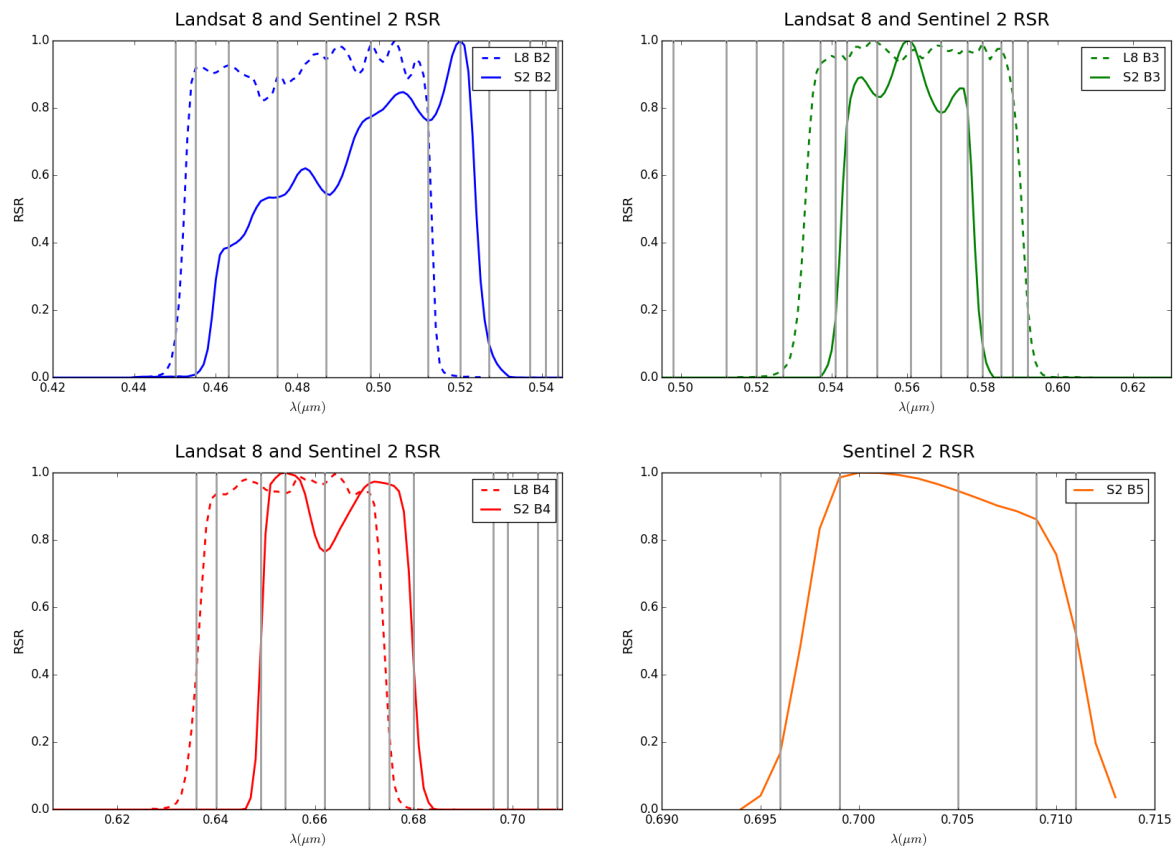
Figure E.2: Wavelength samples chosen for 4 different spectral bands of OLI and MSI sensor. The gray line indicates the sampled wavelength.
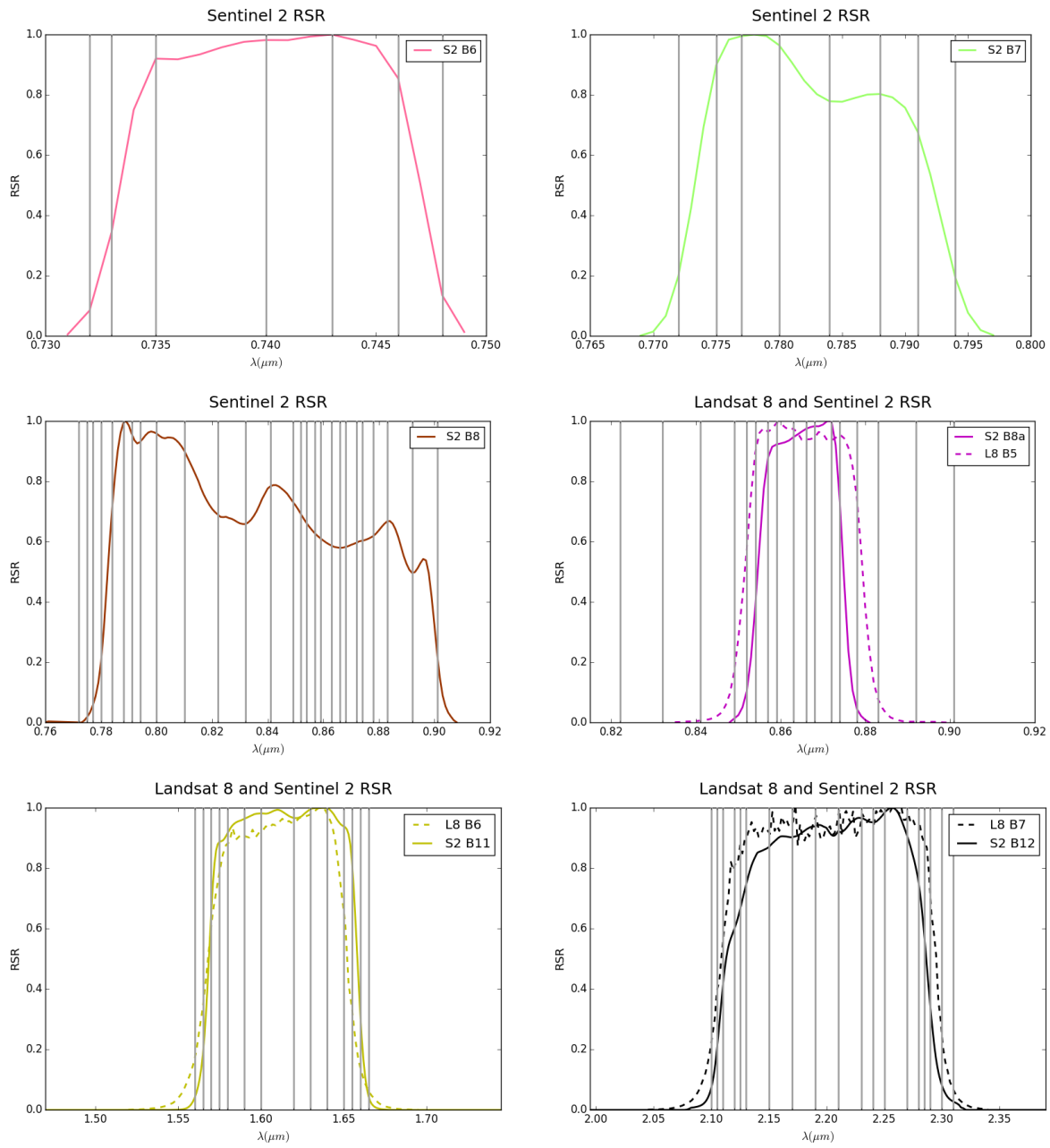
Figure E.3: Wavelength samples chosen for 6 different spectral bands of OLI and MSI sensor. The gray line indicates the sampled wavelength.

3D ROI overlay

(a) VZN : 2 , VAZ : 180

(b) VZN : 2 , VAZ : 0

(c) VZN : 2 , VAZ : 104

(d) VZN : 7 , VAZ : 98

(e) VZN : 13 , VAZ : 86

(f) VZN : 17 , VAZ : 90
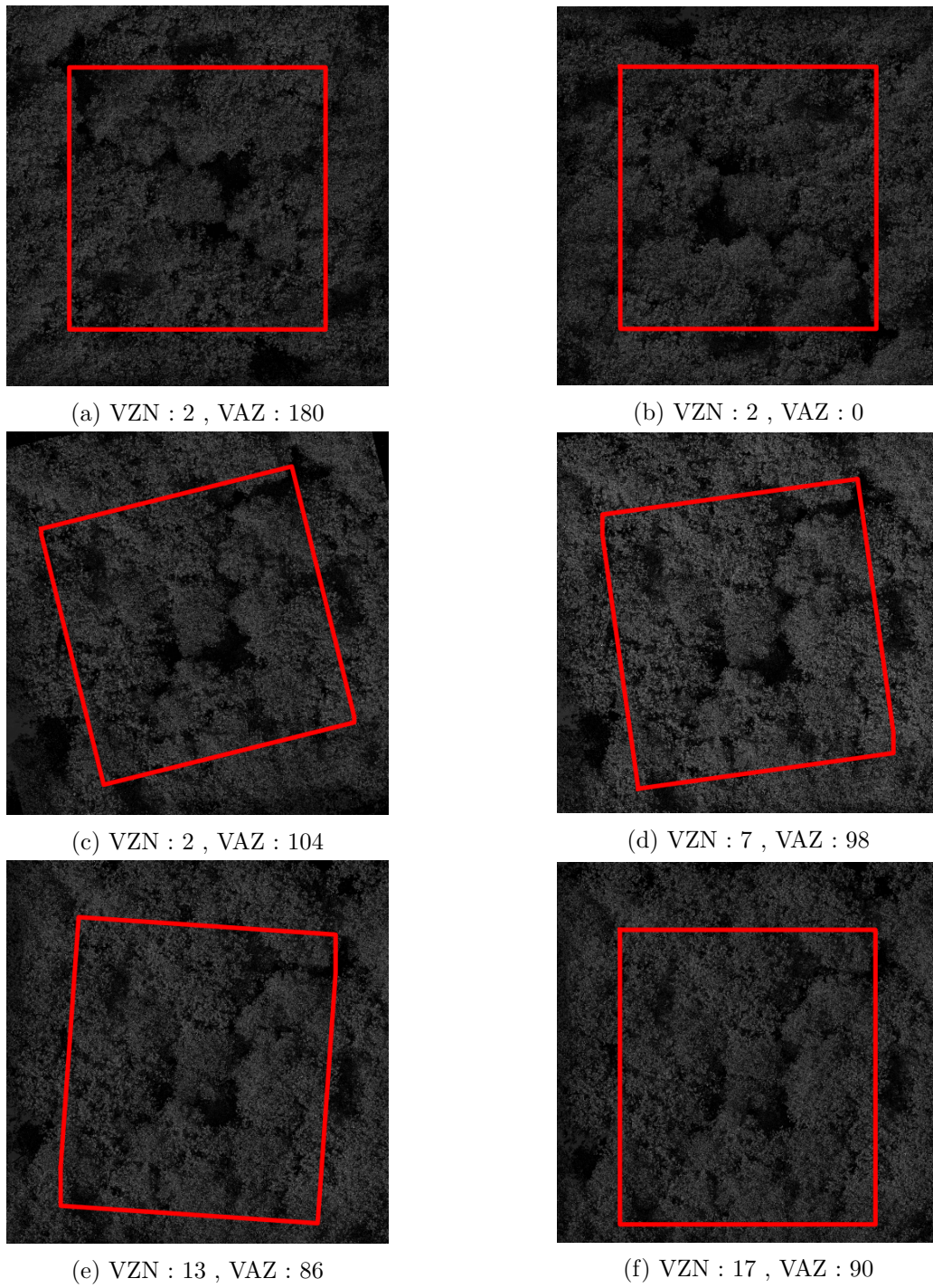
Figure E.4: The 3D ROI overlay of six different view angles when illuminated by the sun from 30° zenith angle and 130° azimuth angle ($\lambda = 0.866\mu m$).

(a) VZN : 15 , VAZ : 98

(b) VZN : 3 , VAZ : 284

(c) VZN : 8 , VAZ : 279

(d) VZN : 12 , VAZ : 263

(e) VZN : 17 , VAZ : 270

(f) VZN : 15 , VAZ : 273
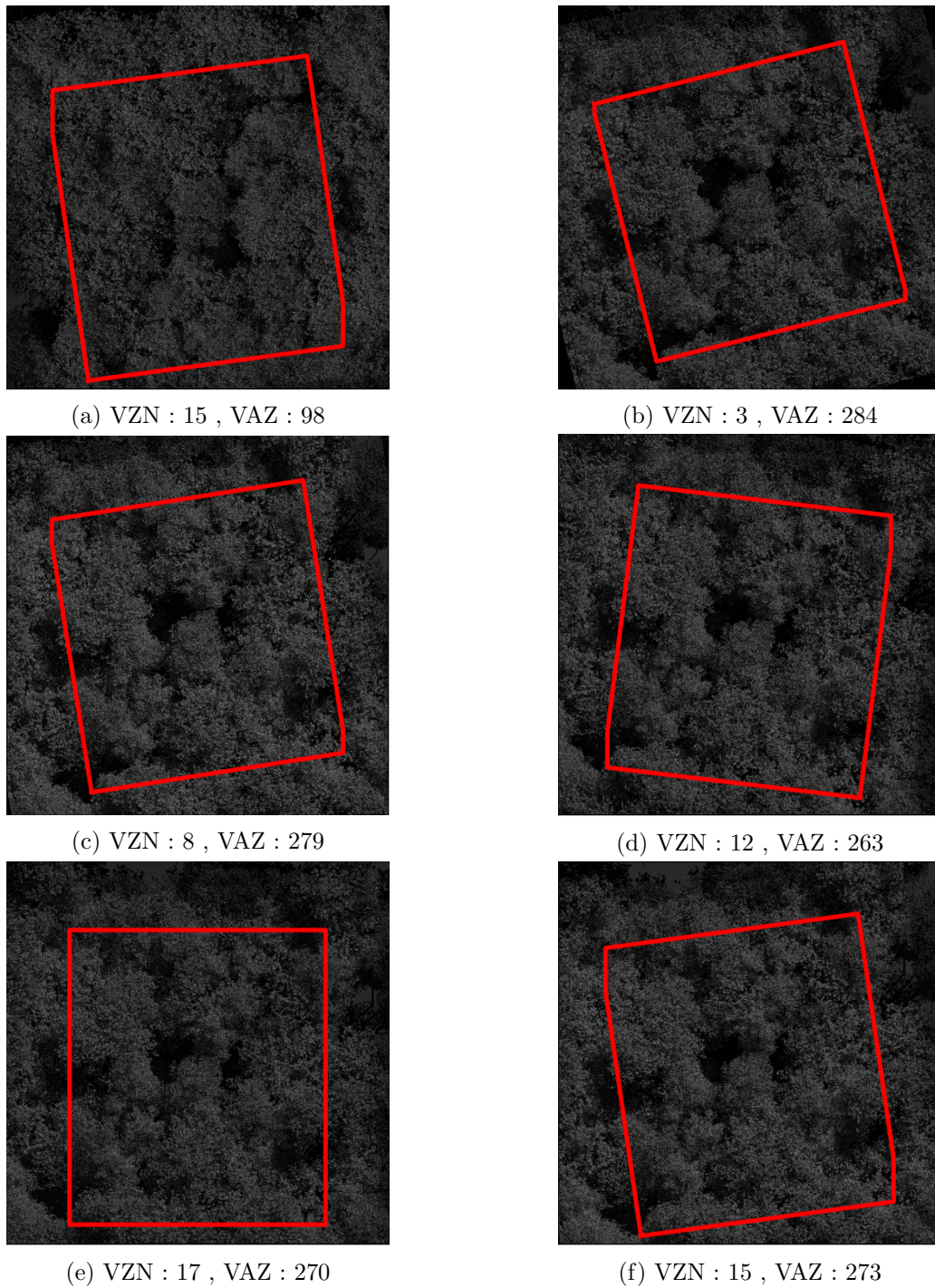
Figure E.5: The 3D ROI overlay of six different view angles when illuminated by the sun from 30° zenith angle and 130° azimuth angle ($\lambda = 0.866\mu m$).

Appendix F

Factorial screening experiment results

2kp Table results

```
Analysis of Variance

Source                   DF   Adj SS   Adj MS    F-Value   P-Value
Model                    18   120527     6696    1260.87     0.000
  Linear                  7   113159    16166    3044.02     0.000
    RSR                   1        1        1       0.16      0.691
    Atrack                1        0        0       0.03      0.873
    Xtrack                1       18       18       3.45      0.070
    Visibility            1     6268     6268    1180.30      0.000
    Terrain               1     1028     1028     193.53      0.000
    SAZ                   1        5        5       0.91      0.344
    SZN                   1   105839   105839   19929.73      0.000
  2-Way Interactions     11     7369      670     126.14      0.000
    RSR*Atrack            1        6        6       1.20      0.279
    RSR*Terrain           1        0        0       0.00      0.971
    Atrack*Terrain        1        0        0       0.09      0.768
    Xtrack*Visibility     1        8        8       1.53      0.223
    Xtrack*Terrain        1       22       22       4.10      0.049
    Xtrack*SAZ            1       18       18       3.32      0.075
    Xtrack*SZN            1       78       78      14.62      0.000
    Visibility*Terrain    1      157      157      29.57      0.000
    Visibility*SZN        1     7073     7073    1331.87      0.000
    Terrain*SAZ           1        1        1       0.12      0.730
    SAZ*SZN               1        6        6       1.12      0.296
Error                    45      239        5
Total                    63   120766
```

Figure F.1: ANOVA table for the center forest factorial experiments for the NIR spectral band. The highlighted(yellow) factors are significant while the orange highlight indicates a factor that is marginally insignificant.

```
Analysis of Variance

Source                DF   Adj SS   Adj MS    F-Value   P-Value
Model                 18   120527     6696    1260.87     0.000
  Linear               7   113159    16166    3044.02     0.000
    RSR                1        1        1       0.16      0.691
    Atrack             1        0        0       0.03      0.873
    Xtrack             1       18       18       3.45      0.070
    Visibility         1     6268     6268    1180.30      0.000
    Terrain            1     1028     1028     193.53      0.000
    SAZ                1        5        5       0.91      0.344
    SZN                1   105839   105839   19929.73      0.000
  2-Way Interactions  11     7369      670     126.14      0.000
    RSR*Atrack         1        6        6       1.20      0.279
    RSR*Terrain        1        0        0       0.00      0.971
    Atrack*Terrain     1        0        0       0.09      0.768
    Xtrack*Visibility  1        8        8       1.53      0.223
    Xtrack*Terrain     1       22       22       4.10      0.049
    Xtrack*SAZ         1       18       18       3.32      0.075
    Xtrack*SZN         1       78       78      14.62      0.000
    Visibility*Terrain 1      157      157      29.57      0.000
    Visibility*SZN     1     7073     7073    1331.87      0.000
    Terrain*SAZ        1        1        1       0.12      0.730
    SAZ*SZN            1        6        6       1.12      0.296
Error                 45      239        5
Total                 63   120766
```

Figure F.2: ANOVA table for the random forest factorial experiment for the NIR spectral band. The highlighted(yellow) factors are significant.
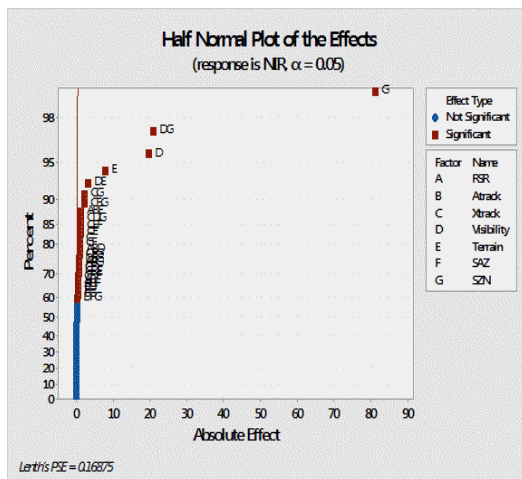
(a) Center forest BRDF
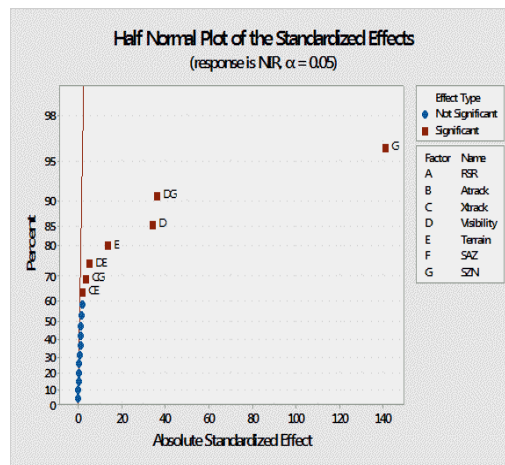
(b) Random BRDF

Figure F.3: Half-Normal probability plot before and after model reduction for center forest factorial experiment. The significant effects can be seen identified in red dots.

(a) Center forest BRDF

(b) Random BRDF
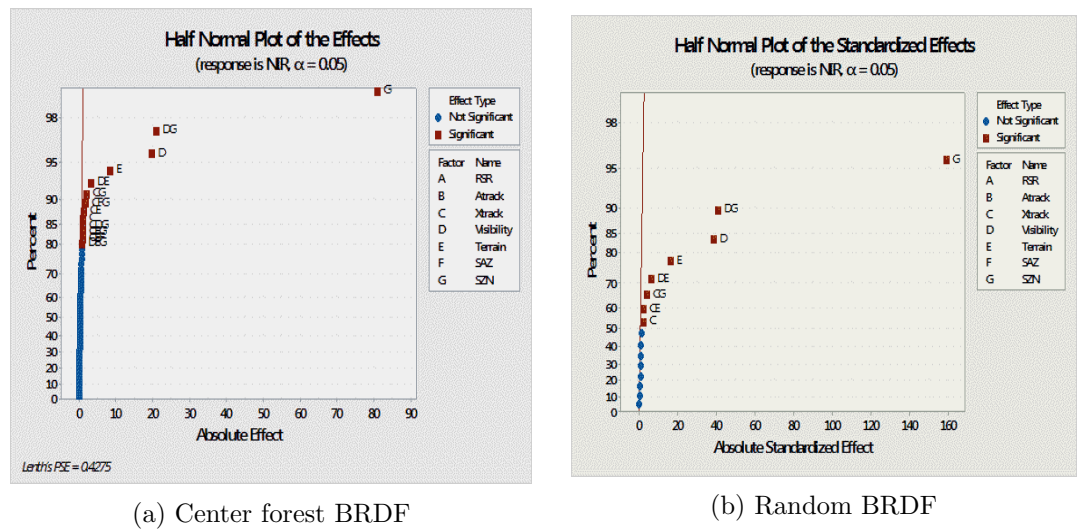
Figure F.4: Half-Normal probability plot before and after model reduction for random forest factorial experiment. The significant effects can be seen identified in red dots.
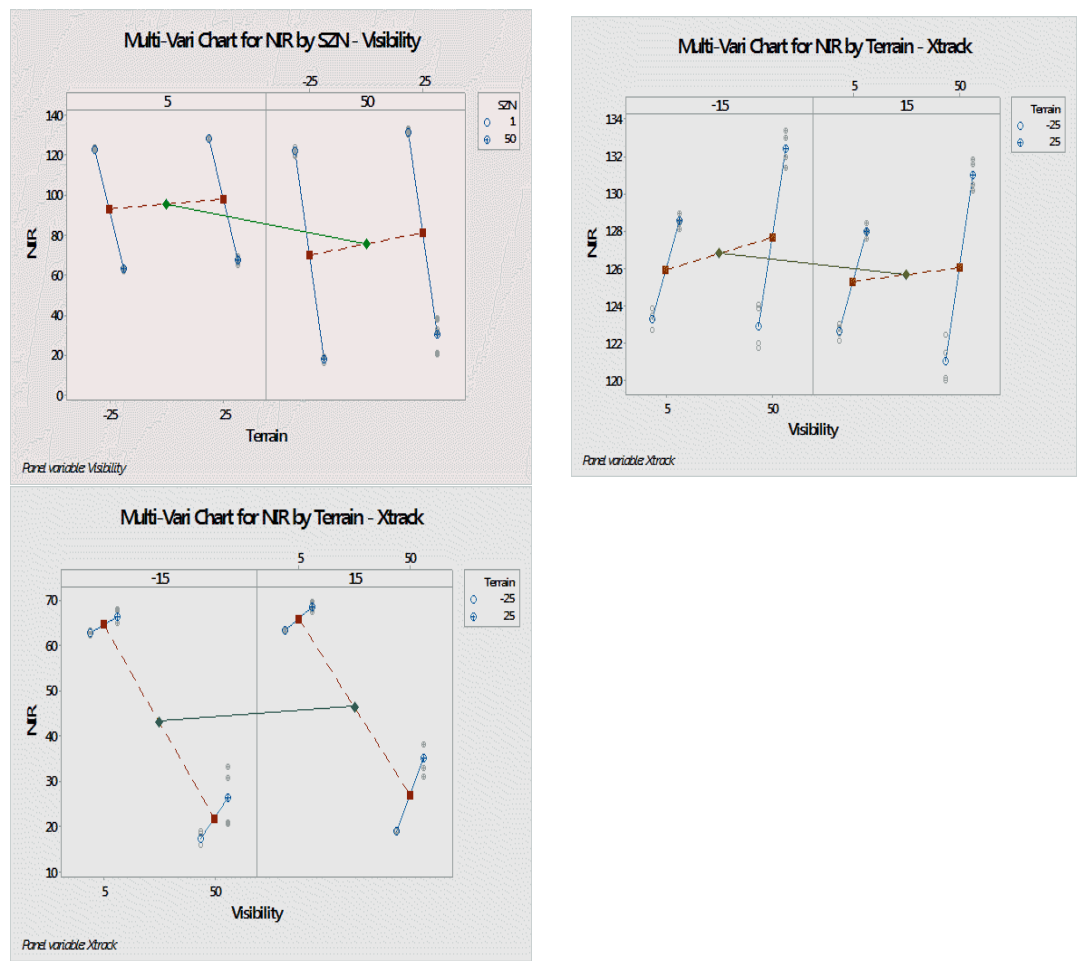
Figure F.5: Interaction plots for center forest
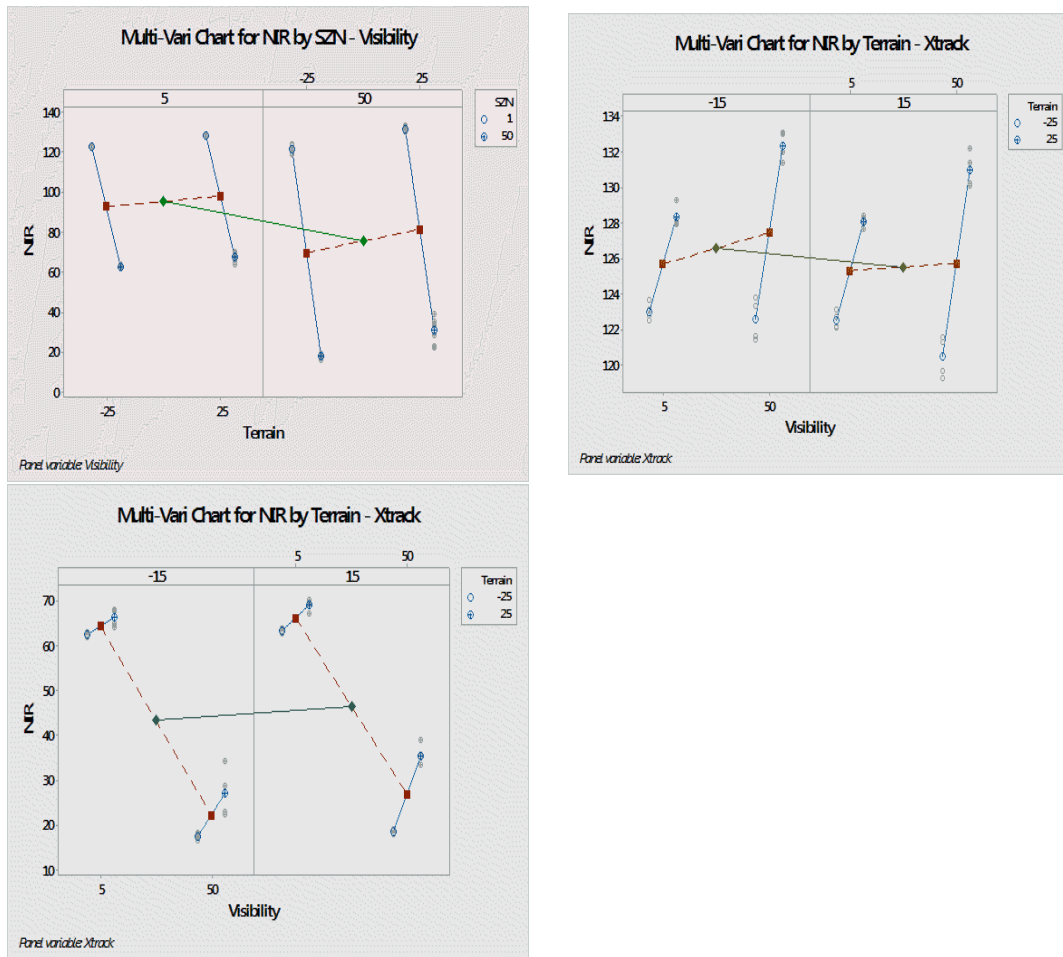
Figure F.6: Interaction plots for random forest

Table F.1: The effects in the blue spectral band using fractional factorial analysis for different forest sections. The "*" indicates the effects of the factor that are insignificant in relation to other factors.

Factors	LR	LL	UR	UL	Center	Random
Visibility	-132.16	-132.07	-132.34	-132.33	-132.22	-132.34
Solar Zenith	-72.41	-72.42	-72.55	-72.56	-72.46	-72.48
RSR	-6.25	-6.23	-6.25	-6.23	-6.23	-6.24
Across Track	-1.2	-1.19	-1.19	-1.12	-1.20	-1.20
Terrain slope	0.69	0.67	0.77	0.70	0.73	0.73
Solar Azimuth	-0.22	-0.22	-0.21	-0.21	-0.21	-0.21*
Along Track	0.11*	0.11*	0.11*	0.11*	0.11*	0.10*

Table F.2: The effects in the green spectral band using fractional factorial analysis for different forest sections. The "*" indicates the effects of the factor that are insignificant in relation to other factors.

Factors	LR	LL	UR	UL	Center	Random
Visibility	-119.43	-119.06	-119.57	-119.33	-119.28	-119.30
Solar Zenith	-61.71	-61.72	-61.77	-61.76	-61.67	-61.67
Terrain slope	1.45	1.58	1.68	1.68	1.67	1.65
RSR	1.26	1.29	1.27	1.29	1.28	1.29
Across Track	-0.58	-0.55	-0.56	-0.56	-0.58	-0.58
Solar Azimuth	-0.07*	-0.07*	-0.07*	-0.07*	-0.07*	-0.08*
Along Track	0.02*	0.03*	0.02*	0.04*	0.04*	0.03*

Table F.3: The effects in the SWIR-1 spectral band using fractional factorial analysis for different forest sections. The "*" indicates the effects of the factor that are insignificant in relation to other factors.

Factors	LR	LL	UR	UL	Center	Random
Solar Zenith	-16.77	-16.37	-15.87	-14.94	-15.92	-15.91
Visibility	-2.94	-2.71	-2.77	-2.50	-2.69	-2.70
Terrain slope	1.15	2.01	1.52	2.18	1.82	1.88
Across Track	0.13	0.30	0.29	0.20	0.20	0.17
Solar Azimuth	-0.09	0.09	0.13	0.11	0.12	0.10
RSR	-0.04*	-0.04*	-0.04*	-0.04*	-0.04*	-0.07*
Along Track	-0.04*	0.04*	0.02*	-0.03*	0.02*	-0.01*

Table F.4: The effects in the SWIR-2 spectral band using fractional factorial analysis for different forest sections. The "*" indicates the effects of the factor that are insignificant in relation to other factors.

Factors	LR	LL	UR	UL	Center	Random
Solar Zenith	-3.02	-2.96	-2.88	-2.79	-2.88	-2.89
Visibility	-0.88	-0.87	-0.88	-0.86	-0.87	-0.87
Terrain slope	0.24	0.36	0.30	0.38	0.34	0.35
Across Track	0.05	0.05	0.04	0.03	0.027	0.03
RSR	0.03	0.03	0.03	0.03	0.03	0.03
Solar Azimuth	0.01	0.02	0.02	0.02	0.02	0.02
Along Track	-0.007*	-0.007*	-0.005*	-0.004*	0.002*	-0.0003*

Appendix G

Sensitivity study of factors

Comparison of real and simulated data

Table G.1: Relative variations measured from the real and the simulated data

Product type	Relative change between the two dates for Harvard forest (%)		
	Real data	simulated data Vis(1) = 20km Vis(2) = 15km	Vis(1) = 20km Vis(2) = 20km
TOA RED	10.2	17.7	9.6
TOA NIR	14.6	6.8	9
TOA REFL RED	5.9	14	5.6
TOA REFL NIR	11	2.6	4.9
ELM-typical RED	4.6	9.2	10.4
ELM-typical NIR	11	5.1	5.8
NDVI (TOA RAD)	1.8	6	0.3
NDVI (TOA REFL)	1.2	3.9	0.2
NDVI (ELM-typical)	0.7	0.5	0.5

Across-Track effects
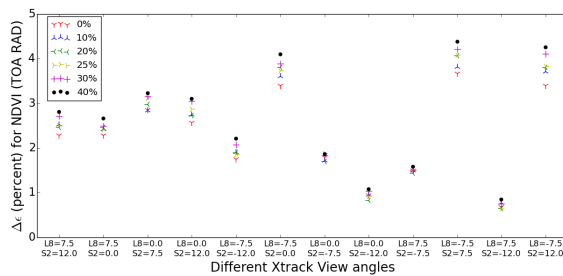
(a) NDVI using TOA RAD
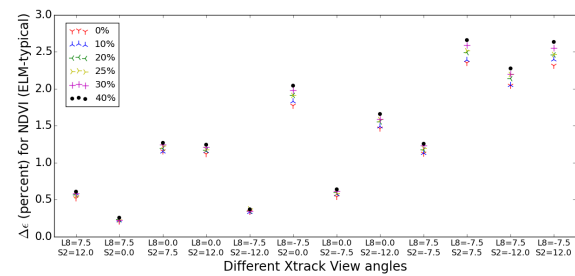
(b) NDVI using ELM-typical

Figure G.1: Across-track effects for NDVI products

(a) NDVI using TOA RAD

(b) NDVI using TOA REFL

(c) NDVI using ELM-typical

(d) NDVI using BRF

Figure G.2: Across-track effects for NDVI products after compensation with SBAF-BRF. The trend for all the NDVI products are consistent for different view angle combinations. The results shown for SBAF-BRF compensation in this Figure is very similar to that observed with SBAF-TOA REFL compensated products (not shown).

Appendix H

Regression Analysis

Table H.1: Orthogonal polynomial contrasts for SZN

Levels	Bias	1st order	2nd order	3rd order	4th order
1	0.4472136	0.4812126	0.4525018	0.4261546	0.4266952
2	0.4472136	0.382624	0.1422977	-0.325178	-0.726368
3	0.4472136	0.0922641	-0.459184	-0.580347	0.4937962
4	0.4472136	-0.198357	-0.594566	0.600899	-0.21463
5	0.4472136	-0.757743	0.4589501	-0.121529	0.0205065

Table H.2: Orthogonal polynomial contrasts for X-Track

Levels	Bias	1st order	2nd order	3rd order	4th order
1	0.4472136	-0.632456	0.5345225	-0.316228	0.1195229
2	0.4472136	-0.316228	-0.267261	0.6324555	-0.478091
3	0.4472136	-1.01E-17	-0.534522	5.644E-17	0.7171372
4	0.4472136	0.3162278	-0.267261	-0.632456	-0.478091
5	0.4472136	0.6324555	0.5345225	0.3162278	0.1195229

Table H.3: Orthogonal polynomial contrasts for visibility

Levels	Bias	1st order	2nd order	3rd order	4th order
1	0.4472136	-0.591556	0.541872	-0.36214	0.1590325
2	0.4472136	-0.325292	-0.163388	0.5918115	-0.563248
3	0.4472136	-0.043036	-0.50662	0.1780592	0.7139879
4	0.4472136	0.2778214	-0.391347	-0.653191	-0.37816
5	0.4472136	0.682063	0.519483	0.2454601	0.0683873

Table H.4: Orthogonal polynomial contrasts for terrain slope

Levels	Bias	1st order	2nd order	3rd order	4th order
1	0.4472136	-0.606338	0.5164761	-0.363805	0.1823525
2	0.4472136	-0.363805	-0.2084	0.6063378	-0.506527
3	0.4472136	3.857E-18	-0.616152	2.581E-17	0.6483488
4	0.4472136	0.3638055	-0.2084	-0.606338	-0.506527
5	0.4472136	0.6063378	0.5164761	0.3638055	0.1823525
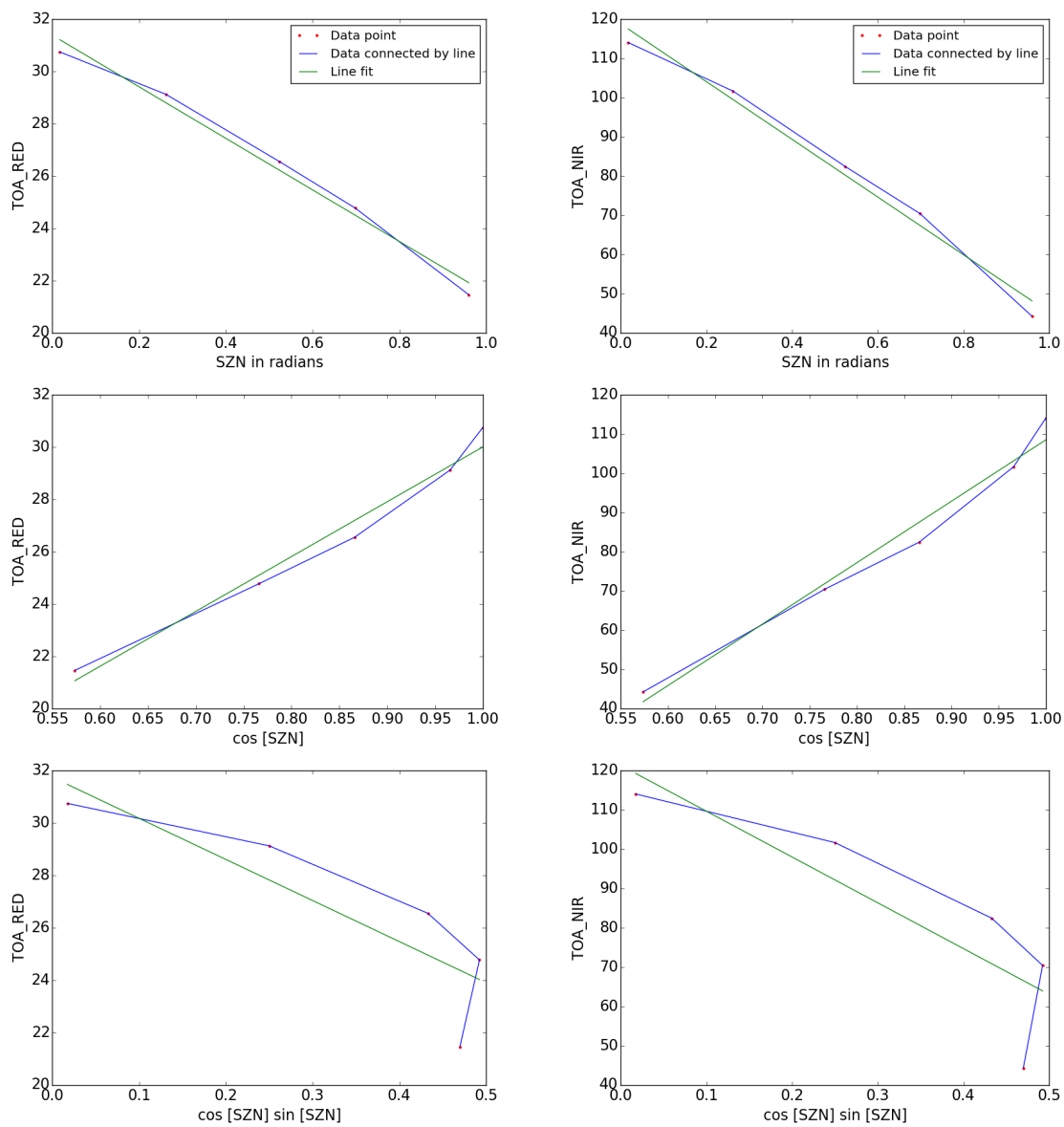
Figure H.1: Various functional variables for SZN

Figure H.2: Various functional variables for terrain slope

Figure H.3: Various functional variables for across-track

Figure H.4: Various functional variables for visibility

Appendix I

DIRSIG simulated data

Table I.1: Simulated data for different atmospheric and sensor conditions when the forest is undefoliated

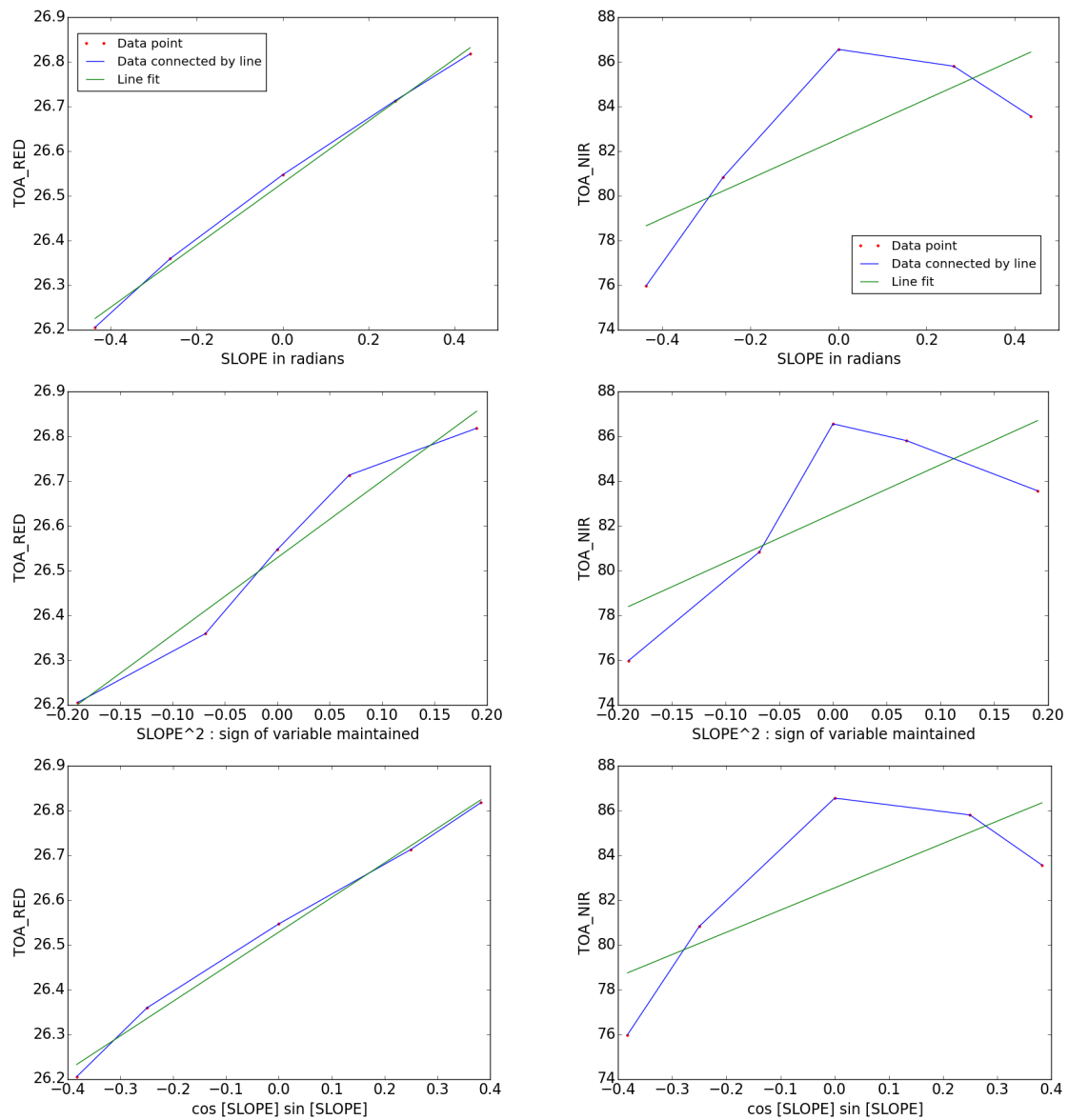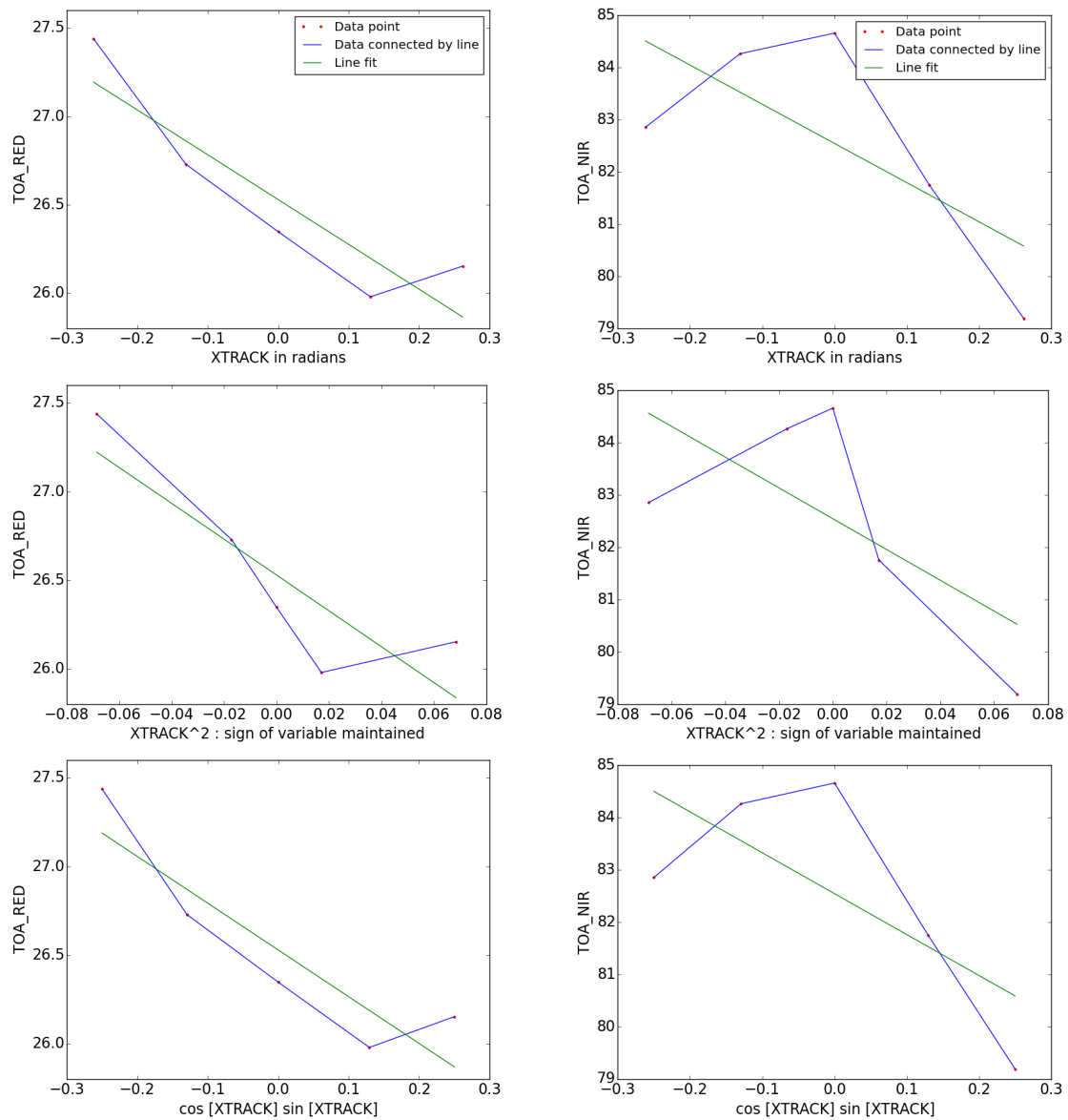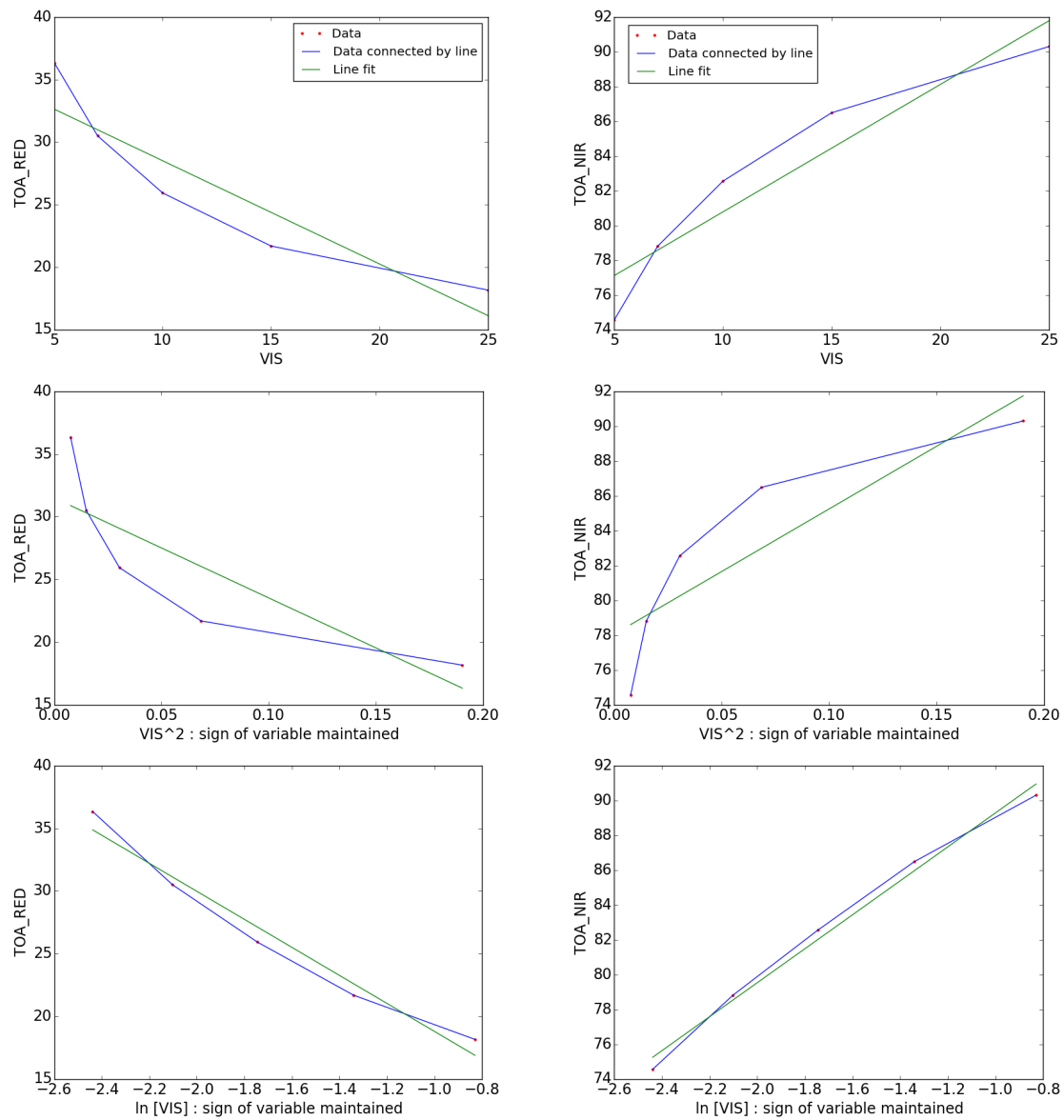| RSR	XT	VIS	szn	saz	L_{red}	L_{NIR}	RSR	XT	VIS	szn	saz	L_{red}	L_{NIR}	RSR	XT	VIS	szn	saz	L_{red}	L_{NIR}
1	-15	10km	35	145	26.45	84.3	1	-7.5	20km	35	145	19.69	89.32	1	7.5	10km	35	145	24.78	82.42
1	-15	10km	30	137	27.6	90.16	1	-7.5	20km	30	137	20.71	95.65	1	7.5	10km	30	137	25.57	87.87
1	-15	10km	33	157	26.48	85.84	1	-7.5	20km	33	157	19.88	91.42	1	7.5	10km	33	157	25.26	84.6
1	-15	10km	25	150	28.29	97	1	-7.5	20km	25	150	21.58	103.31	1	7.5	10km	25	150	26.74	95.23
1	-15	10km	27	135	28.28	94.51	1	-7.5	20km	27	135	21.37	100.25	1	7.5	10km	27	135	26.12	91.46
1	-15	10km	35	165	25.88	82.78	1	-7.5	20km	35	165	19.37	88.29	1	7.5	10km	35	165	24.98	81.86
1	-15	10km	20	150	29.16	103.01	1	-7.5	20km	20	150	22.55	110.23	1	7.5	10km	20	150	27.67	100.23
1	-15	10km	23	132	29.11	100.51	1	-7.5	20km	23	132	22.25	106.75	1	7.5	10km	23	132	26.85	96.02
1	-15	10km	40	150	25.41	78.57	1	-7.5	20km	40	150	18.7	82.67	1	7.5	10km	40	150	24	76.69
1	-15	10km	38	153	25.8	82.25	1	-7.5	20km	38	153	19.1	86.33	1	7.5	10km	38	153	24.45	80.55
1	-15	15km	35	145	22.32	87.86	1	-7.5	25km	35	145	18.37	90.62	1	7.5	15km	35	145	20.81	85.84
1	-15	15km	30	137	23.44	94.2	1	-7.5	25km	30	137	19.37	97.14	1	7.5	15km	30	137	21.62	91.73
1	-15	15km	33	157	22.38	89.59	1	-7.5	25km	33	157	18.57	92.82	1	7.5	15km	33	157	21.29	88.29
1	-15	15km	25	150	24.13	101.03	1	-7.5	25km	25	150	20.23	104.73	1	7.5	15km	25	150	22.76	99.05
1	-15	15km	27	135	24.11	98.85	1	-7.5	25km	27	135	20.02	101.84	1	7.5	15km	27	135	22.17	95.44
1	-15	15km	35	165	21.79	86.45	1	-7.5	25km	35	165	18.07	89.69	1	7.5	15km	35	165	21.01	85.48
1	-15	15km	20	150	25.02	107.52	1	-7.5	25km	20	150	21.2	111.86	1	7.5	15km	20	150	23.66	104.24
1	-15	15km	23	132	24.97	105.29	1	-7.5	25km	23	132	20.91	108.47	1	7.5	15km	23	132	22.88	100.1
1	-15	15km	40	150	21.3	81.62	1	-7.5	25km	40	150	17.38	83.72	1	7.5	15km	40	150	20	79.55
1	-15	15km	38	153	21.68	85.24	1	-7.5	25km	38	153	17.77	87.31	1	7.5	15km	38	153	20.46	83.41
1	-15	20km	35	145	20.17	89.82	1	0	10km	35	145	25.07	83.07	1	7.5	20km	35	145	18.76	87.73
1	-15	20km	30	137	21.26	96.45	1	0	10km	30	137	26	88.6	1	7.5	20km	30	137	19.58	93.9
1	-15	20km	33	157	20.24	91.69	1	0	10km	33	157	25.43	85.11	1	7.5	20km	33	157	19.24	90.35
1	-15	20km	25	150	21.95	103.19	1	0	10km	25	150	27.09	95.87	1	7.5	20km	25	150	20.69	101.1
1	-15	20km	27	135	21.93	101.28	1	0	10km	27	135	26.6	92.19	1	7.5	20km	27	135	20.12	97.67
1	-15	20km	35	165	19.67	88.51	1	0	10km	35	165	25.01	82.25	1	7.5	20km	35	165	18.96	87.53
1	-15	20km	20	150	22.84	109.97	1	0	10km	20	150	28.13	101.95	1	7.5	20km	20	150	21.55	106.39
1	-15	20km	23	132	22.8	107.96	1	0	10km	23	132	27.44	97.32	1	7.5	20km	23	132	20.82	102.35
1	-15	20km	40	150	19.16	83.22	1	0	10km	40	150	24.17	77.26	1	7.5	20km	40	150	17.92	81.06
1	-15	20km	38	153	19.53	86.76	1	0	10km	38	153	24.63	81.11	1	7.5	20km	38	153	18.38	84.86
1	-15	25km	35	145	18.81	91.13	1	0	15km	35	145	21.13	86.57	1	7.5	25km	35	145	17.47	89.01
1	-15	25km	30	137	19.88	97.98	1	0	15km	30	137	22.06	92.53	1	7.5	25km	30	137	18.3	95.37
1	-15	25km	33	157	18.89	93.1	1	0	15km	33	157	21.49	88.84	1	7.5	25km	33	157	17.95	91.76
1	-15	25km	25	150	20.56	104.64	1	0	15km	25	150	23.11	99.78	1	7.5	25km	25	150	19.38	102.48
1	-15	25km	27	135	20.55	102.91	1	0	15km	27	135	22.64	96.24	1	7.5	25km	27	135	18.83	99.17
1	-15	25km	35	165	18.34	89.91	1	0	15km	35	165	21.08	85.9	1	7.5	25km	35	165	17.66	88.93
1	-15	25km	20	150	21.46	111.61	1	0	15km	20	150	24.14	106.31	1	7.5	25km	20	150	20.22	107.83
1	-15	25km	23	132	21.43	109.75	1	0	15km	23	132	23.46	101.57	1	7.5	25km	23	132	19.52	103.86
1	-15	25km	40	150	17.8	84.3	1	0	15km	40	150	20.22	80.21	1	7.5	25km	40	150	16.61	82.08
1	-15	25km	38	153	18.17	87.76	1	0	15km	38	153	20.68	84.03	1	7.5	25km	38	153	17.07	85.84
1	-7.5	10km	35	145	25.8	83.89	1	0	20km	35	145	19.09	88.5	1	15	10km	35	145	24.86	81.88
1	-7.5	10km	30	137	26.88	89.5	1	0	20km	30	137	20.01	94.74	1	15	10km	30	137	25.56	87.24
1	-7.5	10km	33	157	25.97	85.62	1	0	20km	33	157	19.45	90.93	1	15	10km	33	157	25.38	84.23
1	-7.5	10km	25	150	27.77	97.22	1	0	20km	25	150	21.03	101.89	1	15	10km	25	150	26.73	94.44
1	-7.5	10km	27	135	27.54	93.71	1	0	20km	27	135	20.58	98.5	1	15	10km	27	135	26.07	90.94
1	-7.5	10km	35	165	25.44	82.6	1	0	20km	35	165	19.05	87.97	1	15	10km	35	165	25.17	81.68
1	-7.5	10km	20	150	28.72	103.3	1	0	20km	20	150	22.04	108.68	1	15	10km	20	150	27.6	99.28
1	-7.5	10km	23	132	28.42	99.62	1	0	20km	23	132	21.38	103.93	1	15	10km	23	132	26.75	95.35
1	-7.5	10km	40	150	24.81	78.14	1	0	20km	40	150	18.17	81.77	1	15	10km	40	150	24.16	76.23
1	-7.5	10km	38	153	25.21	81.95	1	0	20km	38	153	18.62	85.52	1	15	10km	38	153	24.58	80.07
1	-7.5	15km	35	145	21.78	87.4	1	0	25km	35	145	17.8	89.81	1	15	15km	35	145	20.81	85.23
1	-7.5	15km	30	137	22.83	93.45	1	0	25km	30	137	18.72	96.23	1	15	15km	30	137	21.55	91.03
1	-7.5	15km	33	157	21.96	89.34	1	0	25km	33	157	18.17	92.34	1	15	15km	33	157	21.34	87.86
1	-7.5	15km	25	150	23.71	101.19	1	0	25km	25	150	19.71	103.3	1	15	15km	25	150	22.7	98.27
1	-7.5	15km	27	135	23.49	97.91	1	0	25km	27	135	19.28	100.02	1	15	15km	27	135	22.06	94.83
1	-7.5	15km	35	165	21.45	86.24	1	0	25km	35	165	17.77	89.37	1	15	15km	35	165	21.12	85.25
1	-7.5	15km	20	150	24.67	107.8	1	0	25km	20	150	20.71	110.27	1	15	15km	20	150	23.54	103.28
1	-7.5	15km	23	132	24.37	104.21	1	0	25km	23	132	20.06	105.51	1	15	15km	23	132	22.74	99.38
1	-7.5	15km	40	150	20.79	81.11	1	0	25km	40	150	16.87	82.82	1	15	15km	40	150	20.06	79.03
1	-7.5	15km	38	153	21.19	84.86	1	0	25km	38	153	17.32	86.51	1	15	15km	38	153	20.5	82.88

RSR	XT	VIS	szn	saz	L_{red}	L_{NIR}	RSR	XT	VIS	szn	saz	L_{red}	L_{NIR}	RSR	XT	VIS	szn	saz	L_{red}	L_{NIR}
1	15	20km	35	145	18.72	87.09	2	-7.5	15km	35	165	20.42	86.42	2	7.5	15km	35	145	19.81	86.02
1	15	20km	30	137	19.47	93.16	2	-7.5	15km	20	150	23.47	108.01	2	7.5	15km	30	137	20.57	91.92
1	15	20km	33	157	19.25	89.9	2	-7.5	15km	23	132	23.19	104.41	2	7.5	15km	33	157	20.27	88.47
1	15	20km	25	150	20.6	100.34	2	-7.5	15km	40	150	19.81	81.29	2	7.5	15km	25	150	21.65	99.25
1	15	20km	27	135	19.99	97.01	2	-7.5	15km	38	153	20.18	85.05	2	7.5	15km	27	135	21.08	95.64
1	15	20km	35	165	19.03	87.27	2	-7.5	20km	35	145	18.67	89.51	2	7.5	15km	35	165	20	85.66
1	15	20km	20	150	21.43	105.44	2	-7.5	20km	30	137	19.63	95.84	2	7.5	15km	20	150	22.49	104.45
1	15	20km	23	132	20.66	101.61	2	-7.5	20km	33	157	18.85	91.6	2	7.5	15km	23	132	21.76	100.3
1	15	20km	40	150	17.94	80.51	2	-7.5	20km	25	150	20.45	103.51	2	7.5	15km	40	150	19.05	79.72
1	15	20km	38	153	18.38	84.32	2	-7.5	20km	27	135	20.25	100.45	2	7.5	15km	38	153	19.48	83.59
1	15	25km	35	145	17.4	88.35	2	-7.5	20km	35	165	18.37	88.47	2	7.5	20km	35	145	17.79	87.91
1	15	25km	30	137	18.17	94.61	2	-7.5	20km	20	150	21.36	110.44	2	7.5	20km	30	137	18.55	94.08
1	15	25km	33	157	17.93	91.29	2	-7.5	20km	23	132	21.08	106.96	2	7.5	20km	33	157	18.24	90.54
1	15	25km	25	150	19.28	101.73	2	-7.5	20km	40	150	17.74	82.85	2	7.5	20km	25	150	19.59	101.3
1	15	25km	27	135	18.68	98.48	2	-7.5	20km	38	153	18.11	86.51	2	7.5	20km	27	135	19.06	97.86
1	15	25km	35	165	17.71	88.64	2	-7.5	25km	35	145	17.36	90.8	2	7.5	20km	35	165	17.97	87.71
1	15	25km	20	150	20.09	106.88	2	-7.5	25km	30	137	18.31	97.33	2	7.5	20km	20	150	20.41	106.6
1	15	25km	23	132	19.35	103.1	2	-7.5	25km	33	157	17.55	93	2	7.5	20km	23	132	19.71	102.55
1	15	25km	40	150	16.59	81.51	2	-7.5	25km	25	150	19.11	104.94	2	7.5	20km	40	150	17	81.23
1	15	25km	38	153	17.04	85.28	2	-7.5	25km	27	135	18.92	102.03	2	7.5	20km	38	153	17.43	85.04
2	-15	10km	35	145	25.34	84.49	2	-7.5	25km	35	165	17.08	89.87	2	7.5	25km	35	145	16.51	89.18
2	-15	10km	30	137	26.43	90.36	2	-7.5	25km	20	150	20.03	112.08	2	7.5	25km	30	137	17.28	95.55
2	-15	10km	33	157	25.36	86.03	2	-7.5	25km	23	132	19.75	108.67	2	7.5	25km	33	157	16.96	91.94
2	-15	10km	25	150	27.08	97.21	2	-7.5	25km	40	150	16.44	83.89	2	7.5	25km	25	150	18.3	102.68
2	-15	10km	27	135	27.07	94.71	2	-7.5	25km	38	153	16.8	87.49	2	7.5	25km	27	135	17.78	99.36
2	-15	10km	35	165	24.79	82.97	2	0	10km	35	145	24	83.25	2	7.5	25km	35	165	16.7	89.11
2	-15	10km	20	150	27.9	103.23	2	0	10km	30	137	24.88	88.79	2	7.5	25km	20	150	19.09	108.03
2	-15	10km	23	132	27.86	100.72	2	0	10km	33	157	24.34	85.29	2	7.5	25km	23	132	18.42	104.06
2	-15	10km	40	150	24.36	78.75	2	0	10km	25	150	25.92	96.07	2	7.5	25km	40	150	15.71	82.25
2	-15	10km	38	153	24.72	82.44	2	0	10km	27	135	25.45	92.39	2	7.5	25km	38	153	16.13	86.01
2	-15	15km	35	145	21.26	88.05	2	0	10km	35	165	23.95	82.43	2	15	10km	35	145	23.81	82.06
2	-15	15km	30	137	22.32	94.39	2	0	10km	20	150	26.9	102.16	2	15	10km	30	137	24.47	87.43
2	-15	15km	33	157	21.31	89.78	2	0	10km	23	132	26.25	97.52	2	15	10km	33	157	24.3	84.41
2	-15	15km	25	150	22.97	101.23	2	0	10km	40	150	23.16	77.43	2	15	10km	25	150	25.57	94.64
2	-15	15km	27	135	22.96	99.05	2	0	10km	38	153	23.59	81.29	2	15	10km	27	135	24.94	91.2
2	-15	15km	35	165	20.76	86.63	2	0	15km	35	165	20.11	86.75	2	15	10km	35	165	24.11	81.86
2	-15	15km	20	150	23.8	107.74	2	0	15km	30	137	20.99	92.72	2	15	10km	20	150	26.4	99.49
2	-15	15km	23	132	23.76	105.5	2	0	15km	33	157	20.45	89.02	2	15	10km	23	132	25.59	95.55
2	-15	15km	40	150	20.3	81.79	2	0	15km	25	150	21.98	99.98	2	15	10km	40	150	23.15	76.41
2	-15	15km	38	153	20.66	85.43	2	0	15km	27	135	21.54	96.43	2	15	10km	38	153	23.55	80.25
2	-15	20km	35	145	19.13	90	2	0	15km	35	165	20.07	86.08	2	15	15km	35	145	19.82	85.41
2	-15	20km	30	137	20.16	96.65	2	0	15km	20	150	22.95	106.52	2	15	15km	30	137	20.5	91.21
2	-15	20km	33	157	19.2	91.87	2	0	15km	23	132	22.31	101.77	2	15	15km	33	157	20.31	88.05
2	-15	20km	25	150	20.8	103.4	2	0	15km	40	150	19.26	80.38	2	15	15km	25	150	21.59	98.47
2	-15	20km	27	135	20.8	101.48	2	0	15km	38	153	19.69	84.21	2	15	15km	27	135	20.98	95.07
2	-15	20km	35	165	18.66	88.69	2	0	20km	35	145	18.1	88.68	2	15	15km	35	165	20.11	85.43
2	-15	20km	20	150	21.64	110.19	2	0	20km	30	137	18.96	94.93	2	15	15km	20	150	22.39	103.49
2	-15	20km	23	132	21.61	108.17	2	0	20km	33	157	18.44	91.11	2	15	15km	23	132	21.63	99.58
2	-15	20km	40	150	18.18	83.4	2	0	20km	25	150	19.92	102.09	2	15	15km	40	150	19.12	79.2
2	-15	20km	38	153	18.53	86.94	2	0	20km	27	135	19.49	98.69	2	15	15km	38	153	19.53	83.06
2	-15	25km	35	145	17.79	91.32	2	0	20km	35	165	18.06	88.14	2	15	20km	35	145	17.75	87.27
2	-15	25km	30	137	18.8	98.17	2	0	20km	20	150	20.87	108.89	2	15	20km	30	137	18.45	93.35
2	-15	25km	33	157	17.86	93.28	2	0	20km	23	132	20.25	104.13	2	15	20km	33	157	18.25	90.09
2	-15	25km	25	150	19.43	104.85	2	0	20km	40	150	17.24	81.94	2	15	20km	25	150	19.52	100.53
2	-15	25km	27	135	19.43	103.11	2	0	20km	38	153	17.65	85.7	2	15	20km	27	135	18.93	97.24
2	-15	25km	35	165	17.34	90.09	2	0	25km	35	145	16.83	89.98	2	15	20km	35	165	18.04	87.45
2	-15	25km	20	150	20.27	111.83	2	0	25km	30	137	17.69	96.42	2	15	20km	20	150	20.29	105.64
2	-15	25km	23	132	20.25	109.96	2	0	25km	33	157	17.17	92.52	2	15	20km	23	132	19.57	101.8
2	-15	25km	40	150	16.84	84.47	2	0	25km	25	150	18.61	103.5	2	15	20km	40	150	17.02	80.69
2	-15	25km	38	153	17.18	87.94	2	0	25km	27	135	18.21	100.21	2	15	20km	38	153	17.44	84.49
2	-7.5	10km	35	145	24.71	84.07	2	0	25km	35	165	16.79	89.54	2	15	25km	35	145	16.45	88.52
2	-7.5	10km	30	137	25.73	89.69	2	0	25km	20	150	19.56	110.47	2	15	25km	30	137	17.16	94.79
2	-7.5	10km	33	157	24.87	85.8	2	0	25km	23	132	18.95	105.71	2	15	25km	33	157	16.95	91.47
2	-7.5	10km	25	150	26.57	97.42	2	0	25km	40	150	15.95	82.99	2	15	25km	25	150	18.2	101.92
2	-7.5	10km	27	135	26.36	93.91	2	0	25km	38	153	16.37	86.69	2	15	25km	27	135	17.64	98.7
2	-7.5	10km	35	165	24.36	82.78	2	7.5	10km	35	145	23.73	82.6	2	15	25km	35	165	16.74	88.82
2	-7.5	10km	20	150	27.47	103.51	2	7.5	10km	30	137	24.47	88.06	2	15	25km	20	150	18.96	107.08
2	-7.5	10km	23	132	27.2	99.82	2	7.5	10km	33	157	24.18	84.78	2	15	25km	23	132	18.27	103.29
2	-7.5	10km	40	150	23.77	78.32	2	7.5	10km	25	150	25.58	95.43	2	15	25km	40	150	15.69	81.68
2	-7.5	10km	38	153	24.15	82.13	2	7.5	10km	27	135	24.98	91.65	2	15	25km	38	153	16.11	85.45
2	-7.5	15km	35	145	20.74	87.58	2	7.5	10km	35	165	23.92	82.04							
2	-7.5	15km	30	137	21.73	93.64	2	7.5	10km	20	150	26.46	100.43							
2	-7.5	15km	33	157	20.91	89.53	2	7.5	10km	23	132	25.68	96.22							
2	-7.5	15km	25	150	22.55	101.39	2	7.5	10km	40	150	23	76.87							
2	-7.5	15km	27	135	22.35	98.11	2	7.5	10km	38	153	23.42	80.74							

Bibliography

A. H. Strahler, J. M. et al. (1999). *MODIS BRDF Albedo Product: Algorithm Theoretical Basis Document, Version 5.0.* NASA Goddard Space Flight Center, Greenbelt Maryland.

Anderson, G. P., Berk, A., Acharya, P. K., Matthew, M. W., Bernstein, L. S., Chetwynd, Jr., J. H., Dothe, H., Adler-Golden, S. M., Ratkowski, A. J., Felde, G. W., Gardner, J. A., Hoke, M. L., Richtsmeier, S. C., Pukall, B., Mello, J. B., and Jeong, L. S. (1999). Modtran4: radiative transfer modeling for remote sensing. In *SPIE Proceedings*, volume 3866, pages 2–10.

Barsi, J., Lee, K., Kvaran, G., Markham, B., and Pedelty, J. (2014). The spectral response of the landsat-8 operational land imager. *Remote Sensing*, 6(10):10232–10251.

Biliouris, D., Verstraeten, W. W., Dutré, P., van Aardt, J. A. N., Muys, B., and Coppin, P. (2007). A compact laboratory spectro-goniometer (clabspeg) to assess the brdf of materials. presentation, calibration and implementation on fagus sylvatica l. leaves. *Sensors*, 7(9):1846–1870.

Brown, S. D. and Goodenough, A. A. (2015). Dirsig documentation manual. Technical report, Rochester Institute of Technology, Rochester, NY, USA.

Brown, S. D. and Schott, J. R. (2010). Verification and validation studies of the dirsig data simulation model. Technical Report 1, Rochester Institute of Technology, Rochester, NY, USA.

Chander, G., Mishra, N., Helder, D. L., Aaron, D. B., Angal, A., Choi, T., Xiong, X., and Doelling, D. R. (2013). Applications of spectral band adjustment factors (sbaf) for cross-calibration. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1267–1281.

Chopping, M. J. (1998). *Linear Semi-Empirical Kernel-Driven Bidirectional Reflectance Distribution Function Models in Monitoring Semi-Arid Grasslands from Space*. PhD thesis, University of Nottingham.

Cook, R. (1986). Stochastic sampling in computer graphics. *ACM Transactions on Graphics (TOG)*, 5(1):51–72.

Deering, D. W. (1988). Parabola directional field radiometer for aiding in space sensor data interpretations. In *Recent Advances in Sensors, Radiometry, and Data Processing for Remote Sensing*, volume 0924, pages 249–261.

Disney, M., Holopainen, M., Vastaranta, M., Kaartinen, H., Kaasalainen, S., Åkerblom, M., Kaasalainen, M., Raumonen, P., and Lewis, P. (2013). Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sensing*, 5(2):491–520.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P. (2012). Sentinel-2: Esa's optical high-resolution mission for gmes operational services. *Remote Sensing of Environment*, 120:25–36.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley.

Feng, X., Schott, J. R., and Gallagher, T. (1993). Comparison of methods for generation of absolute reflectance-factor values for bidirectional reflectance-distribution function studies. *Applied optics*, 32(7):1234–1242.

Feret, J.-B., François, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidel, L. P., Ustin, S. L., le Maire, G., and Jacquemoud, S. (2008). Prospect-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment*, 112(6):3030 – 3043.

Frederic Baret (2015). Prospect inversion. `http://teledetection.ipgp.jussieu.fr/prosail/`. Accessed: Jul 2015.

Goodenough, A., Raqueño, R., Bellandi, M., Brown, S., and Schott, J. (2006). A flexible hyperspectral simulation tool for complex littoral environments. In *Proceedings of the SPIE, Photonics for Port and Harbor Security II, Subsurface Change/Threat Detection*, volume 6204, pages 62040F–62040F–12. Proc. SPIE.

Goodenough, A. A. and Brown, S. D. (2015). Development of land surface reflectance models based on multiscale simulation. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI*, volume 9472, pages 94720D–94720D–8. Proc. SPIE.

Google Earth (2015). Google Earth. `http://maps.google.com/`. Accessed: Jul 2015.

Gutman, G. G. (1991). Vegetation indices from avhrr: An update and future prospects. *Remote Sensing of Environment*, 35(2):121–136.

Hasegawa, K., Matsuyama, H., Tsuzuki, H., and Sweda, T. (2006). The effect of bidirectional reflectance distribution function on the estimation of vegetation indices and leaf area index (lai): A case study of the vegetation in succession stages after forest fire in northwestern canada. *Journal of The Remote Sensing Society of Japan*, 26(3):186–201.

Hastie, T., Friedman, J. H., Tibshirani, R., and ebrary, I. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer, New York, second edition.

Herman Tullekem (2015). Poisson disc sampling. `http://devmag.org.za/2009/05/03/poisson-disk-sampling/`. Accessed: Jul 2015.

Hildebrandt, R. and Iost, A. (2012). From points to numbers: a database-driven approach to convert terrestrial lidar point clouds to tree volumes. *European Journal of Forest Research*, 131(6):1857–1867.

Jacquemoud, S. and Baret, F. (1990). Prospect: A model of leaf optical properties spectra. *Remote sensing of environment*, 34(2):75–91.

Jacquemoud, S., Ustin, S. L., Verdebout, J., Schmuck, G., Andreoli, G., and Hosgood, B. (1996). Estimating leaf biochemistry using the prospect leaf optical properties model. *Remote Sensing of Environment*, 56(3):194–202.

Kelbe, D., Romanczyk, P., van Aardt, J., and Cawse-Nicholson, K. (2013). Reconstruction of 3d tree stem models from low-cost terrestrial laser scanner data. In *SPIE Defense, Security, and Sensing*, pages 873106–873106. International Society for Optics and Photonics.

Knight, E. and Kvaran, G. (2014). Landsat-8 operational land imager design, characterization and performance. *Remote Sensing*, 6(11):10286–10305.

Kotchenova, S. Y., Vermote, E. F., Levy, R., and Lyapustin, A. (2008). Radiative transfer codes for atmospheric correction and aerosol retrieval: intercomparison study. *Applied Optics*, 47(13):2215–2226.

Koukal, T. and Schneider, W. (2010). Analysis of brdf characteristics of forest stands with a digital aerial frame camera. In *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVIII, pages 100–105. Proc. ISPRS.

Kuehl, R. (2000). *Design of Experiments: Statistical Principles of Research Design and Analysis*. Statistics Series. Duxbury/Thomson Learning.

Landsat 7 Science Data Users Handbook (2009). *Landsat 7 Science Data Users Handbook*. Landsat Project Science Office. NASA Goddard Space Flight Center, Greenbelt Maryland.

LANDSAT 8 (L8) DATA USERS HANDBOOK (2016). *LANDSAT 8 (L8) DATA USERS HANDBOOK*. Department of Interior, U.S. Geological Survey.

Leroy, M. and Roujean, J. . (1994). Sun and view angle corrections on reflectances derived from noaa/avhrr data. *IEEE Transactions on Geoscience and Remote Sensing*, 32(3):684–697.

Li, X. and Strahler, A. H. (1986). Geometric-optical bidirectional reflectance modeling of a conifer forest canopy. *IEEE Transactions on Geoscience and Remote Sensing*, GE-24(6):906–919.

Li, X. and Strahler, A. H. (1992). Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: effect of crown shape and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2):276–292.

Lucht, W., Schaaf, C. B., and Strahler, A. H. (2000). An algorithm for the retrieval of albedo from space using semiempirical brdf models. *IEEE Transactions on Geoscience and Remote Sensing*, 38(2):977–998.

Markham, B. L. and Helder, D. L. (2012). Forty-year calibrated record of earth-reflected radiance from landsat: A review. *Remote Sensing of Environment*, 122:30–40.

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., Gao, F., Kutler, J., and Lim, T.-K. (2006). A landsat surface reflectance dataset for north america, 1990-2000. *IEEE Geoscience and Remote Sensing Letters*, 3(1):68–72.

McManus, M., Schneeberger, N., Reardon, R., and Mason, G. (1992). Gypsy moth. usda forest service forest insect and disease leaflet 162. `http://na.fs.fed.us/spfo/pubs/fidls/gypsymoth/gypsy.htm`. Accessed: July 2016.

Michael Galyean (2016). Orthogonal Coefficients for Unequally Spaced Treatments . `https://www.depts.ttu.edu/afs/home/mgalyean///`. Accessed: May 2016.

Minitab (2015). Minitab. `http://www.minitab.com/`. Accessed: Aug 2015.

Mitchell, D. (1987). Generating antialiased images at low sampling densities. In *ACM SIGGRAPH*, pages 65–72. ACM.

Montgomery, D. C. (2012). *Design and analysis of experiments*. Wiley, Hoboken, NJ.

Morfitt, R., Barsi, J., Levy, R., Markham, B., Micijevic, E., Ong, L., Scaramuzza, P., and Vanderwerff, K. (2015). Landsat-8 operational land imager (oli) radiometric performance on-orbit. *Remote Sensing*, 7(2):2208–2237.

Onyx Computing (2015). Onyx tree. `http://www.onyxtree.com`. Accessed: Jul 2015.

Painter, T. H. and Dozier, J. (2004). Measurements of the hemispherical-directional reflectance of snow at fine spectral and angular resolution. *Journal of Geophysical Research: Atmospheres*, 109(D18).

Pinty, B., Gobron, N., Widlowski, J.-L., Gerstl, S. A. W., Verstraete, M. M., Antunes, M., Bacour, C., Gascon, F., Gastellu, J.-P., Goel, N., Jacquemoud, S., North, P., Qin, W., and Thompson, R. (2001). Radiation transfer model intercomparison (rami) exercise. *Journal of Geophysical Research: Atmospheres*, 106(D11):11937–11956.

Rahman, H., Pinty, B., and Verstraete, M. (1993). Coupled surface-atmosphere reflectance (csar) model 2. semiempirical surface model usable with noaa advanced very high resolution radiometer data. *Journal of Geophysical Research*, 98(D11):20791–20801.

RAMI (2015). RAdiation transfer Model Intercomparison (RAMI). `http://rami-benchmark.jrc.ec.europa.eu/HTML/`. Accessed: Jul 2015.

Reitberger, J., Schnörr, C., Krzystek, P., and Stilla, U. (2009). 3d segmentation of single trees exploiting full waveform lidar data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6):561–574.

Rengarajan, R. and Schott, J. R. (2016). Modeling of forest canopy brdf using dirsig. volume 9840, pages 98401F–98401F–14.

Rosell, J. R., Llorens, J., Sanz, R., Arnó, J., Ribes-Dasi, M., Masip, J., Escolà, A., Camp, F., Solanelles, F., Gràcia, F., Gil, E., Val, L., Planas, S., and Palacín, J. (2009). Obtaining the three-dimensional structure of tree orchards from remote 2d terrestrial lidar scanning. *Agricultural and Forest Meteorology*, 149(9):1505–1515.

Ross, J. K. (1981). *The radiation regime and architecture of plant stands.* W. Junk, The Hague, Netherlands.

Roujean, J. L., Leroy, M., and Deschamps, P. Y. (1992). A bidirectional reflectance model of the earths surface for the correction of remote-sensing data. *Journal of Geophysical Research-Atmospheres*, 97(D18):20455–20468.

Roujean, J. L., Leroy, M., Deschamps, P. Y., and Podaire, A. (1990). A surface bidirectional reflectance model to be used for the correction of directional effects in remote sensing multitemporal data sets. In *Geoscience and Remote Sensing Symposium, 1990. IGARSS '90. 'Remote Sensing Science for the Nineties'., 10th Annual International*, pages 1785–1789.

Sandmeier, S., Müller, C., Hosgood, B., and Andreoli, G. (1998). Sensitivity analysis and quality assessment of laboratory {BRDF} data. *Remote Sensing of Environment*, 64(2):176 – 191.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R. P., Hu, B., Liang, S., Privette, J. L., and Roy, D. (2002). First operational brdf, albedo nadir reflectance products from modis. *Remote Sensing of Environment*, 83(1):135–148.

Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S., and Martonchik, J. V. (2006). Reflectance quantities in optical remote sensing—definitions and case studies. *Remote Sensing of Environment*, 103(1):27–42.

Schott, J. R. (2007). *Remote sensing: the image chain approach.* Oxford University Press, New York.

Schott, J. R., Gerace, A. D., Brown, S. D., and Gartley, M. G. (2011). Modeling the image performance of the landsat data continuity mission sensors. In *SPIE Optical Engineering+ Applications*, pages 81530F–81530F. International Society for Optics and Photonics.

Shabanov, N. V., Wang, Y., Buermann, W., Dong, J., Hoffman, S., Smith, G. R., Tian, Y., Knyazikhin, Y., and Myneni, R. B. (2003). Effect of foliage spatial heterogeneity in the modis lai and fpar algorithm over broadleaf forests. *Remote Sensing of Environment*, 85(4):410–423.

Smith, D. (1986). *The Practice of Silviculture*. Wiley.

STK (2016). Systems ToolKit, agi . `http://www.agi.com/products/stk//`. Accessed: May 2016.

Storey, J., Choate, M., and Lee, K. (2014). Landsat 8 operational land imager on-orbit geometric calibration and performance. *Remote Sensing*, 6(11):11127–11152.

Storey, J. and Haque, M. O. (2016). Radiometric and geometric performance of sentinel 2a compared to landsat 8.

Storey, J. C. (2001). Landsat 7 on-orbit modulation transfer function estimation. In *International Symposium on Remote Sensing*, pages 50–61. International Society for Optics and Photonics.

Strugnell, N. C., Lucht, W., Hyman, A. H., and Meister, G. (1998). Continental-scale albedo inferred from land cover class, field observations of typical brdfs and avhrr data. In *Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS '98. 1998 IEEE International*, volume 2, pages 595–597 vol.2.

Surface Reflectance Product Guide (2016). *PROVISIONAL LANDSAT 8 SURFACE REFLECTANCE PRODUCT*. Department of Interior, U.S. Geological Survey.

Tan, B., Morisette, J. T., Wolfe, R. E., Gao, F., Ederer, G. A., Nightingale, J., and Pedelty, J. A. (2011). An enhanced timesat algorithm for estimating vegetation phenology metrics from modis data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2(2):361–371.

Teillet, P. M., Markham, B. L., and Irish, R. R. (2006). Landsat cross-calibration based on near simultaneous imaging of common ground targets. *Remote Sensing of Environment*, 102(3):264–270.

Tree diseases (2016). Trees and Shrubs: Diseases, Insects and Other Problems. `http://www.extension.umn.edu/garden/landscaping/maint/tree_disease.html/`. Accessed: May 2016.

United States National Bureau of Standards (1977). *Geometrical considerations and nomenclature for reflectance / F. E. Nicodemus ... [et al.].* U.S. Department of Commerce.

USGS (2016). Data available from the U.S. Geological Survey. `http://earthexplorer.usgs.gov/`. Accessed: May 2016.

Vermote, E., Justice, C., Claverie, M., and Franch, B. (2016). Preliminary analysis of the performance of the landsat 8/oli land surface reflectance product. *Remote Sensing of Environment.*

Walthall, C. L., Norman, J. M., Welles, J. M., Campbell, G., and Blad, B. L. (1985). Simple equation to approximate the bidirectional reflectance from vegetative canopies and bare soil surfaces. *Applied Optics*, 24(3):383–387.

Wang, Q., Adiku, S., Tenhunen, J., and Granier, A. (2005). On the relationship of ndvi with leaf area index in a deciduous forest site. *Remote Sensing of Environment*, 94(2):244–255.

Wanner, W., Li, X., and Strahler, A. H. (1995). On the derivation of kernels for kernel-driven models of bidirectional reflectance. *Journal of Geophysical Research: Atmospheres*, 100(D10):21077–21089.

Warner, T. A., Foody, G. M., and Nellis, M. D. (2009). *The SAGE handbook of remote sensing.* Sage Publications.

Widlowski, J.-L., Côté, J.-F., and Béland, M. (2014). Abstract tree crowns in 3d radiative transfer models: Impact on simulated open-canopy reflectances. *Remote Sensing of Environment*, 142:155–175.

Widlowski, J.-L., Mio, C., et al. (2015). The fourth phase of the radiative transfer model intercomparison (rami) exercise: Actual canopy scenarios and conformity testing. *Remote Sensing of Environment*, 169:418 – 437.

Widlowski, J.-L., Taberner, M., Pinty, B., Bruniquel-Pinel, V., Disney, M., Fernandes, R., Gastellu-Etchegorry, J.-P., Gobron, N., Kuusk, A., Lavergne, T., Leblanc, S., Lewis, P., Martin, E., Mõttus, M., North, P. J. R., Qin, W., Robustelli, M., Rochdi, N., Ruiloba, R., Soler, C., Thompson, R., Verhoef, W., Verstraete, M. M., and Xie, D. (2006). The third RAdiation transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance modelling. *Journal of Geophysical Research*, 112.