# ABSTRACT

Title of dissertation: IMPACT OF SATELLITE GEOMETRIC DISTORTIONS ON LANDSCAPE ANALYSIS: EFFECTS ON ALBEDO

Enrique Lugardo Montaño, Doctor of Philosophy, 2015

Dissertation directed by: Professor Christopher O. Justice
Department of Geographical Sciences

Data from wide field-of-view sensors have been providing information about the Earth's surface since the early 1980's. This manuscript is the result of investigations designed to determine the effective resolution and geometric variability of the NASA Earth Observing System MODerate Resolution Imaging Spectroradiometer (MODIS) and Visible Infrared Imager Radiometer Suite (VIIRS) gridded data. Although the wide field-of-view and high temporal frequency of MODIS provide near-daily global coverage, inconsistent observation assignment in geolocated MODIS pixels measurably demonstrates how spatial accuracy is affected by pixel-size growth (up to 4.8x) along-scan.

For studying the effective resolution, the point spread function of nominal 250m MODIS gridded surface reflectance products (L2G) was estimated from [man-made] large size targets. The findings indicate that in near-optimal locations the resolution of (sinusoidal grid) gridded products varies between 344m-835m along-scan for a range of viewing angles, but also indicate location-dependent variability

with along-scan and along-track ranges of 314m-1363m and 284m-501m respectively.

Albedo was identified as a well-known physical metric to study the effects of geometric variability, thus a broadband albedo using MODIS-like geometry was simulated for five EOS validation sites. Results of each site simulation exhibit compounded uncertainty attributable to the geometric distortion in ranges sufficient to influence climate models (i.e. ranges from 0.01-0.045 albedo).

A second series of broadband albedo simulations was developed for the same five EOS validation sites using VIIRS-like geometries and aggregation zones. Spatially heterogeneous land cover demonstrated a marginally significant difference in the mean albedo between aggregation zones ($< 0.015$). Results from data simulating temporal compositing, demonstrate the influence of geometric artifacts through differing levels of uncertainty between periods (i.e. ranges from 0.01-0.05 albedo).

The variability in both MODIS and VIIRS L2G questions the standard application of a global fixed grid, and indicates that regional projections combined with a representative grid cell 4x the nominal detector size (i.e. 1000m and 1500m for MODIS and VIIRS, respectively) are potentially useful for products using off-nadir views. This work ultimately resolves the surface-feature representation of temporo-spatial wide field-of-view instrument observations and quantifies the results of associating inherently-variable observations into an artificially-fixed and geometrically-regular space.

# IMPACT OF SATELLITE GEOMETRIC DISTORTIONS ON LANDSCAPE ANALYSIS: EFFECTS ON ALBEDO

by

Enrique Lugardo Montaño

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor Christopher O. Justice, Chair/Advisor
Professor Eric F. Vermote
Professor Manuel L. Campagnolo
Professor Samuel N. Goward
Professor Chengquan Huang

# Preface

Chapter 2 of this dissertation has been previously published as a co-authored peer-reviewed journal article. I was responsible for the primary data collection and processing, as well as substantial portions of the literature review in Chapter 2. Manuel Campagnolo was responsible for the model fitting and resolution estimates as well as the simulation processing. The final synthesis of results is the effort of both authors. The remainder of the research presented in this thesis was carried out in its entirety by me.

# Dedication

For Annie. You saw me all the way through.

# Acknowledgments

This work is not only the culmination of several years of intense effort on my part, but also the amazing work of my colleagues. I would like to thank my full committee for seeing me though this process and for their insight. In particular, I thank my committee member and co-author Manuel Campagnolo for all his efforts in developing the work in Chapter 2. The published work would not have been possible without his invaluable contribution and leadership in that effort.

I also owe a great debt of gratitude to Mashihiro Nishihama at the NASA Goddard Space Flight Center. Mash is the author of the code for simulating both MODIS and VIIRS geometry that is the cornerstone of the work in Chapters 3 and 4. My work presented here would have been far more difficult to achieve without this code and Mash's years of experience with MODIS and VIIRS geolocation.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1: Introduction

## 1.1 Research Objectives

"What is the actual resolution of MODerate Resolution Imaging Spectroradiometer (MODIS) data?" is the simple question this dissertation seeks to answer, presenting technical, quantitative findings, and continuing the investigation to include the Visible Infrared Imager Radiometer Suite (VIIRS) instrument, the next generation sensor extending MODIS-type observations into the future. The nominal resolution of the system (e.g. MODIS 250m), rather than a more complete understanding of the properties of the complete sensor system and processing stream employed [Forshaw et al., 1983, Joseph, 2000], is regularly used to characterize the data from these instruments. In order to utilize the finest-possible spatial resolution data from wide field-of-view instruments such as MODIS and VIIRS one must know the actualities of resolutions, in order to avoid extrapolating mischaracterizations into user applications.

MODIS and VIIRS science products and models currently resample and geolocate relative to the nominal sensor resolution (e.g. 500m) instrument observations; these protocols result in products and models with nominal resolutions that are not representative of actual observations. This dissertation examines the relationships

between nominal resolutions, instrument observations, and processing protocols. Recent moves to utilize native-scale observations in science products and models, such as the MODIS MCD43 product that changed from a 1km to 500m resolution as of Collection 5 [Schaaf et al., 2014], makes this research an important line of inquiry. Instruments such as MODIS and VIIRS exhibit a known growth of pixel sizes in off-nadir views [Barnes et al., 1998, Wolfe et al., 1998], questioning whether or not there is consistent representation of surface features in any given observation footprint as viewing geometry varies in different orbit passes.

## 1.2 Background

NASA's Earth Observing Systems (EOS) have been providing data to inform the science community of important changes to Earth's climate system since the early 2000's [NASA, 2014], and builds on the legacy of systems including NOAA's Advanced Very High Resolution Radiometer (AVHRR) and NASA's Sea-viewing Wide Field-of-view Sensor (SeaWiFS). The level of detail has improved with each new generation of sensor providing greater insight along the way. However, despite improvements, the tradeoff between temporal frequency and spatial capabilities remains; in order for a sensor to observe large portions of Earth's surface frequently, sensors must necessarily trade spatial resolution (i.e. pixel size) for scan width. Sensors such as the MODIS instrument are capable of making observations on a daily basis, but at the expense of a spatial resolution that starts at 250m.

Several sensor designs are used to accommodate the viewing requirements of

landscape-scale sensors, those sensors capable of viewing large swaths of Earth's surface at a time. The MODIS instrument employs a scanning mirror to achieve a 2000+ km view of the surface on each orbit [Barnes et al., 1998]. The constantly scanning mirror provides spectral signal to an array of detectors of various size and spectral sensitivity creating a discrete array of pixels. The discrete array is not, however, representative of signal from a fixed space.

An instrument's Point-Spread Function (PSF) describes the response of an imaging system to an idealized point illumination source (2-dimensional). The degree to which the point source is blurred within the boundary of a pixel, for example, defines the spatial response of the system in all directions. The Line-Spread Function (LSF) of an imaging system is the 1-dimensional response in either the along scan direction (x) or in the along track direction (y). It is usual to assume that the PSF can be separated into the two components of the LSF, and therefore can be written as the product $PSF_{xy}(X,Y) = LSF_x(X) * LSF_y(Y)$.

The scanning system in MODIS forms a triangular LSF along scan, integrating radiometric signal from the immediate preceding and subsequent neighbor pixels (12.5%) with the center pixel (75%) [Nishihama et al., 1997]. The portion of physical space observed is dependent on the instantaneous ground-projected sample interval (the spacing between pixels), also referred to as the ground sample distance (GSD) [Schowengerdt, 2007], which is subject to the geometric growth of the observation at increasing view zenith angles (VZA), the angle formed between a line perpendicular to the observed surface (local zenith) and a line pointing back to the instrument. The variable pixel size forms overlapping scans toward the edge of scan, referred

to as bow tie, owing to the larger size at the edges narrowing at the middle, then growing again [Wolfe et al., 1998].

A generally accepted measure of spatial resolution is the maximum distance at which 50% contrast is observed between pixels, known as full width at half maximum (FWHM). The spatial resolution values of 250m, 500m and 1000m often used to describe MODIS data correspond to the FWHM of the MODIS instrument at nadir in laboratory conditions and do not reflect the continuously varying VZA, the overall effect of which is a growth in the GSD of approximately 2.0 times along track and 4.8 times along scan at the maximum observation angle [Wolfe et al., 1998]. The reported pre-launch FWHM along scan spatial resolution for the 250m MODIS bands exceed 290m [Barnes et al., 1998].

The primary concern in this dissertation is the fundamental variability of geographic space being observed as the instrument geometry changes from orbit-to-orbit; a given location on Earth's surface will be observed from different directions and with varying level of contribution from neighboring landscape every time it is observed. Therefore, the assumption of target stability through changing geometry in any spatial sense is negated.

These irregular observation sizes from daily orbit variation make consistent geolocation of the same observation footprint difficult, thus a process was developed to assign the variable space to a fixed map projection with known pixel size and locations. The consequence of this assignment is to fit an oblong peg into a square hole. Figure (1.1) illustrates the assignment of irregular space to a fixed pixel.

Most higher order MODIS biophysical measurement and model products are

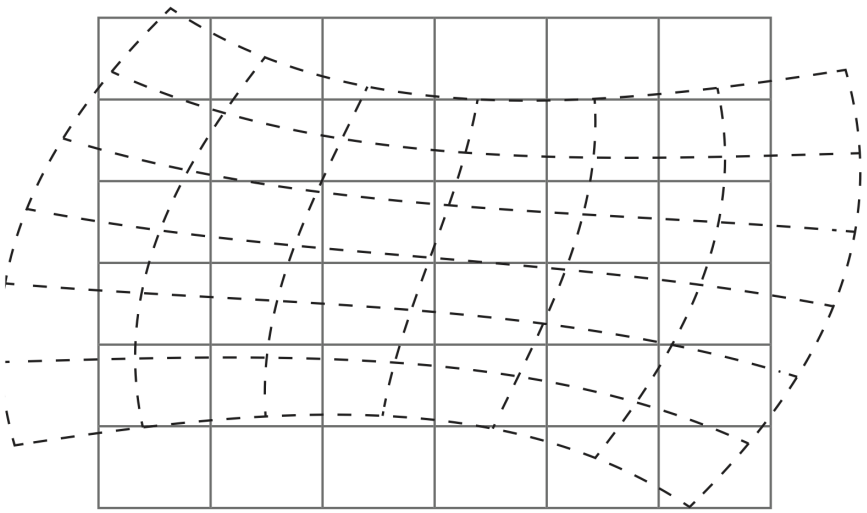


Figure 1.1: Illustration of the relationship between the irregular observation space (dotted line) and the fixed-pixel projected grid space (solid line). The bow tie effect elongates pixels primarily in the scan direction (left to right), increasing their size as a result. The observation at any given time and location is the integration of signal from the [variable] ground sample distance footprint based on the viewing geometry. The resultant observation is assigned to a grid with a fixed pixel size.

assigned to a geo-referenced grid using a custom sinusoidal projection with a fixed pixel size for each of the available nominal resolutions. MODIS geolocation assumes a fixed spatial size based on the subdivision of a 30 arc second grid providing a seamless global grid with resolutions of $\approx$ 231m, $\approx$ 463m, and $\approx$ 926m [Wolfe et al., 2002]. Observations made at resolution $r$ (e.g. 250m) are assigned to a co-located $(x, y)$ coordinate in the geolocation grid of the same resolution.

The reported mean observation-to-grid assignment for MODIS grid data is less than 30%, meaning the value assigned to any given grid pixel is only a portion of the spatial signal from the observation pixel, and not necessarily representative of that smaller geographic grid space [Tan et al., 2006]. Further, single observations are often assigned to more than one grid pixel at large VZA's.

The MODIS instrument's on-orbit spatial response based on Level 1B swath data is well understood [Butler and Xiong, 2006, Rojas et al., 2002, Salomonson et al., 2000], and artifacts from the process of geolocation gridding have been studied by [Tan et al., 2006]. Similarly, on-orbit reports of the VIIRS instrument spatial performance indicate the pixel aggregation scheme is effective [Lin et al., 2013, Wolfe et al., 2012, 2013].

However, what has been missing is an account of an "effective resolution" for Level 2 Gridded (L2G) data, which form the basis for all higher order MODIS and VIIRS science products [Justice et al., 2002, 2013]. Excluding an explicit representation of changing observation contributions through time and space in the L2G product presents uncertainty in the estimation of spatial resolution. That is, if the actual per-pixel observation contribution is assumed to be constant as with the ge-

olocation grid, then the resultant resolution will be underestimated at larger view angles.

## 1.3 Research Questions

This dissertation investigates the compounded effects of a variable observation geometry and geolocation assignment on measurements by posing the following focused research questions:

1. What is the effective spatial resolution of gridded MODIS data? A more informed usability of data is achieved by quantifying the effective resolution of existing gridded MODIS data, the basis of higher-order land biophysical measurement and model products.

2. What impact does variable view geometry have on MODIS observations? A better understanding of error sources is achieved by measuring the effect of variable geometric footprints in simulated MODIS gridded data.

3. Does the VIIRS instrument improve the consistency between variable spatial footprints? Continuity in next generation sensor data is assessed by investigating the effectiveness of applied error mitigation methods.

Both existing science-quality and simulated data were used to demonstrate various aspects of the effects that changing sensor view geometries have on the assumption of spatial resolution and spatial representation.

## 1.4 Outline of Dissertation

This dissertation is organized in five chapters including the introduction and conclusion. The three primary chapters are self-contained and structured in the format of journal articles. Literature pertinent to each chapter is reviewed in that chapter, though there is overlap between chapters since the subject matter in each chapter is closely related.

Chapter 2 establishes the effective resolution of MODIS Level 2 Gridded (geolocated) surface reflectance, which forms the basis of most higher order biophysical products from the MODIS Land Science Team. Chapter 3 takes the results from Chapter 2 for MODIS to test the assumption that the geographic space in the geolocated data is stable across the range of view geometries using broadband albedo as a well-understood physical measurement. Chapter 4 investigates the same assumption by extending the methods from Chapter 3 to VIIRS view geometry. Finally, Chapter 5 summarizes the dissertation and discusses the implications of the findings. The framework of this dissertation is shown in Figure 1.2.

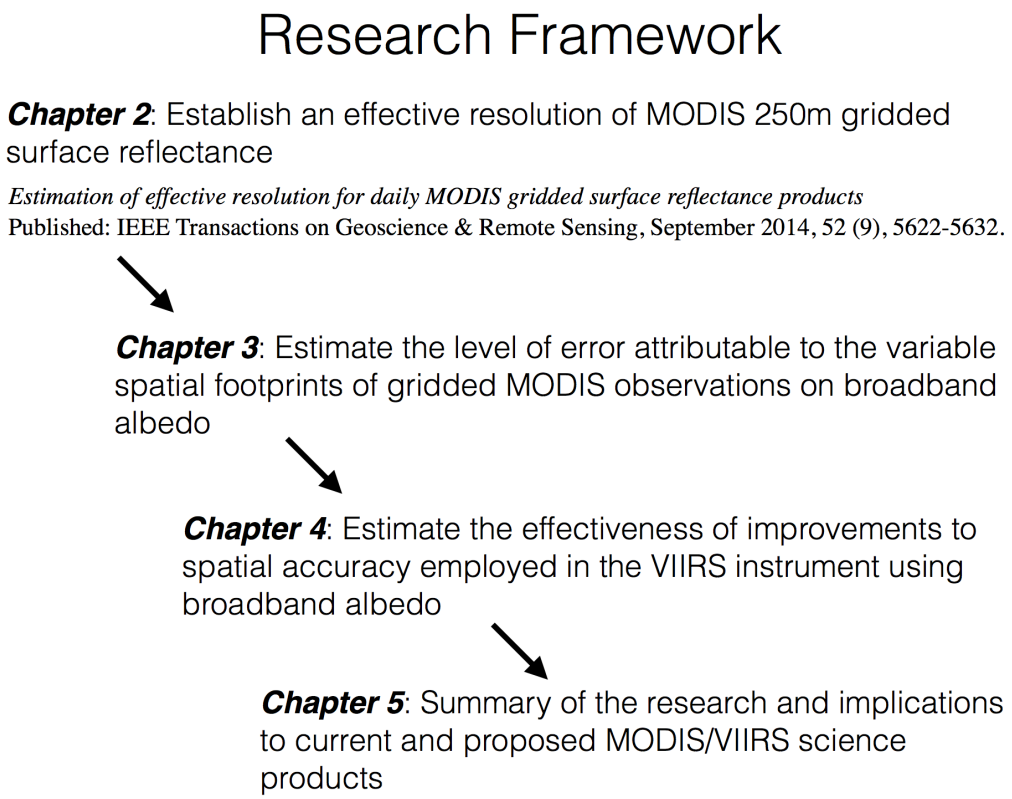



Figure 1.2: Illustration of research framework. Chapter titles are in bold italics.

# Chapter 2: Estimation Of Effective Resolution For Daily Modis Gridded Surface Reflectance Products

The following chapter was published by the IEEE Transaactions on Geoscience & Remote Sensing.  Manuel L. Campagnolo and Enrique L. Montaño, Estimation of effective resolution for daily MODIS gridded surface reflectance products, IEEE Transactions on Geoscience & Remote Sensing, September 2014, 52 (9), 5622-5632.

## 2.1 Abstract

Data from wide field-of-view sensors have provided a wealth of information of the Earth's surface for many years. Of the more recent efforts, the NASA Earth Observing System MODerate Resolution Imaging Spectroradiometer (MODIS) instrument has been operational since 2000 (for Terra, mid-2002 for Aqua). The wide field of view and high temporal frequency of MODIS provides near daily global coverage. However, this coverage comes at the cost of some level of spatial accuracy due to known effects of pixel size growth at increasing view zenith angles (VZA). Further complicating the matter of variable viewing geometry is the application of a fixed grid for the geolocation of MODIS observations. An accurate understanding of this

process is necessary to characterize the effective spatial resolution of daily MODIS gridded products. In this paper we estimate the point spread function (PSF) of nominal 250m MODIS gridded surface reflectance products from sequences of daily images over man-made large size targets. Our results suggest that in near optimal locations the resolution of those MODIS gridded products varies between 344m and 835m along rows and between 292m and 523m along columns when the VZA ranges from, respectively, $0^o$ (nadir view) to $55^o$. We also discuss some implications of the reliance on a global fixed grid, like MODIS sinusoidal grid, on the relation between location and spatial resolution.

## 2.2 Introduction

A key facet in representing features on the ground is understanding how those features are organized in space and how that space is depicted as a function of the spatial resolution of the imaging system [Joseph, 2000]. Spatial resolution in remote sensing is often considered only as the characteristic pixel resolution reported for the sensing system used, but is in fact far more complex [Forshaw et al., 1983]. Schowengerdt [2007] provides extensive definitions and examples of aspects of spatial resolution, representing the cumulative result of optical properties of the sensing system. Several authors have reported on the errors that are due to the contamination of the signal from neighboring pixels and suggested improvements to per-pixel estimates of land cover by incorporating known sensor characteristics [Huang et al., 2002, Townshend et al., 2000]. MODIS, flown on two polar orbiting platforms, is a

whiskbroom system with a scanning mirror providing a wide observation swath at a high temporal frequency for 36 spectral bands at nominal detector pixel resolutions of 250m, 500m, and 1000m [Barnes et al., 1998], and has established the benchmark for high quality landscape to global scale land observations [Gitelson et al., 1996, Huete et al., 1999, Ji et al., 2008]. This accomplishment has been achieved through a well-calibrated sensor, extensive derivative data products along with the development and implementation of quality control [Roy et al., 2002] and validation efforts [Morisette et al., 2002].

The wide field-of-view of 110° [Wolfe et al., 1998] and high temporal frequency of MODIS provides near daily global coverage, but at varying spatial resolution due to known effects of pixel size growth at increasing view zenith angles (VZA) [Tan et al., 2006]. The pixel growth ultimately leads to overlaps toward the edge of scan known as the *bow-tie* effect. The most common measure of a sensor spatial resolution is the ground sample distance (GSD), also known as the ground-projected instantaneous field-of-view (GIFOV), which matches by design the ground-projected sample interval, i.e. the spacing between pixels on the ground [Schowengerdt, 2007]. The reported values of 250m, 500m and 1000m correspond to the GSD at nadir [Barnes et al., 1998] although MODIS observations are made at continuously varying VZA. The overall effect of increasing VZA is a growth in the GSD of approximately 2.0 times along track and 4.8 times along scan for the maximum Earth observation angle [Nishihama et al., 1997, Wolfe et al., 2002]. Higher latitudes provide increasing orbital coverage on any given day presenting the possibility of more than one VZA reported for a single daily product [Vermote et al., 2011]. Further complicating the matter of variable

viewing geometry is the application of a fixed grid for the geolocation of MODIS observations [Wolfe et al., 1998, 2002]. The GSD is independent of the gridding process. Therefore, reporting the nominal GSD of 250m, 500m, or 1000m as the resolution of MODIS gridded products is rather misleading, particularly concerning any attempt at temporal compositing.

MODIS gridded products are the output of a *sampled image system* [Park et al., 1984] that combines an *imaging system* and a *sampling procedure*. The imaging system is characterized by the *sensor point spread function* (PSF), which models the spatial response of the sensor to the measured physical signal [Schowengerdt, 2007]. The sensor PSF is usually assumed to be separable, which means that it can be written as $\text{PSF}_{\text{sensor}} = \text{LSF}_{\text{s}} \times \text{LSF}_{\text{t}}$, where LSF is the *line spread function*, either in the along scan direction ($\text{LSF}_{\text{s}}$) or in the along track direction ($\text{LSF}_{\text{t}}$), and was characterized before launch for the MODIS instrument [Barnes et al., 1998]. To assess the resolution of gridded MODIS products it is essential to take into consideration the sampling procedure as well. The sampling procedure converts the output of the imaging system to an equally spaced, geolocated rectangular grid, producing the images that most users rely on.

There are two alternatives to estimate the PSF of the whole sampled image system. The first approach is to model both the imaging system and the sampling procedure to derive the overall PSF. This is the approach followed in [Park et al., 1984] where the imaging system PSF is assumed to be uniform. However, Park et al. [1984] computes just the *average* system PSF over all possible positions of the physical signal. If the exact position of the source of the signal is known, a more

precise overall PSF estimate can be obtained from a gridded image over a target on the surface. We will follow this latter approach, using a target suitable for MODIS and a technique that does not rely on high resolution imagery.

The well known dikes in Netherlands are large size man-made features that are compatible in scale with MODIS imagery. We consider two of those features, one oriented along the North-South direction (NS edge) and the other along the East-West direction (EW edge), which are straight for over 10km and separate water from agricultural fields. We use two years of observations to estimate the PSF for the 250m MODIS gridded reflectance products. We also use the actual locations of the swath and grid observation cells for the same dates and over the same targets to generate a synthetic MODIS gridded product. Synthesized data provides a template for observations in the absence of noise (e.g. atmospheric affects, edge imperfections) and illustrate in detail the effect of the sampling procedure in the final gridded product. Formally, the PSF of the MODIS gridded product can be written as the convolution (see [Park et al., 1984, Schowengerdt, 2007])

$$\mathrm{PSF}_{\mathrm{grid.product}} = \mathrm{PSF}_{\mathrm{sensor}} * \mathrm{PSF}_{\mathrm{sampling}} * \mathrm{PSF}_{\mathrm{noise}} \tag{2.1}$$

where the three components of $\mathrm{PSF}_{\mathrm{grid.product}}$ model respectively the imaging system, the sampling procedure and the additional noise factors that affect the MODIS daily gridded products. In this paper we estimate directly $\mathrm{PSF}_{\mathrm{grid.product}}$ from MODIS daily Terra and Aqua Collection 5 MOD09GQ and MYD09GQ surface reflectance products, and we estimate $\mathrm{PSF}_{\mathrm{sensor}} * \mathrm{PSF}_{\mathrm{sampling}}$ from the synthetic data mentioned above.

Specifically, we model the PSF of the sampled image system (i.e. the gridded

image, either real or synthetic) as the product of the along row LSF and the along column LSF, each one parameterized as a function of its *full width at half maximum* (FWHM). For the MODIS imaging system as described in [Barnes et al., 1998], the GSD and the FWHM are approximately equal. However, while the GSD characterizes the imaging system but is independent of the sampling procedure, the FWHM can be used to characterize either the imaging system or the whole sampled system (the gridded product). Hence, even if both GSD and FWHM increase with the VZA, only the FWHM is relevant for characterizing the resolution of the gridded products.

The paper is organized as follows: in Section 2.3 we describe the MODIS imaging system PSF and the sampling procedure for MODIS gridded products. In Section 2.4 we present the study area and data and we discuss the central concepts of line, edge and point spread function, and review how those functions can be estimated from the data. We also describe how we synthesize daily MODIS gridded products that simulate ideal observation conditions. Finally, in Section 2.5 we present our results and estimate the PSF for gridded reflectance products as a function of the VZA. We end the paper with a discussion of the implications of using a global grid, like the sinusoidal grid for MODIS, on the relation between location and spatial resolution.

## 2.3 The MODIS instrument and the land level 2 grid

The MODIS instrument includes several components to monitor on-orbit performance [Barnes et al., 1998, Xiong and Barnes, 2006]. The spatial performance of

MODIS is monitored by the on-board spectral radiometric calibration assembly that indicates that the PSF has been stable over time for the 250m bands [Xiong et al., 2006]. The MODIS instrument uses a scanning mirror to satisfy the requirement of a large (2330km) swath. The scanning mirror continuously passes light from the instrument telescope onto three focal plane assemblies containing 10, 20 and 40 detectors for the 1000m, 500m, and 250m bands respectively. The constant scan speed of the mirror causes an integration of the signal from the immediate preceding and following neighbor pixels, with the signal being collected from the center pixel (75%) and from each of its neighbors (12.5%) [Nishihama et al., 1997], and leads to the triangular along scan LSF depicted in Figure 1. The observation footprint grows in size with the scan angle as described in [Tan et al., 2006, Appendix B]. The scan angle is related to the larger *view zenith angle* (VZA) – the angle at the ground position between a ray pointing toward the sensor and one pointing toward the zenith, which is also called the *sensor zenith angle* [Nishihama et al., 1997, Wolfe et al., 2002] – by the expression $\text{VZA} = \arcsin(\frac{\text{R+H}}{\text{R}} \sin(\text{scan angle}))$, where $R$ is Earth's radius and $H$ is the satellite's altitude, for a spherical Earth model.

In this paper, we are concerned mostly with the 250m nominal resolution products, where the GSD is 250m at nadir and increases until it reaches 1207.5m in the along scan direction and 502.5m in the along track direction when the scan angle is $55^{\text{o}}$ ($\text{VZA} \approx 66^{\text{o}}$), as illustrated in [Wolfe et al., 1998, Fig. 2]. Fig. 2.1, which is similar to [Tan et al., 2006, Fig. 2], shows the approximate shape of the sensor PSF according to [Barnes et al., 1998] and illustrates the fact that the FWHM of the imaging system is approximately equal to the GSD and therefore grows with

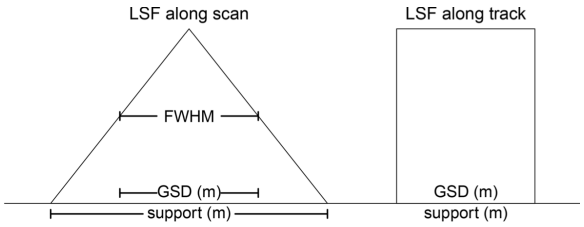



Figure 2.1: The figure describes the approximate $\text{PSF}_{\text{sensor}}$ decomposed into along scan $\text{LSF}_{\text{s}}$ and along track $\text{LSF}_{\text{t}}$ according to [Barnes et al., 1998]. The support is the area on the ground that contributes to the detector response. The FWHM is the width of the LSF at half its maximum height.

the VZA in both along scan and along track directions. The area on the ground that contributes to the response at the detector, i.e. the support of the PSF depicted in Fig. 2.1, is therefore twice as large as the GSD in the along scan direction.

All the above considerations are about the GSD associated to the detectors, and correspond to the *observation cells* (or just *cells*). All MODIS data up to and including level 2 are in an ungridded orbital swath arrangement of those cells [Vermote et al., 2011]. All higher order MODIS biophysical observation and model products (so-called Level 3 and 4) are assigned to a georeferenced grid with a fixed pixel size for each of the available nominal resolutions (250m, 500m, 1000m). As a means of separating geolocating from compositing and averaging, a gridded level 2 data (L2G) was developed, followed by a more user-friendly version (L2G-lite) that provides a minimal level of compositing of daily level 2G data [Vermote et al., 2011]. The data at the core of our analysis is at the L2G-lite daily MOD09GA and MYD09GQ reflectance products, the major daily MODIS reflectance products and also the input data for Collection 5, Level 3, 8-day reflectance composites MOD/MYD09A1 and MOD/MYD09Q1 [Vermote et al., 2011, Table 1]. All publicly available gridded MODIS surface reflectance products are geolocated within the sinusoidal tile grid

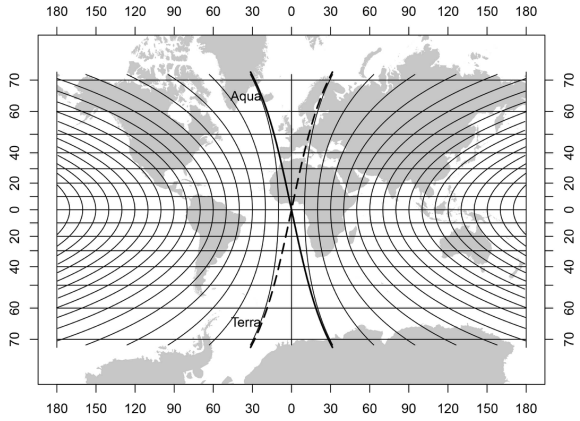



Figure 2.2: Mercator projection of the SIN grid between latitudes 70°S and 70°N. The thick solid and dashed lines represent respectively the tracks of an Aqua ascending orbit and a Terra descending orbit that cross the origin determined with the Orbit Determination Tool Kit (ODTK) software. The Mercator (conformal) projection was chosen to preserve the angles between the grid and the sensor track all over the globe.

(SIN) outside the polar regions, with the exception of the 0.05° product which is based on a Climate Modeling Grid (CMG) [Vermote et al., 2011]. The selection of the SIN grid is discussed in [Yang and Wolfe, 2001]; in particular it is an equal area gridding scheme based on the subdivision of a 30 arc second grid providing a seamless global grid with pixel sizes of $\approx$ 231m, 463m, and 926m. The grid is oriented East-West horizontally, but it is not oriented North-South except at the Equator or at 0° longitude. Its deviation from the vertical direction grows with both latitude and longitude as illustrated in Fig. 2.2.

The sampling procedure associated to MODIS gridding described by [Wolfe et al., 1998] preserves the original value for each cell. In that sense, it is essentially a nearest-neighbor approach. Originally, the values of all cells that intersected a given grid pixel were retained in the L2G products. However, since the support of the $\text{PSF}_{\text{sensor}}$ (see Fig. 2.1) is wider than the cell, the selection in [Wolfe et al.,

1998] was too strict. The L2G approach was improved [Yang and Wolfe, 2001] in a number of ways. Not only the previous limitation was fixed, but a method to limit the size of the revised L2G product while keeping for each grid pixel the cell values with the highest integrated $\text{PSF}_{\text{sensor}}$ over the grid pixel (the so-called *obscov*, which can range from 0 to 1) was devised. This new approach has been used in MODIS gridding starting in Collection 4 products [Tan et al., 2006]. The daily L2G-lite data is derived from L2G with the primary goal of keeping in the L2G-lite first data layer (the one most users rely on) the best possible information from the daily orbits of the satellite [Vermote et al., 2011].

We have described how values for each cell are assigned to pixels in the SIN grid based on the obscov criteria (see Fig. 2.6 for examples of this relationship). Hence, observations are often shifted and/or assigned multiple times because of the mismatch between cells and grid pixels, with mean obscov reported to be less than 0.3 [Tan et al., 2006]. The distortions resulting from assigning irregular observation space data to a regular grid space has been studied, in particular with respect to temporal compositing and band-to-band registration [Tan et al., 2006]. However, what has been missing is an account of the effect of gridding on the overall PSF of L2G data, which form the basis for all higher order MODIS science products [Justice et al., 2002].

## 2.4 Data and Methods

### 2.4.1 Line, Edge and Point Spread Functions

In analogy to the sensor PSF, we consider that the sampled image system PSF is also separable as $\text{PSF}_{\text{system}} = \text{LSF}_{\text{row}} \times \text{LSF}_{\text{col}}$ where $\text{LSF}_{\text{row}}$ and $\text{LSF}_{\text{col}}$ are respectively along the rows and the columns of the gridded product. There are many ways of estimating a directional LSF from remote sensing imagery [Pagnutti et al., 2010]. Schowengerdt *et al.* [Schowengerdt et al., 1985] consider both two-image analysis and target methods to estimate the LSF for Landsat TM images. The former is a technique that requires two simultaneous images, one from the sensor of interest and one from a higher resolution sensor. The two-image technique was also used in [Rojas et al., 2002] to estimate the spatial response from coincident Landsat ETM+ imagery. Since we are interested in analyzing daily time series of MODIS images this method is problematic; indeed, no daily time series of moderate- or high-resolution images are available to us.

The so-called target methods rely on the existence of a target on the ground. Among those we can distinguish the edge method, the pulse method, and the impulse method [Helder et al., 2003], where the high contrast targets are, respectively, a *knife edge* (a straight edge with a wide area), a straight narrow long object, and a point source (a beam of light or mirror). The pulse target can be for example a bridge [Schowengerdt et al., 1985] or an artificial target with known width. The target for the edge method can be artificial, like a high contrast tarpaulin or painted area

on the ground [Schowengerdt et al., 1985], or a natural target, like the edge of a parking lot or a dam [Ruiz and Lopez, 2002]. Given the spatial resolution of MODIS images (250m and larger) it is difficult to find a suitable pulse target. However, we can find an edge target of large enough dimension. In this study we used two edges of Noordoostpolder near Urk, in the Netherlands (see Section 2.4.2), which are over 10km long and perfectly straight, providing excellent targets for the edge method.

The goal of the edge method is to estimate the sensor edge spread function (ESF) along some direction. Then, the corresponding LSF can be derived by differentiation. To estimate the ESF we take a sequence of reflectance measurements that cross either the NS edge (for $\text{ESF}_{\text{row}}$) or the EW edge (for $\text{ESF}_{\text{col}}$). Since a single row (respectively, a column) of the image can lead to a poor estimate of the $\text{ESF}_{\text{row}}$ (respectively, $\text{ESF}_{\text{col}}$), a denser series is needed. Hence, our input is composed of several superimposed rows (or columns) aligned about the edge. To estimate a *non-parametric* ESF we follow the standard approach (see [Helder et al., 2003]) that consists of: (i) aligning the image with respect to the edge (see Section 2.4.2), (ii) adjusting a spline function to each row, and (iii) averaging the splines to obtain a non-parametric estimation of the ESF which is normalized between 0 and 1. Fig. 2.3 illustrates that approach applied to 23 rows of Aqua data over the NS edge of our study area (see Section 2.4.2).

The non-parametric approach described above leads to a sigmoid shaped ESF that can be in general well approximated by a *parametric* expression. In our analysis the major parameter is the FWHM. We denote by $\text{FWHM}_{\text{row}}$ the common parameter for $\text{LSF}_{\text{row}}$ and $\text{ESF}_{\text{row}}$ and we define analogously $\text{FWHM}_{\text{col}}$. The qualitative behavior

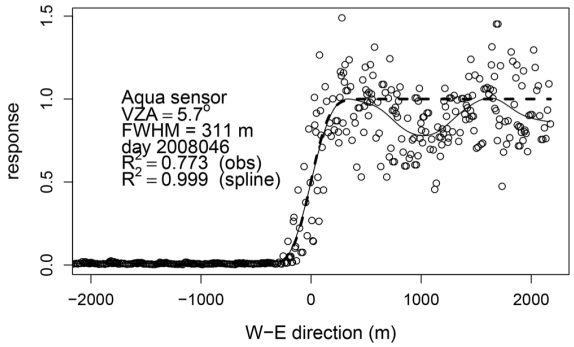



Figure 2.3: Estimation of the along row edge spread function ($\text{ESF}_{\text{row}}$). The dots represent the reflectance values along the aligned $x$ values and the solid line is the estimated non-parametric ESF. All values are normalized so the ESF ranges from 0 to 1. The sensor is Aqua and VZA = 5.7° for the day of acquisition 2008046 (yyyyddd). One can notice that the signal is very uniform on the water side of the edge and it has more variability on the land side, where the land cover is a mixture of forest and agriculture land. The dotted line is the parametric ESF adjusted to the data which has FWHM =311m. The goodness of fit of the parametric ESF is 0.773 with respect to the aligned observations (obs) and it is 0.999 with respect to the non-parametric ESF (spline).

of the imaging system (modeled by $\text{PSF}_{\text{sensor}}$) and the sampled image system (modeled by $\text{PSF}_{\text{system}}$) are similar in the analysis carried out by [Park et al., 1984, see Fig. 6 and 7] (those authors exhibit the modulation transfer function of the systems rather than the PSF) even though the sampled image system is, as expected, blurrier than the imaging system. This leads us to consider as a first approximation of the $\text{PSF}_{\text{system}}$ a trapezoidal family of LSF models, with two parameters: the FWHM and the ratio ($\alpha$) between the bases of the trapezoid. Varying $\alpha$ we obtain shapes ranging from the triangular LSF ($\alpha = 0$) to the rectangular LSF ($\alpha = 1$) depicted in Fig. 2.1. We also consider the well known Gaussian kernel [Schowengerdt, 2007] already used to assess the impact of the PSF in MODIS products [Huang et al., 2002]. To compare the different parametric models we compute the goodness of fit both of a given parametric ESF with the corresponding non-parametric ESF and the sequence

of reflectance values across the edge (see Fig. 2.3). The goodness of fit is only computed over the range where the parametric ESF is effectively increasing, so it's not influenced by the data away from the edge signal. Our analysis show that the best parametric models are piecewise quadratic ($\alpha = 1$) and the Gaussian. It turns out that for any $x$ and any value of the parameter FWHM between 250m and 800m, the expressions for the piecewise quadratic ($\alpha = 1$) and the Gaussian ESF differ at most 0.012 (over the ESF range [0,1]) and, therefore, are in practice interchangeable. Thus, we adopt the Gaussian model for the (unidimensional) LSF, i.e.

$$\mathrm{LSF}(\mathrm{z}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\mathrm{z}^2}{2\sigma^2}\right),$$

where the relation between $\sigma$ and the FWHM is given by $\mathrm{FWHM} = 2.355\,\sigma$. Hence the 2-dimensional PSF of the sampled image system is modeled by Equation 2.2, where $a = \mathrm{FWHM}_{\mathrm{row}}/2.355$ and $b = \mathrm{FWHM}_{\mathrm{col}}/2.355$.

$$\mathrm{PSF}(\mathrm{x},\mathrm{y}) = \frac{1}{2\pi\,\mathrm{ab}} \exp\left(-\frac{1}{2}\left[\frac{\mathrm{x}^2}{\mathrm{a}^2} + \frac{\mathrm{y}^2}{\mathrm{b}^2}\right]\right) \tag{2.2}$$

### 2.4.2 Study area and data

The study area is located at the western edge of Noordoostpolder near Urk, in the Netherlands (Fig. 2.4). The site was specifically chosen for the presence of linear edges between water and agricultural fields. The water feature provides an excellent linear source of contrast to the land features with strong absorption in the NIR spectrum; the edges under consideration provide a very sharp transition between those two features as illustrated by Fig. 2.4. The site presents enough area

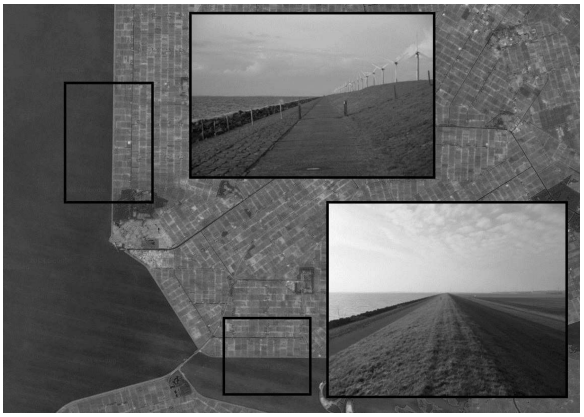


Figure 2.4: The study area in the Netherlands (Lat = 52.697°N, Long = 5.593°E) was chosen for the linear transition between water and agriculture to produce a clear line of contrast between features. The hollow boxes indicate the positions of the sub-images around the NS edge (top) and the EW edge (bottom) where MODIS reflectance products were extracted. The inserted pictures show the sharp transitions between water and a surface mostly covered by vegetation (source: GoogleEarth ID 10896477 for the NS edge and ID 44047111 for the EW edge).

on either side of the edges for sufficient sampling of MODIS observations at multiple view angles. The bounding boxes in Fig. 2.4 shows the extent of the surfaces used for sampling the MODIS data.

Data for this study consists of daily Terra and Aqua Collection 5 MOD09GQ and MYD09GQ (nominal 250m) surface reflectance products for MODIS tile h18v03 during a two-year period from 2007 to 2008 (in the remainder of the paper M09GQ will denote the combined dataset MOD09GQ and MYD09GQ). We use band 2 (NIR) which provides a clearer distinction between the water feature and agricultural fields. Angle data, stored in 1km grids within the companion MOD09GA and MYD09GA products are averaged to provide the mean VZA for each observation date over the study area. In addition we use the cell positions and the corresponding VZA values from level 1 MYD03, collection 5, to generate synthetic observations for Aqua (see Section 2.4.3). All data sources are indicated in Table 2.1.

| Product | Field | Range | |
|---|---|---|---|
| | | full scenes | study area |
| MYD03 | Latitude | 31.62–80.57 | 52.54–52.78 |
| | Longitude | -54.11–65.32 | 5.43–5.83 |
| | SensorZenith | 0.01–65.62 | 1.81–55.90 |
| | SensorAzimuth | -180–180 | -119.9–83.4 |
| MYD09GA | SensorZenith_1 | 0–65.57 | 2.93–55.48 |
| | SensorAzimuth_1 | -180–180 | -110.6–80.4 |
| MOD09GA | SensorZenith_1 | 0.4–65.62 | 2.95–53.46 |
| | SensorAzimuth_1 | -180–180 | -102.0–105.6 |
| MYD09GQ | sur_refl_b02_1 | -100–16000 | -100–7231 |
| MOD09GQ | sur_refl_b02_1 | -100–16000 | -100–5971 |

Table 2.1: Data sources: the fields Latitude and Longitude are the coordinates of the centers of the observation cells (in degrees). The fields SensorZenith and SensorZenith_1 are the view zenith angle or VZA (in degrees), The fields SensorAzimuth and SensorAzimuth_1 represent the sensor azimuth angle at each observation cell. The field SensorAzimuth_1 is listed for completeness since only SensorAzimuth is used in the analysis. The field sur_refl_b02_1 is essentially the surface reflectance at each grid pixel (multiplied by 10000). The ranges in the table correspond to the screened scenes for the years 2007 and 2008.

The study area is known to be persistently cloudy; therefore data were screened for observation quality including cloud and aerosol masking, resulting in the removal of most days from both sensors. Locations at latitudes greater than approximately 30° have multiple scans per day [Wolfe et al., 1998, 2002]. The high latitude of the study area (approximately 52.7° N) yields as many as 3 orbits on a given day. Although the daily composite (L2G-lite, discussed in Section 2.3) can include observations from several orbits, we only retain in our sample data acquired within a single orbit. This is enforced by restricting our analysis to dates where the extracted pixels (delimited in Fig. 2.4) have almost identical SensorZenith_1 values. The final sample group resulted in 62 (NS edge) and 58 (EW edge) observations from Aqua and 51 (NS edge) and 53 (EW edge) from Terra for an overall range of VZA from 2.93° to 55.48°.

The inspection of Table 2.1 shows that the highest available VZA for M09GQ within our study area is 55.48°. Since MOD09GA and MYD09GA use minimum VZA as one of the criteria for selecting observations, it is very unlikely to find VZA values close to the maximum (66°) for a particular location. This will limit our analysis to a range of view zenith angles from 2.93° to 55.48°, which contains, nonetheless, the vast majority of angles that users find in MODIS gridded reflectance products.

To obtain the estimate of the LSF along either rows or columns from daily samples of M09GQ products we use the techniques described in Section 2.4.1 applied to the sub-images delimited in Fig. 2.4. For the NS edge, our target is conveniently positioned in a near north-south direction, forming a small angle of 3.49° with the $y$-direction of the L2G grid. A shift is then applied to every row of the 23 rows by

20 columns corresponding sub-image to align the rows over the edge. For each date, we obtain a dense series of $20 \times 23 = 460$ aligned $x$-values and the corresponding reflectance values, which are used to derive a non-parametric estimation of the ESF and the corresponding parametric adjustment as described in Section 2.4.1 and illustrated in Fig. 2.3. For the EW edge the sub-image has 11 rows and 21 columns (see Fig. 2.4). We proceed similarly, although the sub pixel shift is very small ($\approx 0.33^{o}$). Nonetheless, we claim that the ESF estimates over the EW edge are still reliable (see discussion in Section 2.5.1).

### 2.4.3 Simulated values for observation cells and grid pixels

As described in Section 2.3, MODIS gridded products are resampled from swath observations. The cells are approximately rectangular and their GSD is a function of the view zenith angle. Therefore, we just need to know the geometry of acquisition, the location of the Level 2 grid and the sensor PSF to generate synthetic gridded data over some known target. Since our results for Aqua and Terra don't differ significantly (see Section 2.5) we only consider data from Aqua in this analysis.

The simulation is done in three steps. The goal of the first step is to determine the footprint of each nominal 250m observation cell. First, we select the orbit from which the (gridded) MYD09GQ product is derived. This is done by choosing the orbit which, for the coordinates of our study area, has the same range of values for the SensorZenith field of MYD03 and the SensorZenith_1 field of MYD09GA. Then we extract from the selected MYD03 data set the precise location of each nominal 1km

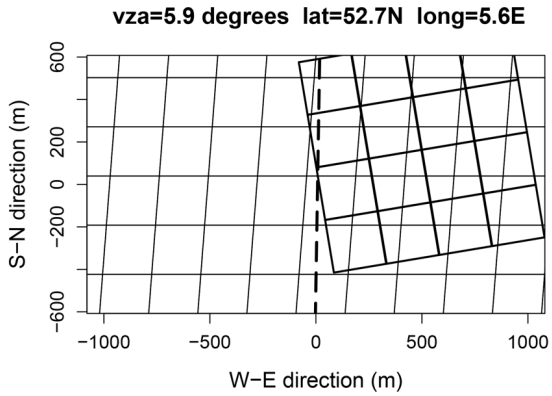



Figure 2.5: Geometry of the observations cells derived from MYD03 data over a portion of the NS edge for day 2007091 (yyyyddd) in relation to the SIN grid. The array of 4 by 4 cells represent a set of nominal 250m observations cells derived from one nominal 1km cell of MYD03, and the background grid represents the SIN grid. The dashed line represents the location of the edge. The coordinates are relative to the middle point of the edge in the study area (Lat = 52.697°N, Long = 5.593°E) . The array of observation cells which is depicted has sensor azimuth angle of 80.52°.

observation cell within the study area. The along scan direction for each nominal 1km cell is obtained from the MYD03 SensorAzimuth value. From the cell center and the values of SensorZenith and SensorAzimuth we determine the footprint of the 1km cell as in [Tan et al., 2006, Appendix B] (those authors use the scan angle as input, which we derive from SensorZenith). Finally, we determine the approximate observation cell footprint for swath data with nominal 250m resolution by generating a regular array of 16 nominal 250m cells that matches the footprint of the nominal 1km cell. This is illustrated by Fig. 2.5 where we represent the SIN grid and a subset of the observation cells obtained as described above.

The second step of the simulation is the 2-D convolution of the signal of the target with the sensor PSF for each observation cell. An *ideal* signal would be a signal which is not contaminated by atmospheric effects or other factors. For our study area, it would be the response over the perfect edge that separates a uniform feature

with value 0 on the water side and a uniform feature with value 1 on the land side. However, due to the geometry of acquisition of MODIS images, such ideal signal still suffers some blurring which increases with the GSD. To generate the response of the imaging system for the ideal signal, i.e. to obtain a simulated reflectance value between 0 and 1, we use the sensor PSF described in Section 2.3 and compute for each observation cell the 2-D convolution of the sensor PSF with the ideal signal. This is illustrated in Fig. 2.6, which background is still the SIN grid but depicts the PSF support, rather than the footprint, of a subset of observation cells for two dates over the NS edge and one date over the EW edge. In particular Fig. 2.6 (top two figures) describes how the signal acquired by the sensor depends on the VZA.

The third and last step is the resampling of the observation cell values into the SIN grid using the maximum obscov resampling criteria discussed in Section 2.3, which determines the values of the Level 2 grid pixels. We illustrate this construction also with Fig. 2.6, where it is noticeable that an increased VZA causes a larger blurring of the edge in the gridded image, depicted by the more heterogeneous distribution of the grid values around the edge in the middle image than on the top image. The bottom image in Fig. 2.6 shows that even if the grid is almost perfectly aligned with the EW edge, the grid values still exhibit a fair amount of variability around the edge. In particular, the observation cells centered near the edge have a response close to 0.5 which is assigned by the maximum obscov rule to grid pixels on either side of the edge. This artifact, which is due to geometry of observation and resampling, explains the blurring of the edge in the gridded product.

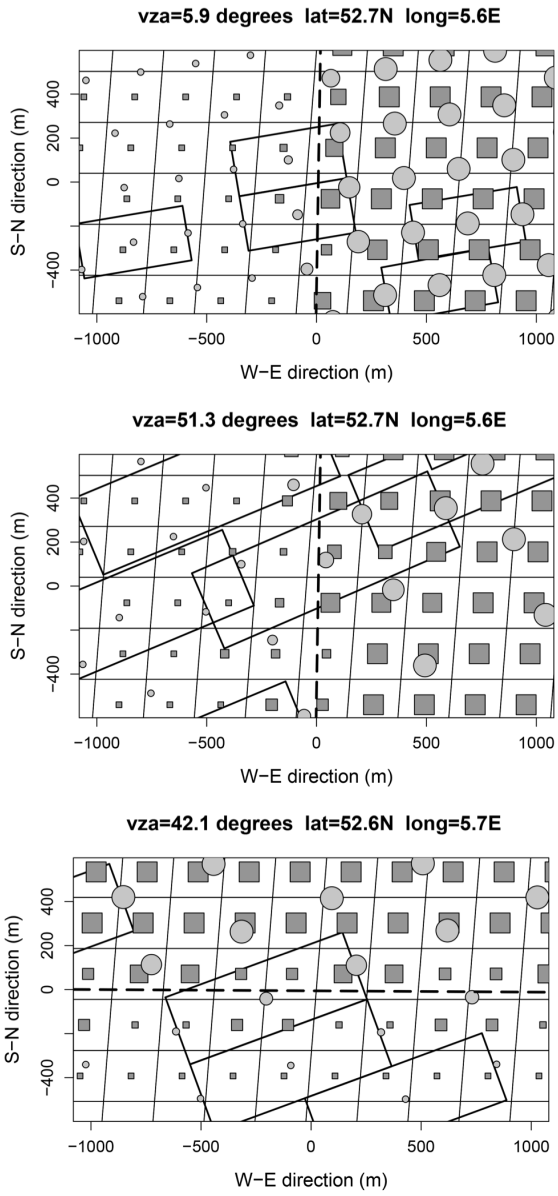



Figure 2.6: Simulation of observation cell and gridded reflectance values based on the location of the observation cell centers and the target. The SIN grid is in the background; the rectangles represent the PSF support for a subset of observation cells and the dashed line represents the position of the target edge between water (on the left) and agricultural fields (on the right). The response at each observation cell is represented by a gray circle; the value at each grid pixel derived from the maximum obscov criteria is represented by a gray square. The two top figures depict a portion of the NS edge respectively for day 2007091 (yyyyddd) with VZA $\approx 5.9^{o}$, as in Fig. 2.5, and day 2007074, with VZA $\approx 51^{o}$. The lower figure depicts a portion of the EW edge (day 2007216 with VZA $\approx 42^{o}$).

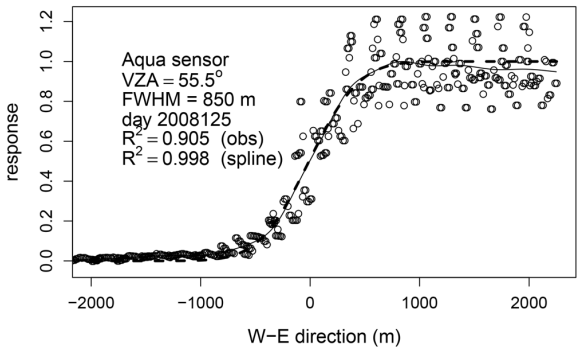



Figure 2.7: Estimation of the along row edge spread function ($\text{ESF}_{\text{row}}$) for MYD09GQ data with VZA = 55.64°. The dots represent the reflectance values along the aligned $x$ values and the solid line is the estimated non-parametric ESF. All values are normalized so the ESF ranges from 0 to 1. The dotted line is the Gaussian ESF adjusted to the data. The goodness of fit of the Gaussian ESF with the non-parametric ESF obtained by adjusting a spline is indicated by "spline"; the $R^2$ between the Gaussian ESF and the aligned observations is indicated by "obs".

## 2.5 Results

### 2.5.1 Line spread function estimation

Given a clip of L2G reflectance (from actual M09GQ products or from our simulations) over an edge we can estimate the non-parametric and the parametric ESF as discussed in Section 2.4.1. Fig. 2.3 illustrates the estimation for a small VZA. For comparison, we include Fig. 2.7 that is also based on MYD09GQ data but for high VZA. The goodness of fit is similar to Fig. 2.3 but the estimated FWHM is clearly higher as expected. These two examples show that the estimation method is robust for different geometry of acquisitions of MODIS products.

For each combination of sensor, edge, and date, we adjust a non-parametric ESF and all candidate parametric ESF described in Section 2.4.1 and compute the goodness of fit as illustrated in Fig. 2.3. We conclude that the best fit is always

| Edge | Data | average $R^2$ (spline) | average $R^2$ (obs) |
|---|---|---|---|
| NS | Aqua (simulated) | 0.998 | 0.870 |
| | Aqua (MYD09GQ) | 0.998 | 0.826 |
| | Terra (MOD09GQ) | 0.994 | 0.812 |
| EW | Aqua (simulated) | 0.998 | 0.812 |
| | Aqua (MYD09GQ) | 0.996 | 0.779 |
| | Terra (MOD09GQ) | 0.996 | 0.772 |

Table 2.2: Goodness of fit of the Gaussian ESF to the data, measured by the average coefficient of determination relative to the non-parametric ESF (spline) and to the aligned observations (obs) for each edge and each data set. Each data set includes all observations for the sampled dates.

obtained with the Gaussian ESF and the similarly shaped piecewise quadratic ESF (see Section 2.4.1). This suggests that the triangular LSF is always better than any combination of triangular LSF and rectangular LSF to model the sampled image system, even in the along column direction. This is an indication that the combined effect of the acquisition geometry and the sampling procedure (see Fig. 2.6) doesn't preserve the sharp rectangular along track response of the imaging system.

As discussed in Section 2.4.1, we model $\text{LSF}_{\text{row}}$ and $\text{LSF}_{\text{col}}$ as Gaussian in the rest of the paper. Table 2.2 lists the goodness of fit measured by the coefficient of determination for the Gaussian ESF for both sensors and edges and show that the Gaussian ESF is essentially as good as the non-parametric model to represent the ESF.

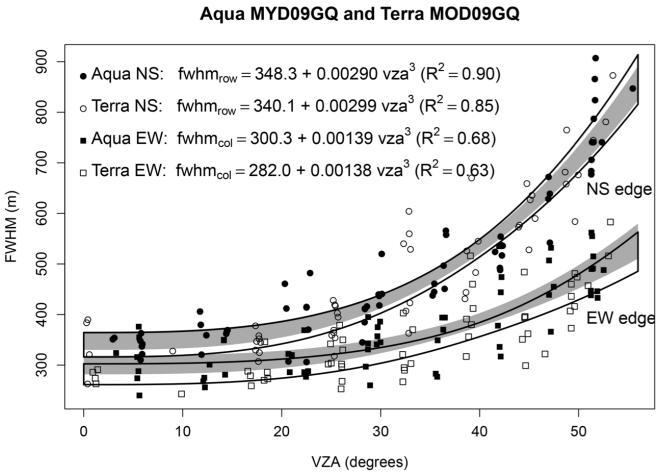



Figure 2.8: Relations between the FWHM and the VZA estimated from actual MODIS gridded surface reflectance products for the years 2007 and 2008 over the NS and the EW edge targets. The closed areas represent 95% confidence bands for the mean value of the FWHM. The gray areas are for Aqua and the hollow areas are for Terra. The estimated FWHM values are represented by circles for the NS edge and by squares for the EW edge, which are filled for Aqua and hollow for Terra. The equations indicate the regression between FWHM and VZA$^3$ for each combination of sensor and edge and are followed by the respective $R^2$.

### 2.5.2 Estimation of the relation between FWHM and VZA

In this section we use the results described earlier and we model the relation between FWHM and VZA. We also compare the actual MYD09GQ with our simulated data set. Fig. 2.8 describes the relation for M09GQ data for both sensors and edges. We adjusted several models to the pairs $\{(\text{VZA}_i, \text{FWHM}_i)\}$ and concluded that a cubic relation is statistically as good or better than other models with more parameters. The confidence intervals depicted in Fig. 2.8 show that there are no significant differences between sensors but there are clear differences between edges. This means that $\text{LSF}_{\text{row}}$ is consistently smoother than $\text{LSF}_{\text{col}}$ for the MODIS gridded products we considered.

As mentioned in Section 2.4.2, a matter of concern is the small sub-pixel shift

used to align the observations about the EW edge. To evaluate its impact on our results, we ran our simulations over an artificial EW edge, which is defined by a rotation of 2.29° of the actual EW edge, in order to optimize the ESF estimation. Fig. 2.9 show that the FWHM estimates for the simulations over the artificial and the actual EW edge are almost coincident. Therefore we conclude that the FWHM estimates for the EW edge are reliable. Since there are not significant differences between sensors, we determine a single relation between FWHM and VZA for the NS edge and another for the EW edge, which we indicate in Fig. 2.9 for the combined M09GQ data set and that we compare with the data set that was obtained through simulation. Some conclusions can be drawn from the inspection of Fig. 2.9. First, there is a significant difference in the mean value of FWHM between the combined M09GQ data set and the simulated data set, for most of the range of VZA values. This is due to noise factors (e.g. atmospheric effects) that alter the signal prior to the sensor acquisition and that we don't incorporate when we generate the synthetic data set. Nonetheless, we can see that the relations for each edge are qualitatively similar, which is further evidence that our estimation techniques are robust. The major concern is that the irregularities of the target edge we consider in this study are affecting the estimation of the LSF. However, Fig. 2.9 shows that the differences between the M09GQ and the synthetic data set grow progressively with the VZA from 30m to 130m for the NS edge, and from 6m to 90m for the EW edge. If the difference was due a lack of definition of the target edge, then it should be essentially constant since it would not depend on the VZA. This suggests that the reported difference is mostly related to atmospheric effects.

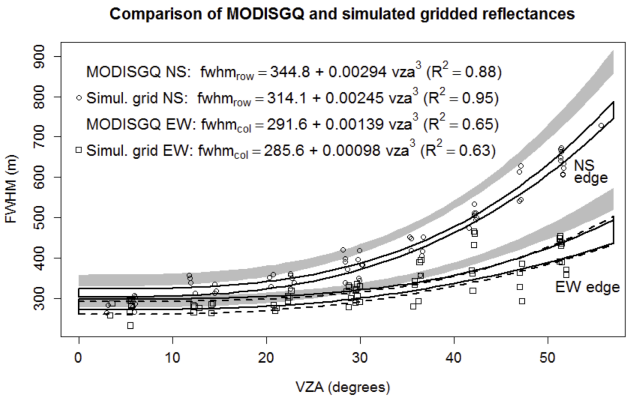



Figure 2.9: Relations between the FWHM and the VZA estimated from the combined M09GQ data set and from the simulated MODIS gridded surface reflectance products over the NS and the EW edge targets. The five curved shapes represent 95% confidence bands for the mean value of the FWHM. The light gray areas are the confidence bands for the FWHM derived from the M09GQ datasets described in Fig. 2.8. The three hollow bands are derived from simulated data over the actual NS edge (top), the actual EW edge (bottom hollow band with a solid border), and the artificial EW edge (hollow band bordered by a dashed line). The fact that both bottom hollow bands are almost superimposed indicates that the actual EW edge can be used to obtain reliable estimates of FWHM in spite of its small sub-pixel shift. The estimated FWHM values for the simulated data are represented by circles for the NS edge and by squares for the EW edge. The equations indicate the regression between FWHM and $\text{VZA}^3$ for each data set, for the NS edge and for the (actual) EW edge.

We use Equation 2.2, where the coefficients $a = \text{FWHM}_{\text{row}}/2.355$ and $b = \text{FWHM}_{\text{col}}/2.355$ are replaced by their dependence on VZA as estimated from the combined M09GQ data set, i.e.

$$\begin{aligned} \text{FWHM}_{\text{row}} &= 344.8 + 0.00294\,\text{VZA}^3 \\ \text{FWHM}_{\text{col}} &= 291.6 + 0.00139\,\text{VZA}^3 \end{aligned} \tag{2.3}$$

to determine the area on the ground that contributes the most to the value at any grid pixel according to our estimation of the sampled image system PSF. Although the minimum size ground area would be elliptical, we compute a rectangular area which is easier to describe. Table 2.3 indicates, for a range of VZA between $0^{\text{o}}$ and $50^{\text{o}}$, the dimensions of rectangles, centered at the grid pixel, that contribute to 50%,

| | | VZA (degrees) | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0 | 10 | 20 | 30 | 40 | 50 |
| 50% | along rows | 308 | 310 | 329 | 379 | 476 | 637 |
| | along columns | 261 | 262 | 270 | 294 | 340 | 416 |
| 75% | along rows | 438 | 442 | 468 | 540 | 679 | 908 |
| | along columns | 371 | 373 | 385 | 419 | 484 | 592 |
| 90% | along rows | 570 | 575 | 609 | 702 | 882 | 1180 |
| | along columns | 483 | 485 | 501 | 545 | 630 | 770 |

Table 2.3: Estimated dimensions (in meters) of an rectangular area on the ground that contributes to 50%, 75% or 90% of the signal of a 250m nominal gridded pixel as a function of the VZA. These values are derived from the combined M09GQ data set depicted in Fig. 2.8 for the study area (Lat = 52.697°N, Long = 5.593°E).

75% or 90% of the signal at that pixel.

We are now in measure to discuss the decomposition of M09GQ PSF described in Equation 2.1. Instead of estimating each component of the PSF, we follow a coarser approach that consists in decomposing the FWHM for M09GQ given by Equation 2.3 into three additive terms: one for the imaging system, one for the sampling procedure, and a third term for the additional noise factors (including atmospheric effects). The GSD, which depends solely on the VZA (see [Tan et al., 2006, Appendix B]), measures the contribution of the imaging system, the FWHM estimated from simulated gridded reflectance accounts for the both the imaging system and the sampling procedure. The additional noise factors correspond then to the difference

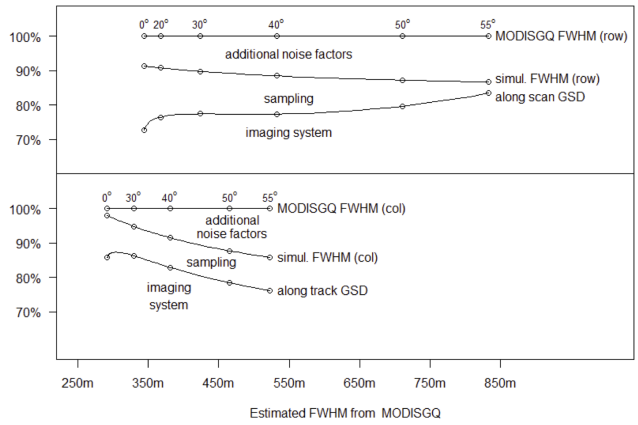



Figure 2.10: Approximate contributions of the imaging system, the sampling procedure and additional noise factors (including atmospheric effects) to the estimated FWHM of M09GQ. The figure shows the decomposition of $\text{FWHM}_{\text{row}}$ (top) and $\text{FWHM}_{\text{col}}$ (bottom). The lower curve indicates the contribution (in %) of the imaging system and the middle curve the cumulative contribution of the both imaging system and sampling procedure. The dots along the curves correspond to the printed VZA values.

between the mean FWHM estimates for M09GQ and simulated datasets. Fig. 2.10 reveals differences in the decomposition of $\text{FWHM}_{\text{row}}$, which combines high VZA and high GSD, and $\text{FWHM}_{\text{col}}$ which combines high VZA with moderate GSD. The figure suggests that the effect of additional noise factors increases with the VZA, but that there is a reduction of the sampling procedure effect for high GSD; this can be explained by the fact that as GSD becomes increasingly larger with respect to the fixed grid size, the contribution of the sampling procedure in the whole imaging system becomes progressively attenuated.

### 2.5.3 Simulations for world wide locations

As discussed in Section 2.3, most MODIS gridded reflectance products are assigned to the SIN grid which intersects the along track direction of the sensor at variable angles (see Fig. 2.2). Since our analysis relies on the combined geometry of

the observation cells and the L2G pixels, one can ask how our previous results apply to areas other than our study area in the Netherlands. To address this question we generalize the simulation described in Section 2.4.3 to an arbitrary location. We replicate the Aqua dataset used for the study area (see Section 2.4.2) but we introduce a correction for the latitude. Since the sensor track direction varies with the latitude (see Fig. 2.2), we rotate the positions of the cell centers and the azimuth angle accordingly. Hence, the rotation is counterclockwise for latitudes higher than 52.7° (the study area latitude) and clockwise otherwise. The L2G geometry can be derived for any location by converting its latitude/longitude into the corresponding row/column of the SIN grid as described in [Giglio, 2008, p.24]. For each location, we simulate the gridded product over two ideal edges which are aligned – up to a small rotation of 2.86° in order to optimize the sub-pixel shift – either with the SIN grid rows or columns. This allow us to estimate the $\text{FWHM}_{\text{row}}$ and the $\text{FWHM}_{\text{col}}$ for the location of interest as a function of the VZA using the approach described in Section 2.4.3.

We illustrate the rotational effects that the combination of track direction and grid geometry create with two opposite examples: Fig. 2.11 (top) shows the result of the simulation with a high VZA = 51.3° at the Earth Observation System (EOS) core validation site [Morisette et al., 2002] of Changbaishan in China, where the rotational effects between track direction and SIN grid are pronounced. In that case, the along scan direction is 67.4°NE while the grid columns are oriented 56.44°NE, leading to a 10.96° angle between those two directions. In a clear contrast with our study area (see Fig. 2.6) the along scan direction is almost aligned with the *columns* of the SIN

grid. The opposite example corresponds to the EOS site of Barrow, Alaska, where the track direction and the columns of the grid are now much more aligned as shown in Fig 2.11 (bottom). The simulation is performed over the same dataset (after the correction for latitude described earlier) and shows that for the same sensor PSF support, the signal of the edge suffers much more blurring. As a matter of fact, while the simulation represented by Fig. 2.11 (top) leads to an estimated $\text{FWHM}_{\text{row}}$ of 594m, the one for the bottom figure leads to a 1233m value for $\text{FWHM}_{\text{row}}$.

We perform the above simulation for 145 locations regularly distributed between latitudes below 72°, and the whole range of longitudes, and for an additional 29 EOS site locations well distributed around the globe to estimate the dependences of both $\text{FWHM}_{\text{row}}$ and $\text{FWHM}_{\text{col}}$ on the VZA using the techniques described in Section 2.4.1. Fig. 2.12 shows some contour lines of the estimated $\text{FWHM}_{\text{row}}$ at nadir (VZA = 0°) and Table 2.4 indicates the estimated values of $\text{FWHM}_{\text{row}}$ and $\text{FWHM}_{\text{col}}$ at VZA = 0° and VZA = 55° for a subset of EOS locations around the globe. The inspection of Fig. 2.12 shows that the along row resolution is very much dependent on the latitude, mainly for higher longitudes, while the along column resolution is much more stable. This is the result from the geometry of the SIN grid, where the rows are always along the parallels but the direction of the columns vary as described by Fig. 2.2. Moreover, Fig. 2.12 shows that there is a asymmetry in the degradation of the along row resolution, which is related to the track direction of the sensor. Therefore, it is important stress that Fig. 2.12 illustrates the results obtained from simulations over daily Aqua data sets, although it is reasonable to admit that simulations based on Terra data would be similar, but inverted with

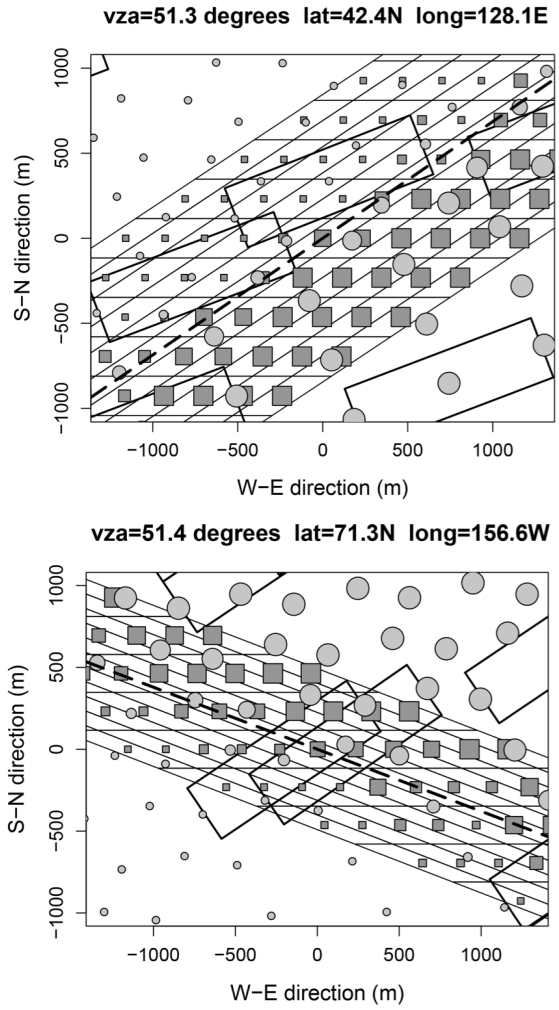



Figure 2.11: Simulated values for Aqua observation cells and SIN grid pixels over a perfect edge (depicted as the dashed line) aligned with the SIN grid columns (represented by the thin solid lines in the background). A subset of the PSF supports is depicted (wide rectangles). The circles are located at the centers of the observation cells; their size indicates the estimated response over the edge and ranges from 0 (to the left of the edge) to 1. The squares are located at the centers of the SIN grid pixels; their size indicates the estimated grid pixel value which is derived from the observation cell values by the maximum obscov criteria. The top figure corresponds to the EOS Changbaishan site location (Lat=42.4025°N and Long=128.0958°E) and the bottom figure corresponds to the EOS Barrow site location in Alaska (Lat=71.281°N and Long=156.612°W).

respect to the prime meridian, due to the symmetry of the orbits.

| Site | Lat | Long | FWHM$_{row}$ (m) | | FWHM$_{col}$ (m) | |
|---|---|---|---|---|---|---|
| | | | nadir | 55$^o$ | nadir | 55$^o$ |
| Ak | 71.28 N | 156.61 W | 631 | 1363 | 281 | 475 |
| Or | 44.43 N | 121.56 W | 460 | 1169 | 279 | 494 |
| Me | 45.20 N | 68.73 W | 390 | 935 | 279 | 469 |
| Br | 10.08 S | 61.93 W | 321 | 720 | 281 | 475 |
| An | 52.61 N | 0.52 E | 319 | 746 | 284 | 424 |
| St | 59.80 N | 30.80 E | 317 | 690 | 284 | 435 |
| Si | 57.27 N | 91.60 E | 415 | 733 | 285 | 407 |
| Ch | 42.40 N | 128.09 E | 446 | 757 | 285 | 436 |
| Zb | 15.43 S | 23.25 E | 326 | 752 | 288 | 466 |
| Au | 34.39 S | 145.30 E | 451 | 1153 | 280 | 501 |

Table 2.4: Simulations based on Aqua data sets for 10 EOS site locations: Ak (Barrow), Or (Cascades), Me (Howland), Br (Ji-Parana), An (Barton Bendish), St (St. Peterberg), Si (Krasnoyarsk), Ch (Changbaishan), Zb (Mongu), and Au (Uardry). The listed values are the estimated FWHM$_{row}$ and FWHM$_{col}$ at nadir and at VZA = 55$^o$ for each location.

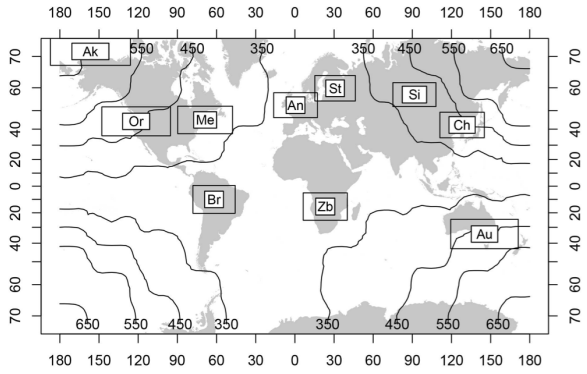



Figure 2.12: Simulations based on Aqua data sets for 174 locations distributed over the globe. The countour lines represent the estimated $\text{FWHM}_{\text{row}}$ (in meters) at nadir (VZA = 0°). Each box represents the location of an EOS site. The inner and outer boxes dimensions for each location are proportional to $\text{FWHM}_{\text{row}}$ and $\text{FWHM}_{\text{col}}$ at nadir and at VZA = 55°; their exact values are described in Table 2.4. The map projection is as in Fig 2.2.

## 2.6 Conclusions

The main goal of this paper is to characterize the spatial resolution for the MODIS L2G surface reflectance product incorporating the known properties of the variable viewing geometries. We have shown that the FWHM is a reasonable estimate of resolution for gridded products over the range of VZA for MODIS. Furthermore, based on the estimation of the along row and along column FWHM, we propose a complete description of the overall sample image system PSF.

The results in this paper rely on two main assumptions. Firstly, we consider that the anisotropic Gaussian distribution is a good model to represent the PSF of the gridded products, as long as the support of the PSF is allowed to vary. We found that this simple model for the PSF fits very well with both actual MODIS L2G products and MODIS synthetic products over an ideal target. The second assumption is that our study area is a suitable target for the edge method with sufficient resolution for

MODIS. In addition to the very sharp geometry of the target, illustrated in Fig. 2.4, our results with real data match nicely with our synthetic results, which we consider an additional evidence of the applicability of our techniques.

The results obtained in this paper indicate a resolution near nadir coarser than the reported nominal GSD of 250m. The pixel offsets and noise shown around a stable, simulated edge target allow us to evaluate how sampling and geometric effects from the MODIS instrument contribute to image degradation, as illustrated in Fig. 2.6 and Fig. 2.11 that describe the artifacts that can occur when the grid is aligned with the target.

The reported results assert the importance of distinguishing the imaging system, associated to the "swath" data, from the sampling procedure that leads to the gridded products. While the former is beyond the control of the final user, the latter is somewhat arbitrary, since it depends on the chosen projection and resampling criteria. Our results show that the projection associated to the MODIS SIN grid causes a degradation of the resolution for locations away from the Equator and the central meridian of the projection, as illustrated by Fig. 2.12. The study area in the Netherlands happens to be situated near the central meridian of the SIN grid and, therefore, our results described by Fig. 2.8 and Table 2.3 give a good indication of the best possible resolution that can be obtained for 250m nominal gridded products, with a suitable projection. These findings support the idea that users should re-project the data depending on the location of interest. Tools like LPDAAC MODIS Reprojection Tool Swath or NSIDC MODIS Swath-to-Grid Toolbox are readily available to re-project swath data (e.g. level 2 MOD09 or MYD09) into

a grid with minimal distortion for a given location. Those gridded data could then be processed to obtain higher level (level 3 or level 4) MODIS products, although this last step would require a much greater level of expertise from the user.

There are several open questions that could be addressed in future work. One question is related to the resampling criteria since L2G products depend on the selection of both grid and resampling criteria. It would be interesting to analyze how a criteria other than maximum obscov (e.g. bilinear resampling) would affect the resolution of gridded products. Another direction to explore is related to the fact that our study was only focused on daily MODIS products. Most higher order MODIS products are temporal composites created from the best quality pixels (e.g. cloud free, low aerosol, or minimum VZA) over a composite period (e.g. 8 or 16 days). The outcome of the composite product is a scene often comprised of observations acquired from multiple days and geometries [Solano et al., 2010]. The scenes over persistently cloudy regions, e.g. the Amazon or our own study area, must frequently use off-nadir observations but near-nadir observations are prevalent in areas with little clouds. While our results for VZA $= 0^{o}$ provide lower bounds for the resolution of higher order multi-day products, more research is needed to assess the spatial resolution of temporal composites in general.

The results presented in this paper establish the methods to account for the effective spatial resolution of geolocated grid data from wide field-of-view instruments. The VIIRS instrument will provide data with a more uniform along scan GSD than MODIS does currently for varying VZA. The methods used in this paper will be used to assess the gridded products generated from these new data as soon as a

representative dataset is available. The results from this future analysis is essential to evaluate continuity between the MODIS and VIIRS sensors for long-term data records.

# Chapter 3: Demonstration of Uncertainty Resulting from MODIS-Like Geometries: an Albedo Case Study

## 3.1 Abstract

This research examines uncertainty in MODerate Resolution Imaging Spectroradiometer (MODIS) observations, and demonstrates the direct influence of geometric distortions resulting from the standard practice of geolocating swath observations. MODIS observations vary dependent on the Ground Sample Distance (GSD), which varies dependent on the View Zenith Angle (VZA) that changes with every orbit. MODIS Level 2G (L2G) land products are generated by applying a geolocation algorithm that resamples the variable observation geometries to a consistent grid of fixed pixel size and location, a process which itself introduces variability associated with the changing observational footprint. For this study, broadband albedo was simulated for five validation sites, representing five distinct land cover types, exhibiting quantifiable variability, with additional seasonal variability exhibited in some sites. All site simulations exhibit compounded uncertainty attributable to the geometric distortion sufficient to influence climate models (i.e. ranges from 0.01-0.045 albedo). These results indicate there is a minimum level of uncertainty

associated with the variable geometry that should be factored into L2G-based products, particularly for nominal 250 meter band data. Aggregating the data to coarser resolutions and smoothing the data through average resampling can mitigate the uncertainty.

## 3.2 Introduction

Coarse resolution sensors (4km-250m) commonly used for land remote sensing, such as the NOAA AVHRR, SPOT VGT, MODIS have broad swath-widths (2,200–2,600km). The resulting off-nadir viewing allows for near-daily viewing, increasing the opportunity for cloud-free observations [Townshend and Justice, 2002]. However the off-nadir viewing results in variable geometries, which needs to be addressed in data pre-processing and production, and provides an inherent uncertainty in the resulting products. For this study, uncertainty is examined with respect to a hypothetical albedo product from MODIS. Albedo was selected as a physical quantity measurable by satellite with importance to the climate change community.

Albedo is the fraction of incident solar irradiance reflected by the Earth's surface across the solar spectrum [Goward, 2005], measured as a ratio of upwelling to downwelling radiance, providing estimates of the energy budget available for driving heat fluxes in climate system modeling. Radiative forcing is the change imposed to Earth's radiation balance as a result of some perturbation of the surface [Hansen et al., 1997], measured in Watts per meter squared ($Wm^{-2}$), providing a useful measure for positive (warming) and negative (cooling) changes, and can be

derived from measurements of albedo. The relationship of albedo in a basic net radiation budget model is:

$$R_{net} = (1 - albedo) * E_{solar} + \varepsilon_s * E_{lwd(Ta)} - \varepsilon_s * M_{lwu(Ts)}, \tag{3.1}$$

where *albedo* is the surface albedo, $E_{solar}$ is directly absorbed solar radiation, $\varepsilon_s$ is surface emissivity, $E_{lwd(Ta)}$ is downwelling longwave atmospheric emission ($Ta$ = air temperature), and $M_{lwu(Ts)}$ is upwelling longwave surface emission ($Ts$ = surface temperature).

Monteith and Unsworth [2008] demonstrate the linear relationship between irradiance and air temperature for both direct and diffuse radiative scattering models; multiple studies have either included, or modeled, their climate system response to changing land albedos with respect to direct temperature effects and carbon sequestration [Dixon et al., 2003, Montenegro et al., 2009, Schurgers et al., 2008, Waggoner and Sokolik, 2010].

The role and importance of albedo in Earth Systems Science has been established since the early days of remote sensing [Sagan et al., 1979], and is now key to the ongoing discussion regarding climate change and anthropogenic effects upon Earth systems [Goward, 2005, Lacis et al., 2010]. Recorded albedo estimates derived from satellite observations contribute to our understanding of changes in the climate, through a relation to the net radiative flux over different landscapes. The absolute error for satisfactory albedo parameters in climate models ranges from ±0.05 to ±0.02 [Oleson et al., 2003]. As climate models continue to improve their spatial

and temporal resolution and incorporate new data sources [Pachauri and Reisinger, 2007], identifying and minimizing model error becomes increasingly necessary.

Two studies, each using simulated cropland land cover class albedos, exemplify the dependency between albedo accuracy and model uncertainty. Myhre and Myhre [2003] ran sequential models using albedo values from 0.15, 0.18 and 0.20 to simulate baseline albedo scenarios representative of the potential natural vegetation (PNV); the resultant radiative forcing was calculated at $-0.06$, $-0.20$ and $-0.29$ $\mathrm{Wm}^{-2}$ cooling, respectively. Matthews et al. [2003] ran models using albedo values between 0.17 and 0.20 to simulate historical baseline land covers from the 1700's CE to 1992, demonstrating increased radiative forcing from $-0.15$ to $-0.28$ $\mathrm{Wm}^{-2}$, calculated as a difference in cooling of $-0.09\,^{\circ}\mathrm{C}$ and $-0.17\,^{\circ}\mathrm{C}$ respectively. Both studies conclude that model outcomes are dependent on small changes in albedo parameters and describe a general relationship where $+0.01$ albedo corresponds to $-0.05 Wm^{-2}$, and $-0.03^{\circ}\mathrm{C}$ cooling.

The MODerate Resolution Imaging Spectroradiometer (MODIS), flown on two polar orbiting platforms, has provided useful information regarding planetary albedo since the early 2000's [Schaaf et al., 2002]. The Fifth Edition of the Intergovernmental Panel on Climate Change (IPCC) describes the uncertainties of the albedo parameters used in global climate models [Myhre et al., 2013], exhibiting the sensitivity of model runs to changes in albedo parameters, and demonstrating the dependency of model outcomes on the accuracy of albedo parameters, some of which are derived from MODIS. Calculations of albedo are sensitive to observational characteristics, such as atmosphere and land cover. For example, the Myhre [2009] albedo

study demonstrates diminished negative forcing, $-0.8$ to $-0.65$ $\mathrm{Wm}^{-2}$, when atmospheric models incorporate aerosol optical depth (AOD) from MODIS observations. Liang et al. [1999] report albedo uncertainty exceeding $\pm 0.05$ attributed to sources including changing AOD conditions, variable zenith angles, and complex landscapes.

Given the sensitivity of climate models to small changes in albedo, this paper simulates a series of MODIS-like gridded broadband albedo for several land cover types to test the measurement consistency of a given gridded pixel through changing view geometries. The following section describes the MODIS instrument characteristics and known sources of error. Methods for simulating MODIS geometries and a simulated albedo time series are discussed in Section 3.4. Results for the various simulations are discussed in Section 3.5. Methods useful in mitigating the error are explored in Section 3.5.5. Finally, conclusions and recommendations are presented in Section 3.6.

## 3.3 MODIS Observation Characteristics

MODIS is a whiskbroom system with a scanning mirror providing a wide observation swath (2330km) for 36 spectral bands at nominal detector pixel resolutions of 250m, 500m, and 1000m [Barnes et al., 1998]. The wide field-of-view ($\pm 66°$) and high temporal frequency (at least 2 passes a day at the equator) of MODIS provides near daily global coverage [Salomonson et al., 2000, Xiong and Barnes, 2006], but comes at the cost of a loss in signal accuracy due to known effects of pixel size growth at increasing view zenith angles (VZA). VZA is the angle formed between a

line perpendicular to the observed surface (local zenith) and a line pointing back to the instrument.

The ground-projected sample interval (the spacing between pixels), known as the ground sample distance (GSD), is a common measure of instrument spatial resolution [Schowengerdt, 2007]. The reported values of 250m, 500m and 1000m correspond to the nominal GSD of the MODIS instrument at nadir and do not reflect the continuously varying VZA, the overall effect of which is a growth in the GSD of approximately 2.0 times along track and 4.8 times along scan at the maximum observation angle. The pixel growth ultimately leads to overlapping scans toward the edge of scan, known as the bow-tie effect. The constant scanning of the mirror forms a triangular line spread function (a sensor's one-dimensional spatial response to the measured physical signal) integrating radiometric signal from the immediate preceding and subsequent neighbor pixels (12.5%) to the center pixel (75%) along scan [Nishihama et al., 1997, Wolfe et al., 2002].

To help facilitate use, all higher order MODIS products are assigned to a geo-referenced grid using a custom sinusoidal projection with a fixed pixel size for each of the available nominal resolutions. MODIS geolocation assumes a fixed spatial size based on the subdivision of a 30 arc second grid providing a seamless global grid with resolutions of $\approx$ 231m, $\approx$ 463m, and $\approx$ 926m [Wolfe et al., 2002]. Observations made at resolution $r$ (e.g. 250m) are assigned to a co-located $(x, y)$ coordinate in the geolocation grid of the same resolution. Observational coverage ($obscov$) is the integrated signal response during a sample interval overlapping a grid cell divided by the full observation footprint, which accounts for the changing sample sizes at

increasing GSD. Grid pixels are assigned the value of those observations (swath pixels) with the highest *obscov* values [Wolfe et al., 1998, Yang and Wolfe, 2001]. However, the reported mean *obscov* for MODIS grid data is less than 30% [Tan et al., 2006], meaning the value assigned to any given grid pixel is only a portion of the spatial signal from the larger observation pixel, and not necessarily representative of the small geographic area defined by the grid location. Further, single observations are often assigned to more than one grid pixel at large VZA.

There is no specific problem with the assignment of single or partial values to multiple grid cells. Conversely, there is no benefit from producing a finer resolution product given the cell will only ever represent a portion of spatial signal present because the GSD of the instrument is not consistent and is always greater than the grid cell size. The portion of the spatial signal will also differ as a function of location, since the observation pixel center location and associated geometry will be different for each orbit between nadir-to-nadir views, making the defined grid space an arbitrary subdivision of the swath data.

The results from [Campagnolo and Montaño, 2014] introduced an inconsistent, location dependent effective GSD that is far different than the nominal grid resolutions suggested by product descriptions. The author concludes the effective resolution of MODIS 250m grid data varies between 344m and 835m along scan and behave in the same basic manner as swath data (i.e. the effective resolution grows in the same fashion as swath data).

Many MODIS products are generated as *discrete* temporal composites, using only the single best quality per-pixel observation (e.g. cloud-free and/or minimum

view zenith angles) for a given composite period - usually 8-16 days. Data in these composites ultimately suffer from geometric distortion in that a seemingly spatially continuous scene nonetheless represents a spatially discontinuous surface comprised of observations from any number of disparate GSD.

Figure 5.1 illustrates the different days of data and the associated VZA used in a typical discrete composite MOD13 NDVI product. The composite day of year seen in the figure is an interesting illustration of the false assumption of spatial continuity of data within a composite scene. There are 8 days worth of data in the scene that are not necessarily spatially contiguous. Further, and of particular import to this study, the range of VZA in the composite is far larger than one might expect. Specifically, the MOD13 product series applies a minimum VZA criterion in the selection of best quality pixels [Solano et al., 2010]. However, there are clearly circumstances overriding this criteria (e.g. a limited number of cloud-free observations) making the use of less desirable observation days (e.g. those with large VZA) a necessity.

Therefore, the goal of this paper is to demonstrate the level of measurement uncertainty imposed by the variable GSD of the MODIS instrument and geolocation process attributable to geometric distortion at different spatial scales using broadband albedo as a measure. The work is not a validation study for any current MODIS science product, and does not employ the same production methods or quality assessment as those products. Reference to any given published product accuracies is solely to compare the relative error contribution.

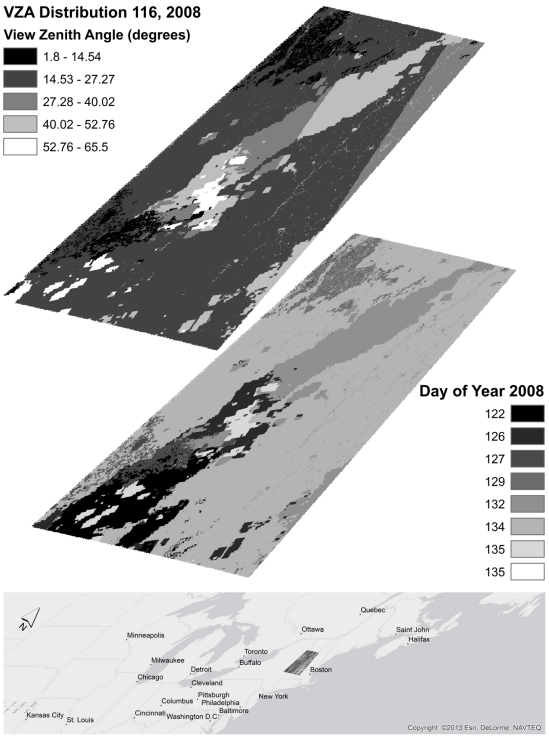



Figure 3.1: Sample illustration of days used in compositing the MODIS MYD13Q1 NDVI product from the April 27, 2008 (day 116) 16-day composite. The top figure shows the view zenith angles associated with the days used. The lower figure shows the day within the period that was used in the composite. The MOD13 product uses a minimum VZA as part of the quality selection, but nonetheless some large angles are used as a result of coincident factors such as persistent cloud contamination.

## 3.4 Data and Methods

Assessing the impact of geometric artifacts on measures of albedo requires isolation of the geometric influence from other potential sources of error. To accomplish this, synthetic time series of MODIS data were generated from 11 individual scenes of Landsat 5 TM images for five Earth Observation System (EOS) validation sites [Morisette et al., 2002], representing five distinct land cover types (Table 3.1).

A broadband albedo was calculated using a simple linear model albedo and methods for non-snow land cover described in Brest and Goward [Brest and Goward, 1987]. The broadband albedo method uses a linear combination of visible red and near infrared bands calibrated to top-of-atmosphere reflectance - originally bands 4 and 7 from Landsat 1-3 in the paper translated into bands 3 and 4 for the Landsat 5 data used here - that are represented in 250m MODIS bands 1 and 2. The method first uses a simple ratio (SR) of band 4 divided by band 3, where a ratio greater than 2 indicates vegetation versus non-vegetation. The broadband albedo is then calculated as follows:

$$albedo = \begin{cases} 0.526(B3) + 0.362(B4) + 0.112[0.5(B4)] & \text{if } SR > 2.0 \text{ (vegetation)} \\ 0.526(B3) + 0.474(B4) & \text{if } SR < 2.0 \text{ (non-vegetation)} \end{cases} \tag{3.2}$$

The difference in bandwidth between Landsat instruments and potential changes to the linear model was not a factor, nor was the lack of atmospheric correction, since absolute albedo model accuracy is not the goal of this paper. Albedo was

selected for this study as it is a physical quality that can be estimated from satellite instruments. Thus, the simple broadband model is sufficient to serve as a context from which geometric variability can be assessed.

The surface composition and spatial variability of the land cover at any site will influence the amount of spatial signal incorporated into any given GSD; a uniform, homogeneous canopy will present a more stable spatial signal over coarsening scales than a highly variable, heterogeneous landscape with small crop fields. Spatial variation further depends on seasonality for those canopies that experience different levels of canopy green-up and senescence. Therefore, seasonal profiles were generated for those sites expected to demonstrate seasonally variable phenology, Harvard Forest (broadleaf forest), Konza Prairie (grassland/cereal crop), and Bondville (broadleaf cropland). Images containing snow cover are beyond the scope of this analysis, and were excluded.

### 3.4.1 Study Areas

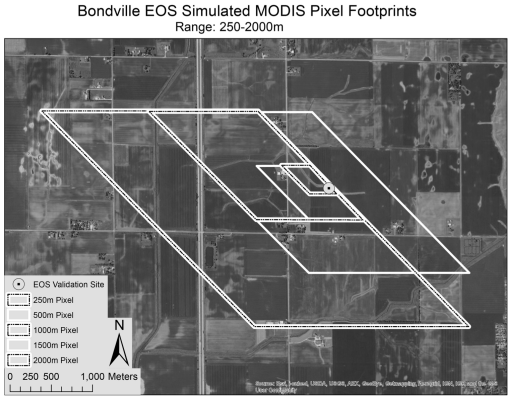
The Bondville, IL EOS agricultural site is located in the Midwest United States and is characterized by agricultural land with generally rectangular fields of agricultural land that can be in various states of development based on both time of year and crop type in each field. A high level of spatial heterogeneity can be expected from this site as a function of the GSD being larger than the field size from day-to-day. The pixel grid cells containing the site, seen in the Figure 3.2a, clearly demonstrate the MODIS signal incorporated at each given stage of spatial

Table 3.1: EOS Validation Sites and Landsat TM data used in the study

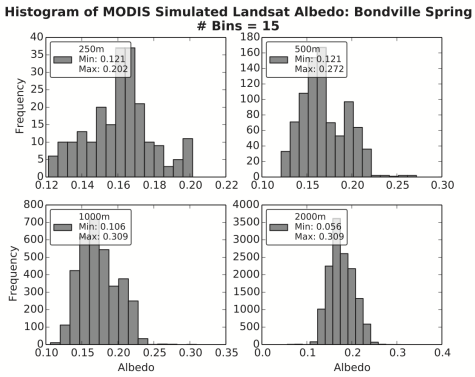
| Site Name | Lat | Lon | Path/Row | Date | Land Cover |
|---|---|---|---|---|---|
| Bondville, IL | 40.00 | -88.29 | 22/32 | 2010-04-14<br>2010-07-03<br>2010-11-08 | Broadleaf Cropland |
| Harvard Forest LTER, MA | 42.54 | -72.18 | 13/30 | 2008-04-26<br>2008-06-12<br>2006-10-13 | Broadleaf Forest |
| Konza Prairie LTER, KS | 39.08 | -96.56 | 28/33 | 2011-03-10<br>2011-07-01<br>2011-09-02 | Grassland/Cereal Crop |
| Jornada LTER, NM | 32.60 | -106.86 | 33/37 | 2011-04-30 | Shrubland/Woodland |
| Lake Tahoe, NV | 39.17 | -120.104 | 43/33 | 2009-09-21 | Needleleaf Forest/Clear Water |

aggregation.

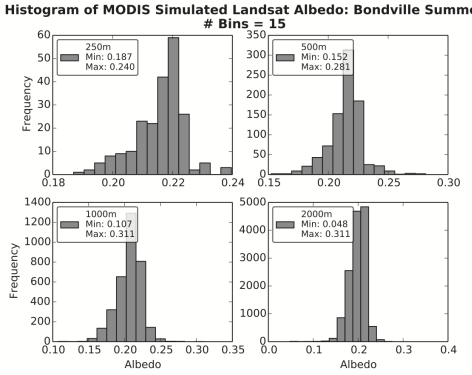
The histograms in Figure 3.2 are used to understand the level of spatial heterogeneity of the study sites at varying scales. Each group of histograms illustrate the distribution of Landsat albedo pixels centered around the site extending in a circle to 250m, 500m, 1000m, and 2000m and for each season. Non-normal distributions will incorporate varying amounts of outlier data as the MODIS GSD changes, causing irregularity in the simulated MODIS albedo; a normal distribution (i.e. spatially homogeneous) provides a more stable input signal for each simulated footprint, thus



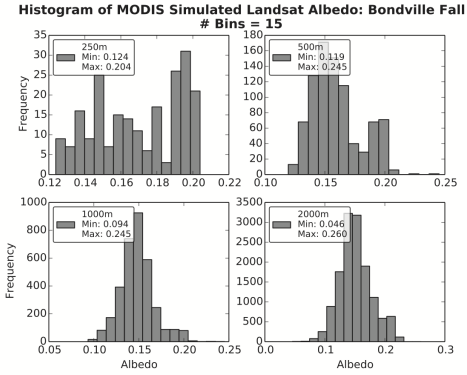

(a) Bondville, IL




(b) Spring albedo frequency




(c) Summer albedo frequency

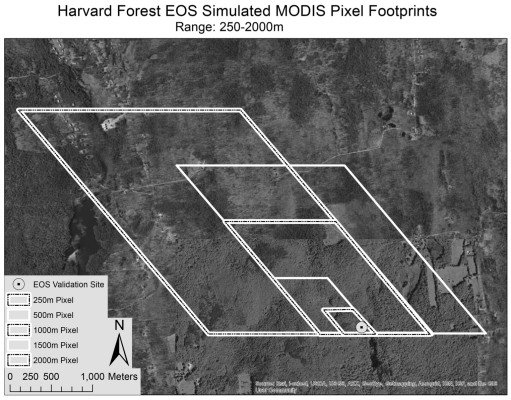



(d) Fall albedo frequency

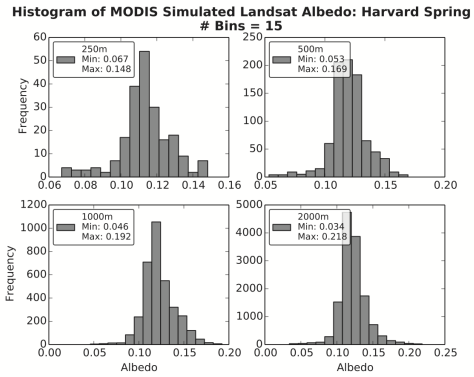
Figure 3.2: The site characteristics respective of the geolocation grid aggregation scheme for the Bondville site are seen in 3.2a. The remaining plots are distributions of the source Landsat albedo for each season at increasing spatial scales (250m-2000m) centered around the site.

less variability as the GSD changes. The distributions for all season at the Bondville site indicate an irregular distribution at finer scales that tend to smooth at coarser resolutions, indicating inherent spatial heterogeneity at the site.

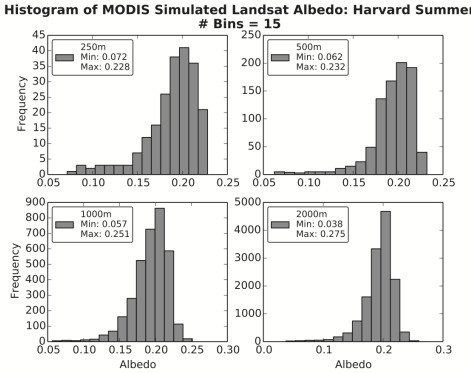
The Harvard Forest LTER (EOS Test Site) is a deciduous forest in the northeast United States, which would be expected to demonstrate some seasonality as a result of leaf-on/leaf-off periods. The site relative to the geolocation grid is illustrated in Figure 3.3a. The histograms for the three seasons in Figure 3.3 are



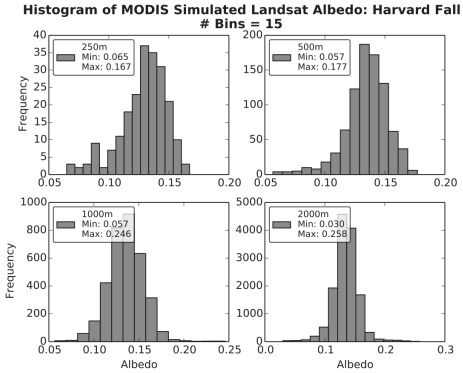

(a) Harvard Forest, MA




(b) Spring albedo frequency




(c) Summer albedo frequency

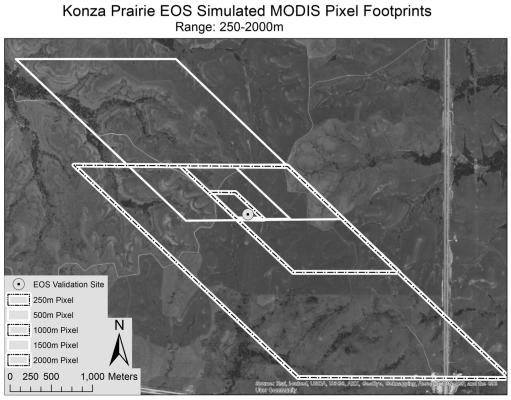



(d) Fall albedo frequency

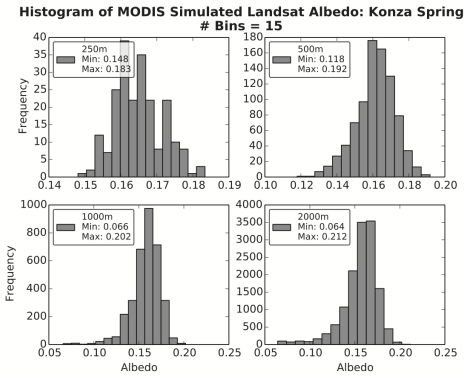
Figure 3.3: The site characteristics respective of the geolocation grid aggregation scheme for the Harvard Forest site 3.3a and histograms of albedo frequency for the different seasons and scales.

generally more normally distributed than the Bondville site, suggesting this site is more spatially homogeneous. However, the distributions demonstrate a skew at finer resolutions (250m and 500m), particularly in the Summer period, which can manifest as greater variability with varying GSD.

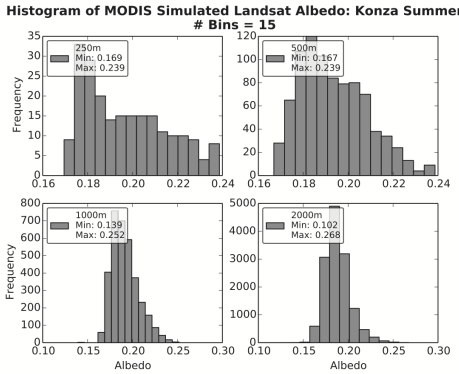
The Konza Prairie LTER site is a grassland land cover class in the Midwest United States. The site relative to the geolocation grid is illustrated in Figure 3.4a suggests the site is relatively spatially homogenous at all scales. The histogram distributions in Figure 3.4 have a relatively normal distribution for all scales, similar



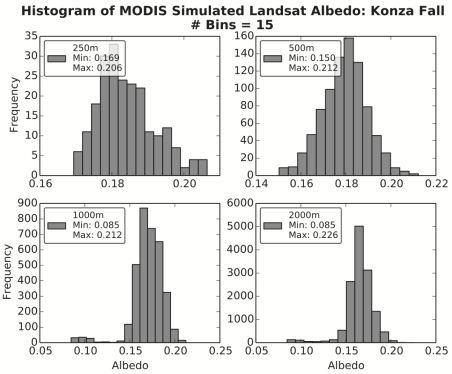

(a) Konza Prairie, KS




(b) Spring albedo frequency




(c) Summer albedo frequency




(d) Fall albedo frequency

Figure 3.4: The site characteristics respective of the geolocation grid aggregation scheme for the Konza Prairie site 3.4a and histograms of albedo frequency for the different seasons and scales.

to that of Harvard Forest. The Summer period shows a skew at finer resolution, though the frequency is denser than the skew in the Harvard site, which suggests the site will continue to be homogeneous.

The site characteristics respective of the geolocation grid aggregation scheme for Lake Tahoe are seen in Figure 3.5a and Jornada in Figure 3.5c. Land cover at the Lake Tahoe site was not expected to demonstrate high levels of seasonality. Therefore, only a single Fall image was simulated for Lake Tahoe. The Lake Tahoe EOS site is a clear water and needle leaf land cover class in the western United States.

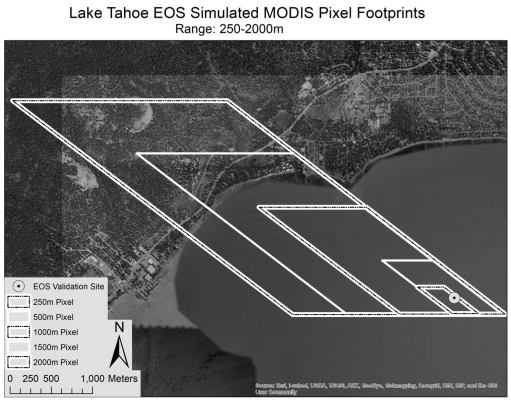
The Lake Tahoe site was chosen for the interesting feature of having the exact site location defined in the lake, but the aggregation of pixels ultimately incorporating land features at coarser resolutions. The Jornada LTER, located in the southwest United States, is classified as a semi-arid grass/scrubland. A single Summer image was selected for the Jornada site, since seasonality was not expected to influence the land cover.

The histograms for both Lake Tahoe and Jornada are seen in Figures 3.5b and 3.5d respectively. The histograms for the Lake Tahoe site are very irregular, which is an artifact of the incorporation of very different land albedos within the scene at coarser scales, and is highly spatially homogeneous at fine scales. The spatial homogeneity of the Jornada site is indicated through the generally normal distributions of albedo at all scales.

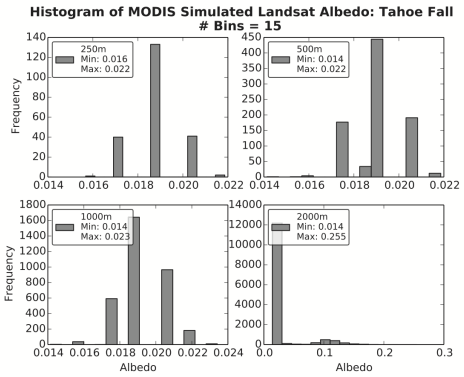
### 3.4.2 MODIS Albedo Simulation

MODIS bands 1 and 2 used for the simulation are nominal 250m bands, thus the decision was made to start the analysis for this research at the finest available pixel size. The process of generating the synthetic MODIS data from the Landsat scenes incorporates the known spatial integration of signal by creating a triangular weighting scheme for all Landsat pixels that fall into a given MODIS swath observation footprint. The simulation is accomplished in several steps.

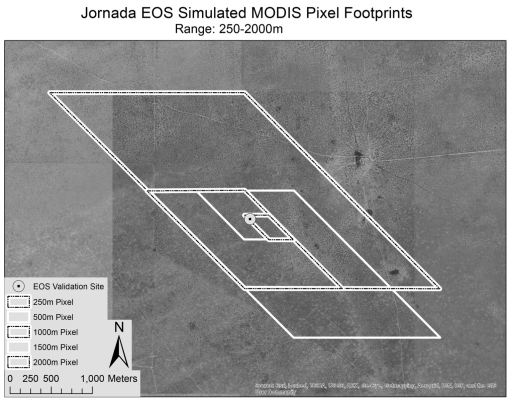
A MODIS observation center for the 250m resolution must first be determined from the input geolocation file (MOD03/MYD03), which is stored at 1km resolution.



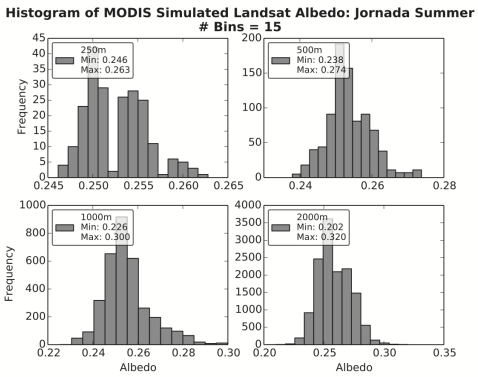

(a) Lake Tahoe, NV




(b) Lake Tahoe Fall albedo frequency




(c) Jornada LTER, NM




(d) Jornada Summer albedo frequency

Figure 3.5: Figure 3.5a for Lake Tahoe illustrates the composition of pixels at coarser resolutions and the level of land features that are added to the signal. The relative homogeneity of the Jornada site can be seen in Figure 3.5c.

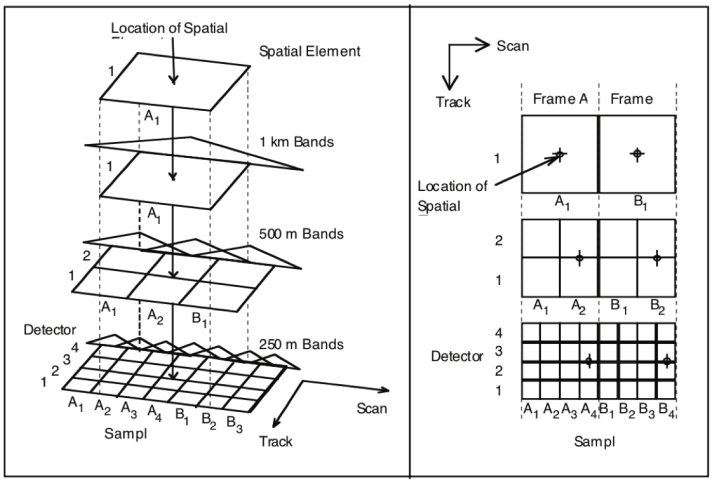



**Figure 2-11. MODIS Spatial Element Location**

Figure 3.6: The spatial relationship between 1km geolocation data and the 500m and 250m resolutions. The cell center of the 1km data is co-registered with the second and fourth 500m and 250m pixels respectively. An accurate geolocation for each of the resolutions can be determined though bilinear interpolation of the 1km values. (This figure is derived from [Nishihama et al., 1997, Figure 2-11]

The cell centers for the first along-track 1km earth scans are co-registered with the cell centers of the second 500m and fourth 250m spatial samples. Figure 3.6 from [Nishihama et al., 1997, Figure 2-11] shows the relationship between the spatial resolutions. Accurate geolocation must occur for each scan independently due to the overlap of successive scans. Therefore, each scan is calculated individually by dividing the number of rows in the MOD03 file by 10, i.e. the number of scans in each file. The 250m longitude and latitude for each pixel center is then calculated using bilinear interpolation of the 1km longitude and latitude values. The first two and last two rows in the 250m geolocation fall outside the actual 1km frame, so these values are approximated using bilinear extrapolation.

From these 250m cell centers the footprint of individual pixels can be de-

termined using methods described in [Tan et al., 2006, Appendix B], using the approximate scan angle and sensor azimuth obtained from the MOD03 geolocation file. The observation footprints derived using this method vary in size based on the given viewing geometries.

The MODIS along-scan Line Spread Function (LSF) is approximately triangular, integrating 12.5% of each neighboring pixel to the 75% pixel currently being observed [Nishihama et al., 1997]. The weighting scheme $W$ = [0.2, 0.4, 0.6, 0.8, 1.0, 0.8, 0.6, 0.4, 0.2] describes a representation of the relative contribution of signal for each MODIS pixel along-scan; the along-track LSF is assumed to be relatively rectangular Barnes et al. [1998], with equal weight in that direction. Therefore, a MODIS pixel is simulated by defining a 5 row x 5 column grid for each MODIS footprint, with a companion 5 row x 2 column grid defined for each of the adjacent pixel footprints starting from the pixel center. Weights from $W$ are defined along each respective vector of the grid columns, representing the relative contribution to the triangular LSF (Figure 3.7).

Landsat pixels located inside the observation pixel footprint and adjacent pixels are assigned weights from the nearest column of the weighting grid using nearest neighbor resampling. Along-scan weights are assumed constant and discrete for the sake of computational consistency and ease, so no effort was made to interpolate between column weights (e.g. bilinear interpolation). The weighting scheme remains unchanged regardless of the observation size, but the number of Landsat pixels assigned a given weight will increase or decrease based on the size of the footprint. The final observation value is calculated using the weighted average of all Landsat pixels

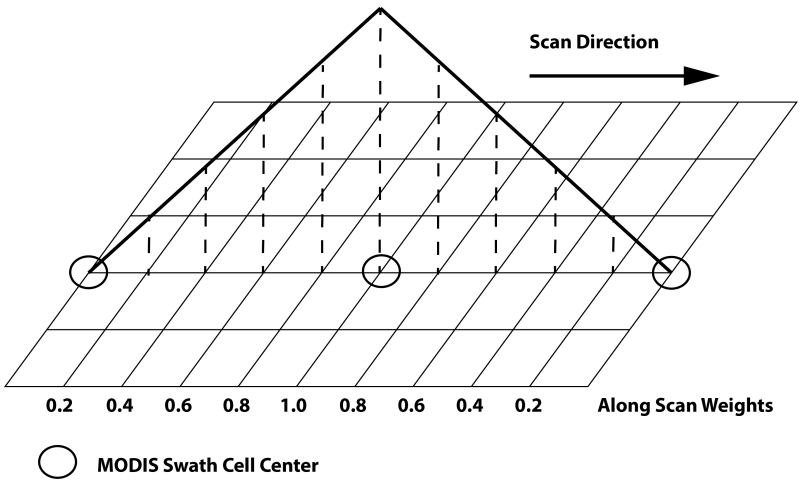



Figure 3.7: Illustration of swath observation simulation. Landsat pixels are assigned weights from a 5x9 grid centered inside the swath cell and incorporating signal from neighboring swath cells. The value for a given cell is then the weighted average of all Landsat pixels. The simulated swath observation is consistent with the LSF of the MODIS instrument.

using the weighting scheme $W$ for each respective column, respecting the 12.5% / 75% / 12.5% neighborhood contribution and the triangular shape of the MODIS LSF [Nishihama et al., 1997].

The individual Landsat albedo scenes were transformed to MODIS swath geometries for each MOD03/MYD03 file intersecting a given site during an $\approx$ 30 day period, representing at least two nadir-to-nadir periods (see Table 3.2). The resultant synthetic MODIS swath data were maintained in a WGS84 geographic projection, preserving the variable pixel size for the daily swath. The synthetic swath data were projected from geographic to the custom MODIS sinusoidal projection with a pixel resolution of 250m. The choice was made to keep the grid resolution at an even 250m for easy manipulation even though the actual MODIS grid resolution is $\approx$ 231m. The reprojection uses a nearest neighbor pixel assignment method to approximate the *obscov* method used in actual MODIS product generation. The as-

Table 3.2: MOD/MYD03 time periods used for simulating data

| Site | Platform | Season | Begin Date | End Date |
|---|---|---|---|---|
| Bondville | Aqua | Fall | 2010-11-08 | 2010-12-09 |
| | Aqua | Summer | 2010-07-03 | 2010-08-03 |
| | Aqua | Spring | 2010-04-14 | 2010-05-14 |
| Harvard Forest | Aqua | Fall | 2006-10-13 | 2006-11-14 |
| | Aqua | Summer | 2010-07-03 | 2010-08-03 |
| | Aqua | Spring | 2010-04-14 | 2010-05-14 |
| Konza Prairie | Terra | Fall | 2011-09-02 | 2011-10-01 |
| | Terra | Summer | 2011-06-30 | 2011-07-31 |
| | Terra | Spring | 2011-03-10 | 2011-04-08 |
| Jornada LTER | Terra | Spring | 2011-04-30 | 2011-07-08 |
| Lake Tahoe | Aqua | Fall | 2009-09-21 | 2009-10-26 |

signment difference between the resampling methods is expected to be small and is consistent with tools like the LPDAAC MODIS Reprojection Tool Swath [LPDAAC User Services, 2015].

The daily MODIS Surface Reflectance (MOD09GQ) processing steam selects the highest quality individual pixels from multiple orbits for the first layer based on quality criteria including overall observation quality, maximum *obscov* and minimum VZA [Vermote et al., 2011]. However, the simulated data in this study are considered to have equal observation quality and *obscov* was not generated. Therefore, only

the minimum VZA criteria was applied to select individual daily observations.

Nominal 250m band MODIS data used in coarser resolution science products (e.g. MODIS Albedo MCD43 at 500m resolution) are aggregated using the weighted average of all observations intersecting a given coarse resolution cell determined by the *obscov* values for the input pixels. For this study, a simpler method was applied assuming any given resolution is an equal subdivision of the original sinusoidal grid - a 250m pixel in the sinusoidal grid is the 4x4 subdivision of the 1km grid for example. Each resolution is calculated from the original 250m grid data, with the simple mean value calculated from all pixels falling into the new, coarser pixel. The full process is illustrated in Figure 3.8. Finally, the pixel value for the grid cell containing the study site was extracted for each gridded scene in the series.

Although the process is not completely consistent with the production code used for MODIS land products, the process does provide an approximation of the variability. The results of the simple method demonstrate a similar effect of data smoothing through the averaging of data to coarser resolution even though the method of aggregation differs.

## 3.5 Results

The goal of the simulation process is to hold albedo constant, with the geometric effects from the MODIS instrument as the only variable; geometric variation and inconsistent swath-to-grid assignment from orbit-to-orbit leads to differing measurements of the same underlying albedo at the same location, as the GSD changes.

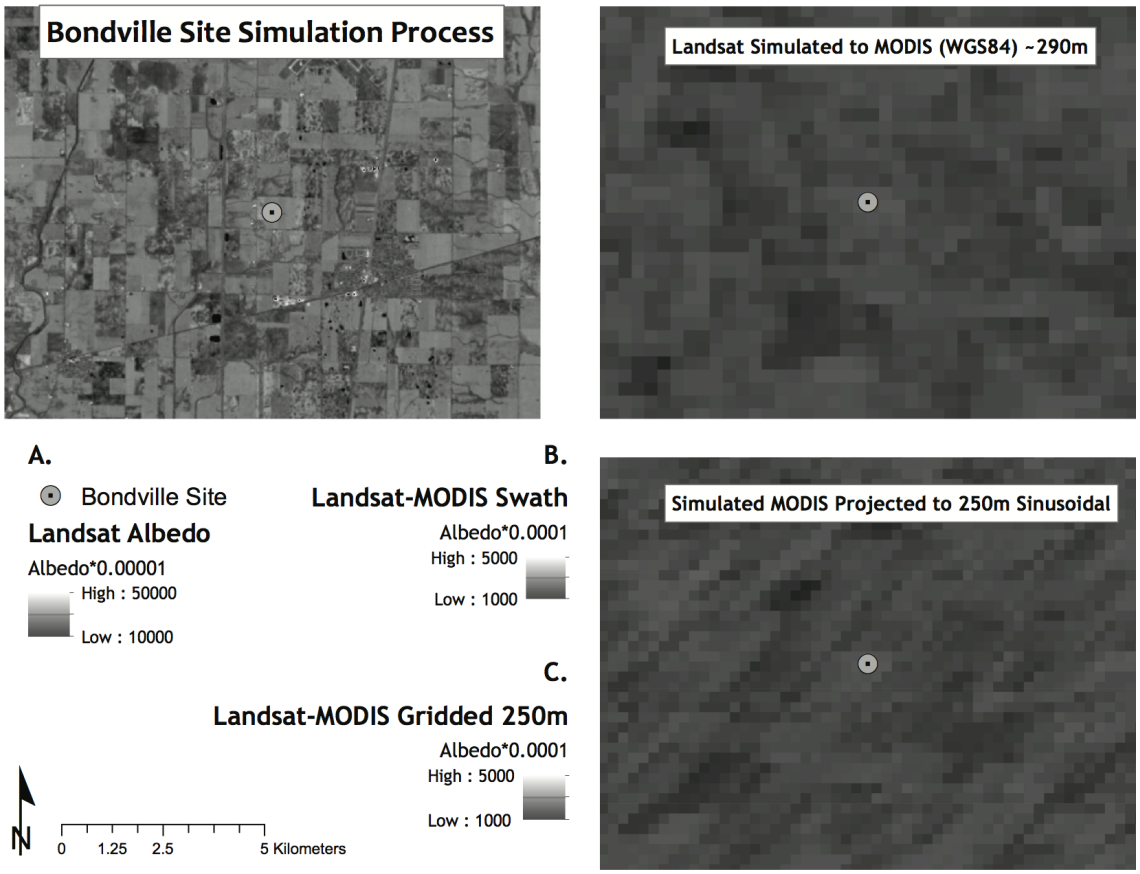



Figure 3.8: The full process of synthesizing data. Clockwise from top left: A) The original Landsat synthetic albedo at 30m resolution. B) The next step transforms the Landsat at 30m resolution to daily MODIS observation geometry derived from daily MOD03 geolocation data and projected to a geographic coordinate system preserving the variable swath pixel sizes. This example uses data from Aqua and has a VZA of 26.47°, resulting in an $\approx$ 290m pixel. C) Finally the daily synthetic swath data are gridded to the MODIS sinusoidal grid at 250m resolution.

Data from the MODIS instruments on both Aqua and Terra are used to illustrate any differences to be seen in the viewing geometries from a morning descending orbit (Terra) versus an afternoon ascending orbit (Aqua).

Two measures are used to determine variability in the data, range and standard deviation. Standard deviation provides the basic metric of variability around the observed mean values. The range of measured albedo (*maximum value* − *minimum value*) is a useful measure in this case to assess the geometric error. The range of signal variation from peak to trough through time is related to the observation footprint as a function of the changing GSD from the variable VZA, since the source Landsat albedo only varies due to relative contribution in the changing GSD. Observations made at increasing VZA will necessarily incorporate more signal from adjacent space as the observation footprint grows to be greater than the grid pixel size. The range of values for homogeneous land covers should be small, since the underlying albedo would vary little with changing GSD. However, spatially heterogeneous land covers are expected to demonstrate increased ranges, influenced by the variable GSD and the fluctuating albedos present in the footprint.

An example of simulated swath data from the Harvard Forest, MA site in Figure 3.9 clearly shows the extreme values are not observed at the extreme range of VZA. Rather, the extreme values are observed in the middle range of VZA between $\approx 5^\circ$ and $35^\circ$, while the smallest is near nadir. The range incorporating these values is relatively large at 0.038, though the standard deviation is far smaller at 0.009. Excluding the extreme values still retains a range of values $\approx 0.025$. These measurements are made with data prior to the grid process, suggesting a significant

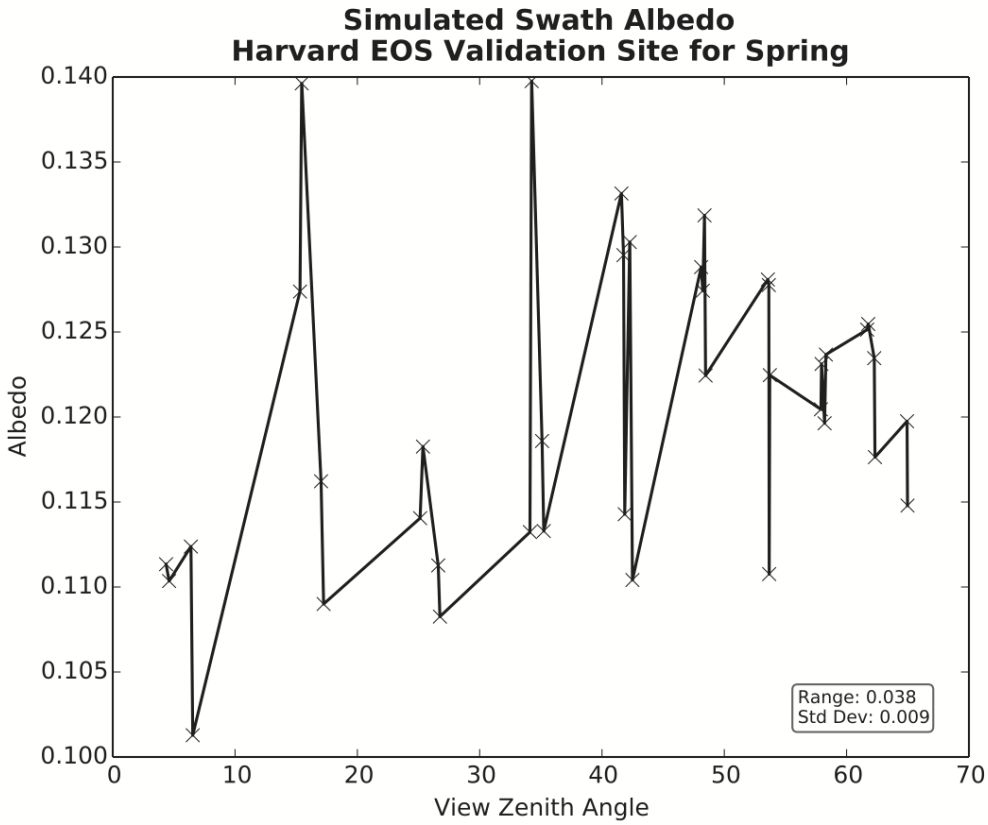



Figure 3.9: Albedo plotted as a function of VZA for Harvard Forest site. The results show the largest values are observed in the mid range of VZA between $\approx 15°$ and $35°$. Overall variability for these simulated swath data appears low reporting a measured standard deviation of 0.009. The range of values is larger at 0.038, which is a significant contribution to uncertainty from the variable GSD before gridding occurs.

variability from the GSD alone.

The results for all sites and seasons during the entire simulated period, Figure 3.10, are organized in ascending order of increasing pixel size ranging from 250m to 2000m in 500m increments showing the range and standard deviation of daily measured albedo. The results clearly demonstrate a measurable level of variability from the geometric properties of the MODIS instrument, which would carry forward to the final grid product. Reduction of both standard deviation and range is achieved

through aggregating data to coarser grid resolutions, but full removal is not achieved at any resolution in the simulation. The increasing values for the Lake Tahoe site are an artifact of the center of the EOS site being located offshore, but ultimately incorporating land features at increasing grid resolutions.

The overall low values for standard deviation, generally $< 0.01$ albedo, suggest a given temporal series of data is relatively stable. However, the relationship between climate model and albedo parameters was previously described as $+0.01$ albedo corresponding to $-0.05\ Wm^{-2}$, and $-0.03$°C cooling. The level of variability seen at the finest resolution, 250m, are sufficient to influence radiative forcing in climate models attributed to the variable observation geometry of the instrument and geolocation, independent of atmospheric or albedo model related sources.

The reduction of both range and standard deviation from aggregating the data to 500m in this study, suggests products such as the 500m MODIS albedo product [Schaaf et al., 2014, 2002], MCD43, incorporating multiple days of observations from both Aqua and Terra, and therefore GSD, to approximate a nadir adjusted value are not impacted greatly from the varying geometry. However, the higher values for the range indicate significant levels of uncertainty remain when individual geometries are integrated into temporally discrete composite data products like the NDVI in Figure 5.1. This applies to both assignment of varying GSD in a given scene, or from composite period-to-period. The variability observed from fine-resolution GSD in a scene could affect land cover classification accuracy for example.

The focus of the above research was to demonstrate the impact of geometric distortions as a result of off-nadir viewing. The analysis shows the impact, and

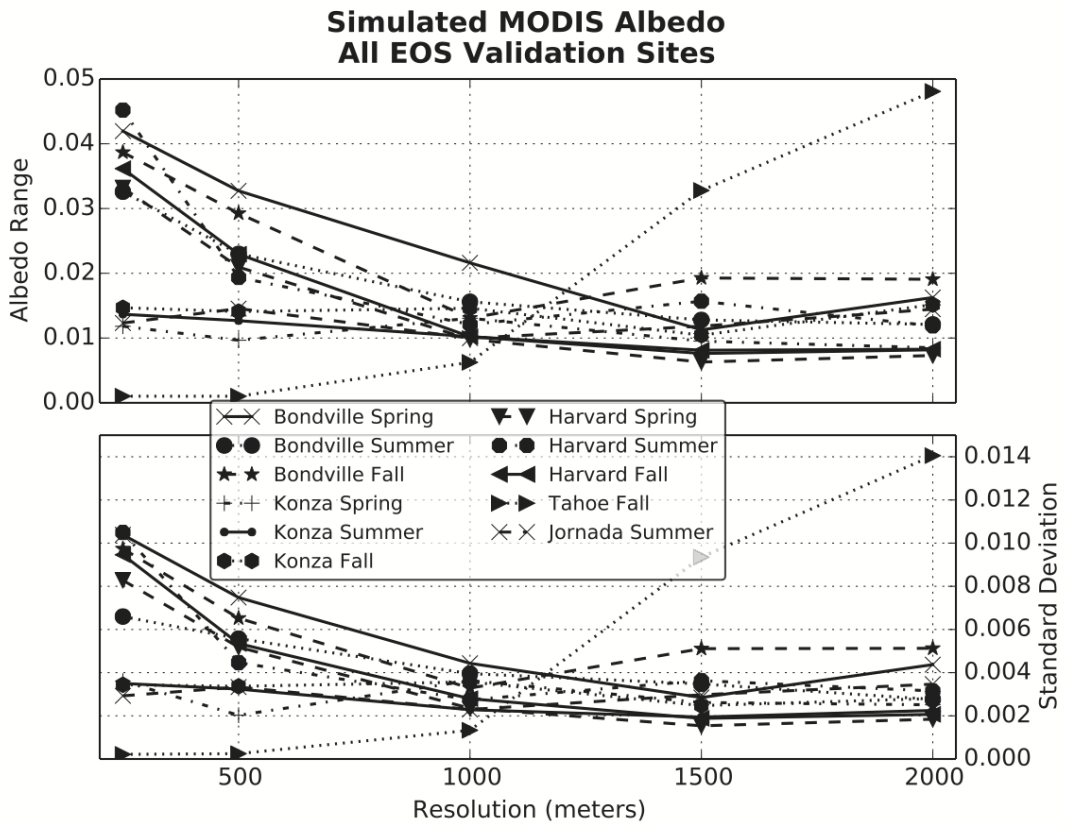



Figure 3.10: Albedo ranges and standard deviation for all sites. The plot is arranged in ascending order of increasing pixel size from 250m-2000m illustrating the albedo variability for the three seasons tested.

that, in general, the smoothing effect of aggregating the data to coarser resolutions is expected to reduced the resultant measurement variability. The analysis also revealed a relationship between land cover spatial heterogeneity, the variability in the signal and the level of spatial aggregation required to reduce albedo variability at different sites, which is described individually in the following sections.

### 3.5.1 Bondville, IL

The results for the Bondville site, Figure 3.11, show the range (upper figure) and standard deviation (lower figure) for the three seasonal periods for all data in each of the respective simulations. The values indicate the expected smoothing effect from pixel aggregation, though the variability is never totally removed. The Spring and Fall values demonstrate the highest variability. This variability is likely a function of the spatial heterogeneity seen in the irregular Landsat albedo frequency seen in Figures 3.2b and 3.2d, and is expected since the Spring and Fall periods would present the most heterogeneous canopy as crops are planted, then either harvested or senesced. The Summer period has a smaller variability from a more fully developed canopy providing a more uniform signal. The uncertainty for all seasons is sufficient to indicate potential model influence at fine scales, up to at least 1000m.

Several MODIS biophysical products are generated as *discrete* temporal composites, using only the single best quality per-pixel observation (e.g. cloud-free and/or minimum view zenith angles) for a given composite period - usually 8-16 days. The series of plots in Figure 3.12 group the simulated data into discrete 8-day groups and calculate a range for each period. The 250m and 500m groups demonstrate persistent variability greater than 0.01 albedo for all seasons, which is sufficient to influence an expected result within any given 8-day group, and suggest care is needed when using observations from different geometries during a given composite period.

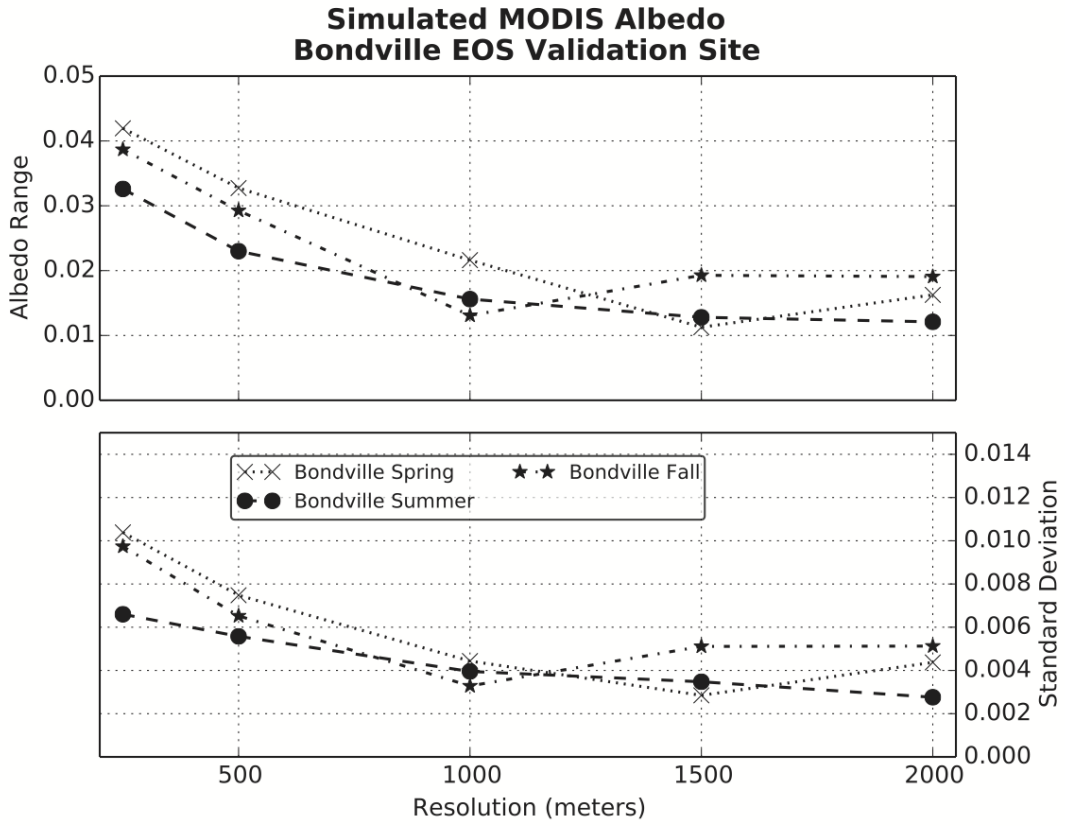
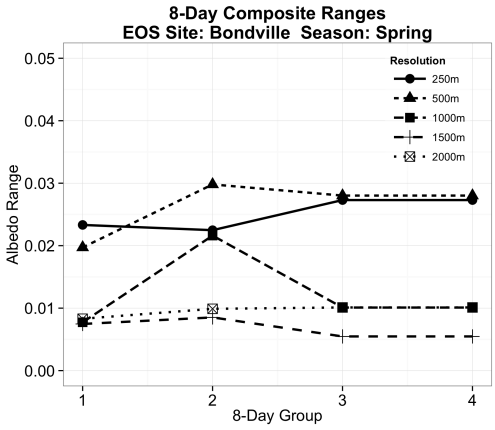
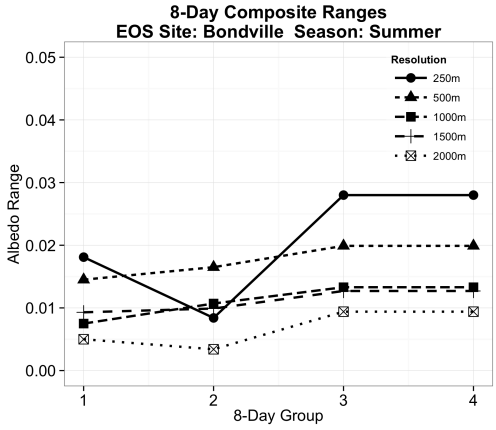


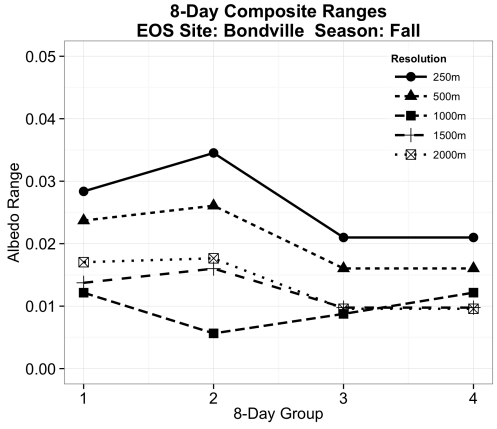

Figure 3.11: Albedo ranges (upper figure) and standard deviation (lower figure) for the Bondville site. The results for Bondville indicate a resolution of at least 1000m to stabilize the effect from geometric distortion, though it is not fully removed even at 2000m.




(a) Bondville Spring




(b) Bondville Summer




(c) Bondville Fall

Figure 3.12: Discrete 8-day groups of simulated albedo for the Bondville site.

### 3.5.2 Harvard Forest LTER, MA

The variability in Figure 3.13 appears similar to the Bondville results for the 250m resolution. However, the reduction of variability achieved more rapidly than Bondville, the majority occurring at 500m, and a relative minimum reduction is achieved at 1000m, similar to the Bondville site. The level of spatial heterogeneity at the Harvard site appears constrained to finer resolutions, and becomes more homogeneous at coarser resolutions. The Landsat albedo frequencies for the site, Figure 3.3, support this, particularly the Summer results that demonstrated a skewed distribution that is seen here as the highest level of uncertainty. The standard deviation values are small, suggesting the data are stable when using a mean value. However, the uncertainty remains a potentially significant contribution to the overall accuracy of the signal at finer resolutions.

Figure 3.14 groups the simulated data into similar discrete 8-day groups as 3.12. The Spring and Fall scenes demonstrate low variability from period-to-period and remain below 0.02 for all resolutions, contrary to the global variability in Figure 3.13, which is cumulative over the range of all values in the simulation. The discrete results suggest the site is far more homogeneous and stable through geometry, though some care should be taken with fine resolution data that present the highest per-group uncertainty. The results for Summer, however, demonstrate high levels of variation in the 250m data. The effect is mitigated through aggregation, becoming more stable as the other seasons. The results for Summer suggest 250m data are sensitive to GSD-to-grid assignment in developed tree canopies.

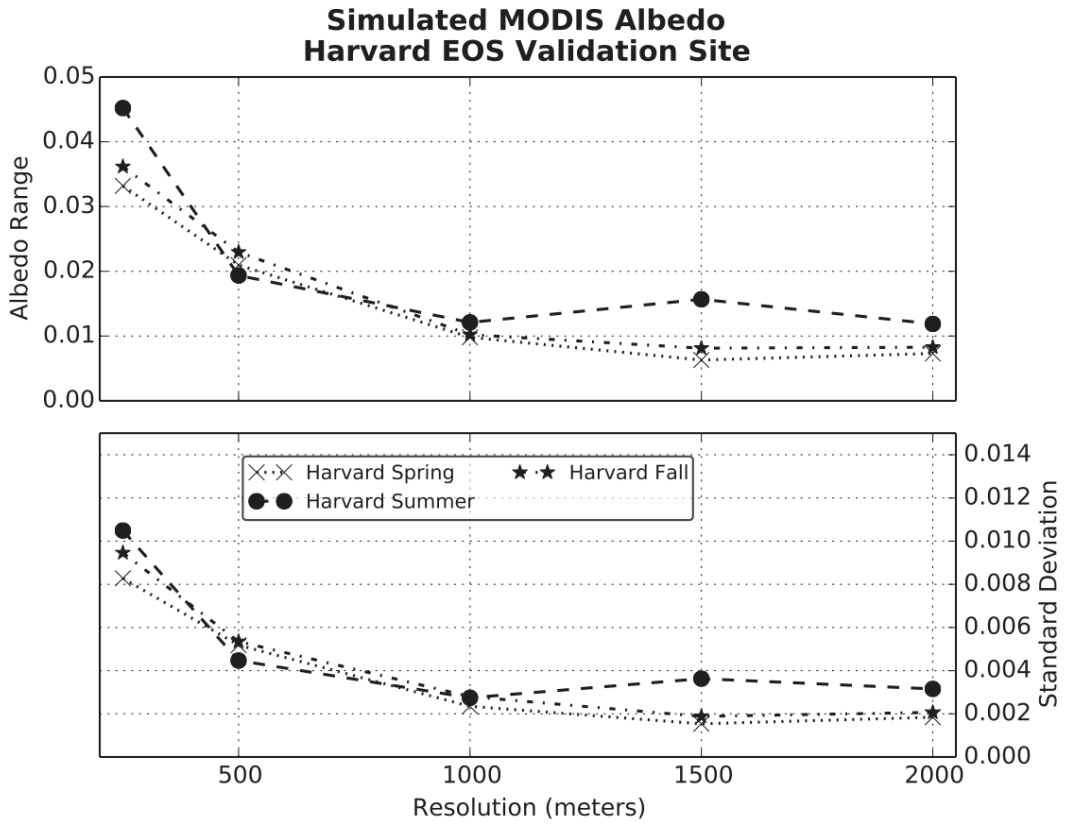
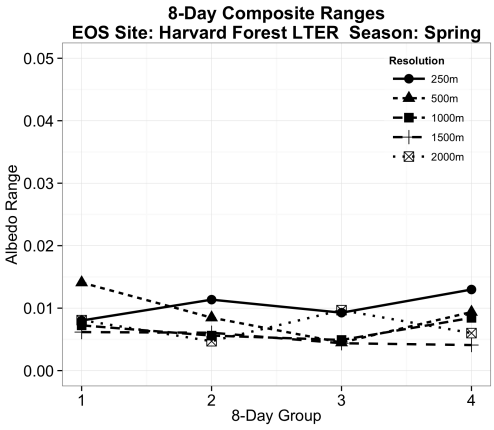
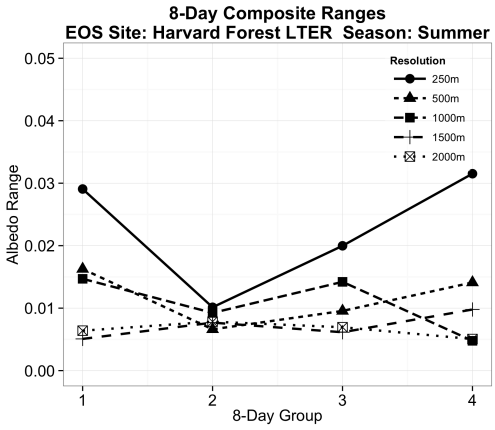


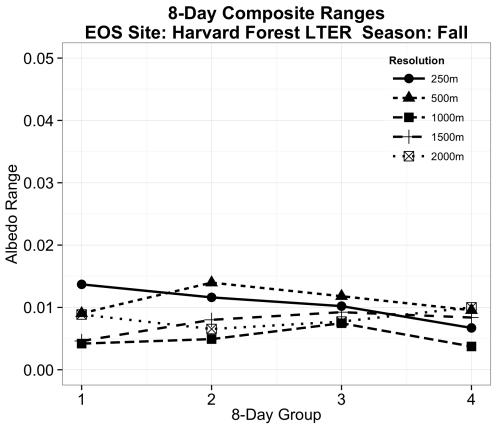

Figure 3.13: Range (upper plot) and standard deviation (lower plot) for the Harvard Forest site. The albedo data for Harvard Forest demonstrate a continuing reduction of error as pixel size increases.




(a) Harvard Forest Spring




(b) Harvard Forest Summer




(c) Harvard Forest Fall

Figure 3.14: Discrete 8-day groups of simulated albedo for the Harvard site.

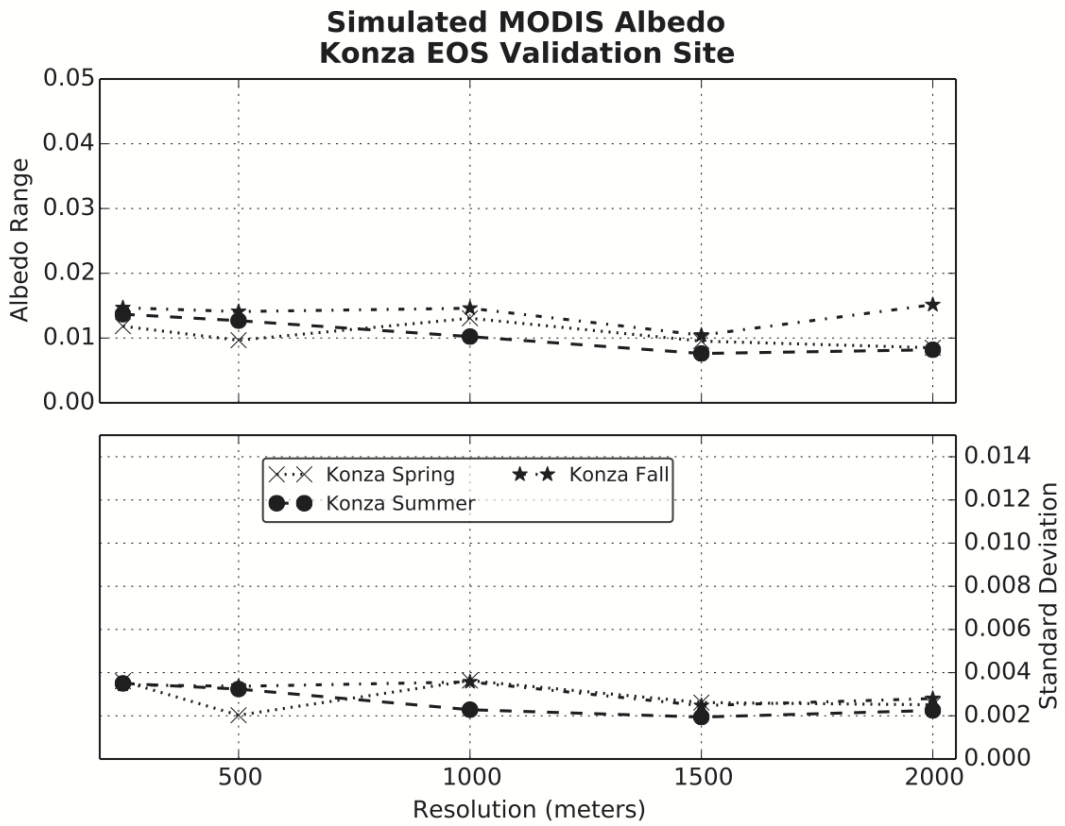
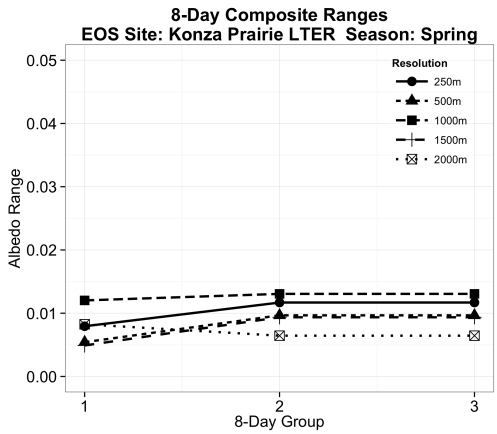



Figure 3.15: Albedo ranges and standard deviation for the Konza Prairie site. The relatively homogeneous land cover at the Konza Prairie site demonstrates an overall small and consistent uncertainty for all resolutions.

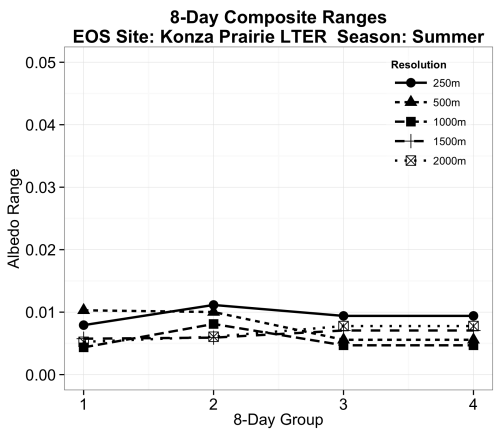
### 3.5.3 Konza Prairie LTER, KS

The results for Konza Prairie, Figure 3.15 demonstrate a low, and stable overall uncertainty for all resolutions. The standard deviation is insignificant, but the range data, while small, remains a constant source of significant uncertainty despite aggregation. The land cover seen in Figure 3.4a appears relatively uniform compared to Bondville and Harvard Forest indicating a high level of spatial homogeneity for all seasons, and is supported in the Landsat albedo frequencies in Figure 3.4.

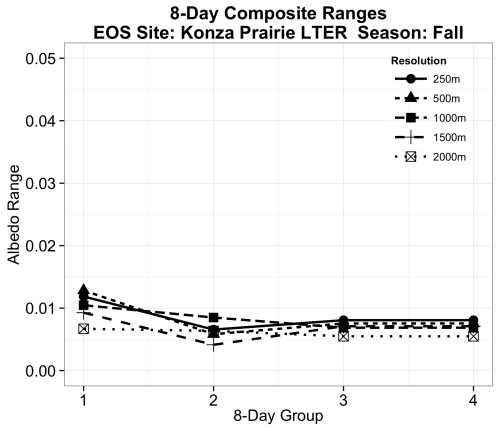
The discrete analysis, Figure 3.16 shows the same basic pattern of low un-




(a) Konza Prairie Spring




(b) Konza Prairie Summer




(c) Konza Prairie Fall

Figure 3.16: Discrete 8-day groups of simulated albedo for the Konza site.

certainty between groups similar to the Harvard site. The Fall data show the largest variation, though this is small. Overall, the results for Konza Prairie support MODIS data present a consistent uncertainty for homogeneous landscapes.

### 3.5.4 Lake Tahoe, NV and Jornada, NM

The final two sites are presented together since only one season was analyzed for each. The Lake Tahoe site was chosen for the interesting feature of having the

exact site location defined in the lake, but the aggregation of pixels ultimately incorporating land features at coarser resolutions. Figure 3.5a illustrates the composition of pixels at increasing resolutions and the level of land features that are added to the signal. The level of variability is inverse to the expected smoothing for resolutions greater than 500m seen in Figure 3.17, as the spatial variability of land features is introduced. Ultimately, the variability is the highest level for all the measured sites, but at the coarsest resolution. The integration of potentially disparate land covers into any given GSD is a consequence of the MODIS instrument's wide field-of-view. The variable land cover integration is carried forward into the gridded product, so care should be taken making assumptions about per-pixel stability.

The Lake Tahoe site discrete results, Figure 3.18a, are similar to the global in that nominal 250m or aggregating data to 500m is the most representative for this type of extreme heterogeneous land cover. Aggregation to coarser resolutions introduces a level of uncertainty to each period, unlike the other land covers represented in this research.

The Jornada site is relatively homogeneous, which can be seen in Figure 3.5 and is confirmed by the Landsat albedo frequencies in the figure. The homogeneity of the site is further confirmed with an low level of uncertainty for all resolutions in Figure 3.17

The discrete analysis of the Jornada data, Figure 3.18b, shows the 250m data return the lowest uncertainty of all the resolutions. This result, and a similar finding for the Konza Prairie site, suggests aggregation introduces heterogeneity from adjacent areas not otherwise incorporated in the nominal pixel. However, aggrega-

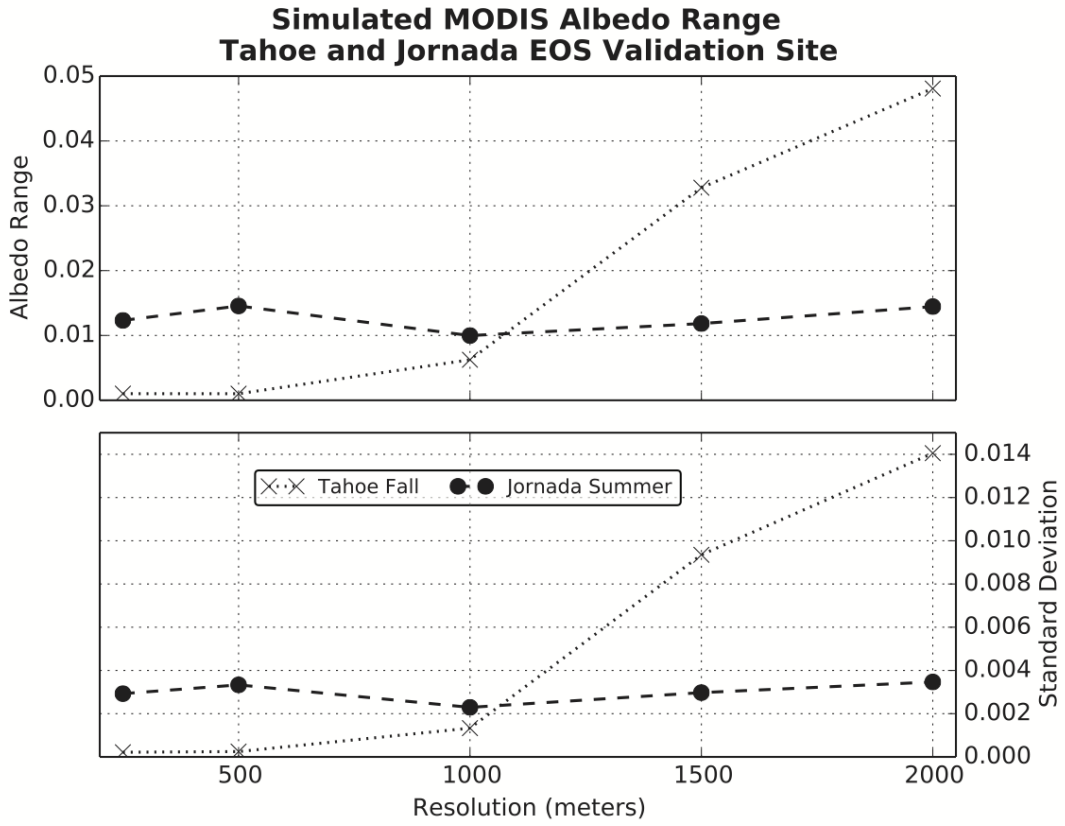
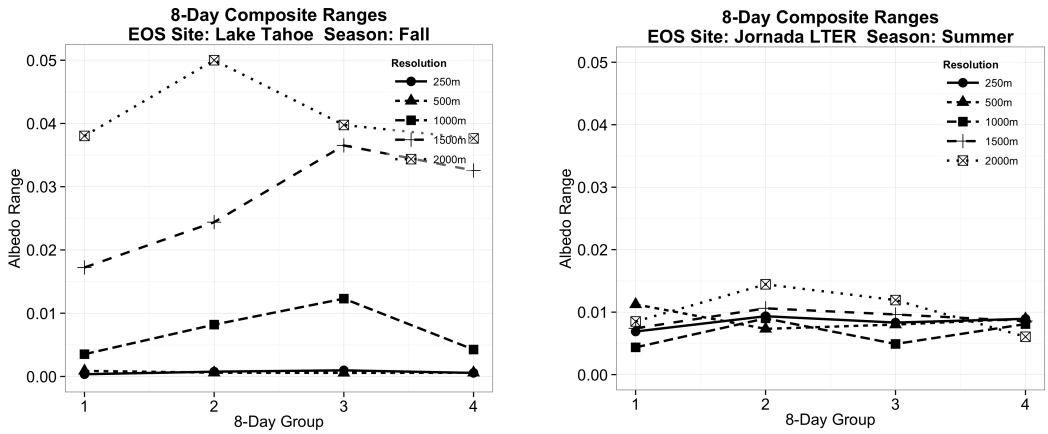



Figure 3.17: Albedo ranges and standard deviation for the Lake Tahoe and Jornada sites. The actual site center for Lake Tahoe is located offshore, but inclusion of signal from land features is steadily introduced at increasing spatial resolutions increasing the variability in the signal. The Jornada data present a stable level of uncertainty for all resolutions.




(a) Lake Tahoe Fall

(b) Jornada Summer

Figure 3.18: Discrete 8-day groups of simulated albedo for the Lake Tahoe and Jornada sites.

tion to resolutions needed to stabilize variability in other sites does not necessarily introduce a significantly different uncertainty.

### 3.5.5 Average Resampling

Current MODIS land data products use the highest proportion of a GSD intersecting a given grid space, *obscov*, to assign a grid value [Wolfe et al., 1998, Yang and Wolfe, 2001]. However, the resampling of a widely variable GSD to a fixed 250m grid cell does not necessarily represent the sample space. Specifically, the misalignment of observations occur when two observations partially intersect a grid cell and only a single value, with the largest coverage, is assigned; the feature of interest may be mis-assigned if the observation coverage is insufficient.

A second method of data processing was attempted as part of this study, to mitigate the inconsistent assignment resulting from the nearest neighbor (*obscov* in

production) resampling. This method uses the average of all observations intersecting the grid cell to assign the grid a value. The average resampling technique is more representative of the changing GSD through the inclusion of all data informing a given grid space. Average resampling is processed from swath at all resolutions rather than aggregating the results of resampling a single resolution to progressively coarser resolutions as was done above. In a production environment, the simple average used here should be replaced with a weighted average using *obscov*, the same method used for aggregation in production. The results are seen in Figure 3.19.

The immediate effect of the average resampling method is a reduction of range by at least half for the worst sites and a far more consistent reduction for all sites except Lake Tahoe, which demonstrates an increase. The standard deviation is similarly reduced for finer resolutions. However, some level of geometric error remains. The Bondville site for the Fall period demonstrates a clear increase at coarser resolutions, which is likely reflective of the introduction of spatially heterogeneous land cover similar to the Lake Tahoe site and the inherent sensitivity of averages to outlier values. The Lake Tahoe site does not benefit, and in fact suffers from the average resampling method.

These results indicate a more uniform choice of aggregating data to 1000m to mitigate geometric effects and produce a more consistent error across land covers. The average resampling swath method also presents a better tradeoff between resolution and accuracy.

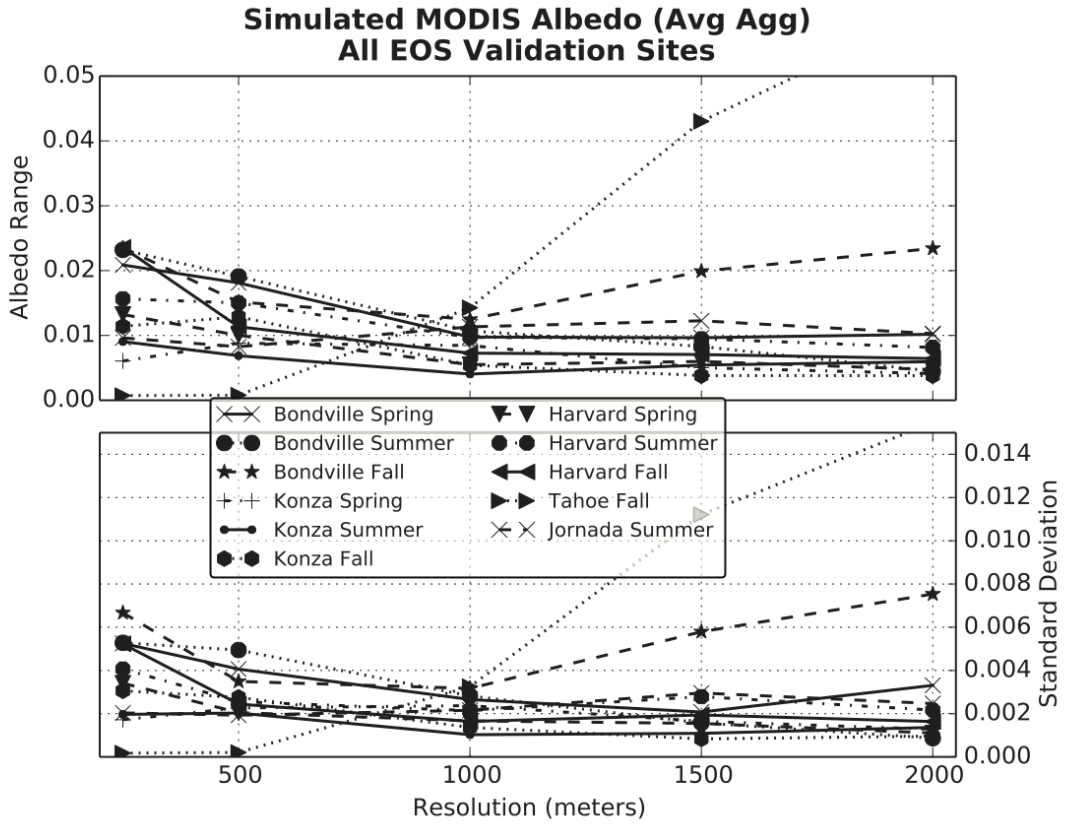



Figure 3.19: Albedo profiles for all sites using an average resampling method for all resolutions. All sites except for Lake Tahoe demonstrate a reduction of range and standard deviations at fine resolutions. The average resampling method presents a more consistent and smaller error at a finer resolution.

## 3.6 Discussion and Conclusion

The objective of this research was to investigate the effect of geometric distortions inherent in the scanning MODIS system using broadband albedo as an example to quantify uncertainty. It should be noted that the purpose of this paper is not to assess the accuracy or model methods used for the MODIS standard albedo product (MCD43) or any other MODIS standard product. Rather, the findings of this work are meant to inform users of MODIS land products of the level of uncertainty added by the distortion inherent in MODIS data due to off-nadir viewing, using albedo as a meaningful physical variable. A quantified error budget associated with the spatial representativeness of wide field-of-view data like MODIS will improve the accuracy and inform end use of current and future product generations through a better understanding of specific sources of error. The work has led to several valuable conclusions.

First, the assignment of a variable daily geometric footprint to a consistent geolocation grid has a significant influence on daily observations of synthetic albedo. Variability in albedo attributable to geometric effects alone ranged as high as $\approx 0.045$ to $< 0.01$. Even applying a more robust assumption of maximum error that is half the range the results, still present uncertainty in albedo $< 0.015$ for most sites at nominal resolution. These ranges from geometric and geolocation effects alone would be sufficient to influence climate models utilizing MODIS data. The standard deviation is much smaller for the same data and range from $\approx 0.01$ to $< 0.002$, suggesting algorithms integrating multiple GSDs into a single, adjusted value suffer

less from the inconsistent geometry.

Second, the research illustrates that error from geometric distortions can be mitigated through the smoothing of data by aggregation of pixels to coarser resolutions. Pixel resolutions in excess of 1000m are required to mitigate the geometric influence from some land cover types, though 1000m provides a relatively consistent result for all sites. The level of aggregation is consistent with previous findings using gridded data [Tan et al., 2006], as well as the known 4.8 times GSD growth along scan.

The fact that complete removal of noise is not fully achieved even at coarse resolutions up to 8 times the nominal 250m resolution when the known GSD of the instrument grows up to 4.8 times nominal resolution, indicates models using these data must incorporate an uncertainty attributable to geometric distortions before model error can be assessed. It should be noted that the current MCD43 BRDF/Albedo is generated at 500m grid resolution rather than the 250m geometries used for the above simulation [Schaaf et al., 2014].

The applicability of the results presented here is significant despite this difference in resolution for several reasons. The Red and NIR bands in the MCD43 product are nominal 250m MODIS bands that are aggregated to 500m using a more robust method than that presented here. The results for 250m bands presented here clearly demonstrate the smoothing from just a simple aggregation method to 500m helps stabilize the geometric uncertainty, and the 500m nominal bands could likewise benefit from aggregation to 1000m. Additionally, the geometries for all MODIS detector resolutions function the same, but at different scales. Therefore, even the

simple assumption of doubling of the minimum effective GSD for 250m grid data [Campagnolo and Montaño, 2014] at nadir (e.g. $250m \approxeq 340m \therefore 500m \approxeq 680m$) suggests a 500m grid product is not representative of the space being observed through varying geometry. While MODIS products are designed to accurately present the portion of the space observed in a given grid cell, there is no specific benefit from using a smaller grid cell size than a cell size more representative of the maximum variable observation footprint for example; a resolution increasing the average *obscov* to greater than 30% [Tan et al., 2006] would be beneficial to stabilize the spatial signal.

Finally, the assessment of an alternate resampling method demonstrates averaging all swath pixel that fall into a grid space rather than nearest neighbor (as a proxy for *obscov*) is a more consistent spatial representation across the range of GSD. Future research of alternate data processing methods should also investigate the efficacy of other equal area projections that can improve the swath-to-grid alignment over the sinusoidal projection. Campagnolo and Montaño [2014] demonstrate the sinusoidal grid causes a degradation of the resolution for locations away from the Equator and the central meridian of the projection from inconsistent swath and grid orientation, which might be improved with a more localized projection.

The results demonstrate the gridded data from different GSDs are not inherently interchangeable, and selection of data for any given period of geometry introduces a level of uncertainty that should be factored into product accuracy assessment for these data. This is particularly true of nominal 250m data, but also for other resolutions. Spatial homogeneity for the site is an important factor when

considering observation stability across the range of GSDs. Relatively homogeneous sites like Konza Prairie, Jornada, and Harvard Forest provide a stable, though potentially significant, level of uncertainty for all resolutions.

Spatially heterogeneous sites like Bondville can be smoothed through aggregation to similar levels of uncertainty to accommodate the spatial variability. Field size is also a factor when considering the most appropriate scale for MODIS data. Sites like Bondville, with agricultural fields smaller than a given GSD, will present different portions of those fields with every orbit at finer resolutions. Thus these land cover types benefit from a coarser scale to stabilize the signal. Conversely, some care should be taken not to aggregate the data to level that no longer represent the land cover of interest, as the Lake Tahoe results demonstrated with two disparate adjacent land cover types.

The MODIS instruments have proven to be valuable contributions to the ongoing monitoring of Earth's surface [Gitelson et al., 1996, Huete et al., 1999, Ji et al., 2008]. There remains an important tradeoff between achieving high-temporal frequency observations and spatial accuracy; the results shown in this work clearly demonstrate the importance of minimizing measurement error arising from known instrument artifacts. However, the presence of known artifacts does not negate the use of existing nominal-scale data. Rather, the final use informs the level of utility.

Nominal 250m gridded data are adequate when the final accuracy needed is relatively low as might be the case for a simple forest/non-forest classification or visual analysis. The results using standard deviation also indicate the uncertainty of measurements are stable when considering products that integrate multiple GSD

into a single measurement. Physical variables or analysis assuming a stable spatial signal through time, per-pixel comparison from two different 8-day composites for example, and requiring a high level of accuracy or measurement sensitivity can benefit from adding a geometric uncertainty factor based on the known variability of the GSD to increase the precision.

The use of whisk-broom sensors will continue and there is considerable interest in developing long term consistent climate data records from satellites [Justice et al., 2011]. The Visible Infrared Imaging Radiometer Suite (VIIRS) instrument, part the orbiting Suomi NPP mission and to be flown on the upcoming JPSS mission will provide data that are aggregated along scan to help control for the known effects of GSD growth. The intention of the aggregation is to provide a more consistent observation footprint than MODIS does currently, and early spatial performance [Wolfe et al., 2012] indicates the method is effective. Standard products from VIIRS, similar to the current MODIS suite, are available in beta form now. Future research will investigate efficacy of the cross-scan pixel aggregation using the methods presented here.

# Chapter 4: Evaluating VIIRS Product Uncertainty associated with Scanning Geometry: an Albedo Case Study

## 4.1 Abstract

This research examines uncertainty in Visible Infrared Imager Radiometer Suite (VIIRS) observations, and demonstrates the direct influence of geometric distortions resulting from the standard practice of geolocating swath observations. The VIIRS data are unique amongst coarse resolution earth-sensing systems, as a result of the user requirements to exhibit uniform pixel size across scan. This was addressed by on-board spatial aggregation of pixels toward nadir, to mitigate the growth of Ground Sample Distance (GSD). The pixel aggregation scheme employed for VIIRS data in an effort to mitigate the growth of GSD is assessed. No significant difference was found between aggregation zones for homogeneous land covers, but spatially heterogeneous land cover demonstrated a marginal difference in the mean albedo between zones. VIIRS observations vary dependent on the Ground Sample Distance (GSD), which varies dependent on the View Zenith Angle (VZA) that changes with every orbit. Gridded VIIRS land products (L2G) are generated by applying a geolocation algorithm that resamples the variable observation geometries

to a consistent grid of fixed pixel size and location, a process which itself introduces variability associated with the changing observational footprint. For this case study, broadband albedo was simulated for five validation sites, representing five distinct land cover types, exhibiting quantifiable variability, with additional, seasonal variability exhibited in some sites. Results from discrete 8-day groups of data simulating temporal composite data demonstrate the influence of geometric artifacts through differing levels of uncertainty between periods sufficient to influence climate models (i.e. ranges from 0.01-0.05 albedo). Results further indicate that uncertainty should be factored into L2G-based products, particularly for nominal 375 meter band data.

## 4.2 Introduction

The Visible Infrared Imager Radiometer Suite (VIIRS) instrument is a rotating telescope assembly onboard The Suomi National Polar-orbiting Partnership (NPP) satellite. The Suomi NPP mission will provide VIIRS data from several agency sources. NOAA provides a core set of Environmental Data Records (EDRs) primarily for the existing community of weather-prediction and hazard-response users who require rapid turn-around time. NASA Land Product Evaluation and Analysis Tool Element (Land PEATE) further processes data continuity and climate quality data [Justice et al., 2013]; currently released at beta-quality, the data will eventually be processed at full production quality for the broader user community. Part of the MODIS and VIIRS sensor-to-sensor continuity is the application of a consistent geolocation process, ensuring that the VIIRS data produced by the NASA

Land PEATE will continue to geolocate data in the same manner as MODIS data [Justice et al., 2013], and perpetuating the assumption of a fixed pixel size based on the subdivision of a 30 arc second sinusoidal grid providing a seamless global grid with resolutions of $\approx$ 231m, $\approx$ 463m, and $\approx$ 926m [Wolfe et al., 2002].

The VIIRS instrument is expected to continue the important legacy of climate and environmental measurements established by several predecessor missions [Justice et al., 2013], including the Advanced Very High Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) [Justice et al., 2011] missions. This continuity between missions and protocols enables the VIIRS-derivative science products to benefit from the expertise of established science teams, standard science quality products, and validation strategies [Justice et al., 2013, Morisette et al., 2002] previously developed for the MODIS missions, while obviating the imperative to fully understand mitigation strategy impacts and effects on the accurate use of VIIRS data.

Composed of 22 spectral and thermal bands, which are split between imagery bands (I-Bands) representing a nominal 375m ground sample distance (GSD), and moderate resolution bands (M-Bands) representing a nominal 750m GSD [Wolfe et al., 2013], the VIIRS instrument, like the MODIS instrument [Barnes et al., 1998], suffers growth in pixel ground sample distance as a function of the view zenith angle (VZA) of its sensor, with sample interval size increasing farther from nadir [Schueler et al., 2013]. To control this growth of pixels, the VIIRS instrument employs rectangular detectors with the smaller dimension aligned in the along-scan direction combined with an along-scan pixel aggregation scheme [Baker, 2014, Wolfe

et al., 2013], and introduces a deletion scheme for the bow tie effect (redundant data resulting from overlapping scans), seen at increasing VZA [Wolfe et al., 2012].

The pixel aggregation scheme performed onboard the satellite [Wolfe et al., 2013] averages three pixels for along-scan regions from nadir to scan angle 31.72°, two pixels for regions between 31.73° and 44.86°, and native observations are used for the remaining data (44.87° to 56.28°). It is important to note the difference between the scan angles used to define the aggregation zone here and the larger VZA used throughout this paper. The relationship between the VZA and the scan angle can be described by VZA = $\arcsin(\frac{R+H}{R}\sin(scan\ angle))$, where $R$ is Earth's radius and $H$ is the satellite's altitude, for a spherical Earth model. The corresponding VZA for the aggregation zones are nadir to 35.75° (3-pixel), 35.75° to 51.75° (2-pixel) and 51.75° to 67.48° (1-pixel). Each detector is ≈131m in the short dimension to accommodate known pixel growth along-scan, thus providing a footprint of ≈393m, 617m, and 800m for each of the 3x1, 2x1 and 1x1 aggregation zones respectively (Figure 4.1a). The breakpoints between aggregation zones along the scan, and the overall effect of pixel growth, are illustrated in Figure 4.1b.

On-orbit reports of spatial performance indicate that the aggregation scheme is effective [Lin et al., 2013, Wolfe et al., 2012, 2013], presenting significant improvement in constraining the spatial footprint along-scan compared to MODIS (co-plotted in Figure 4.1b). I-bands results show some variability in pixel sizes along scan, generally ranging within one-tenth of a pixel, with pixel size at the edge of scan still greater than 2 nominal pixels. Some bands show even greater variability, likely resultant from problems finding and processing adequate ground targets [Lin

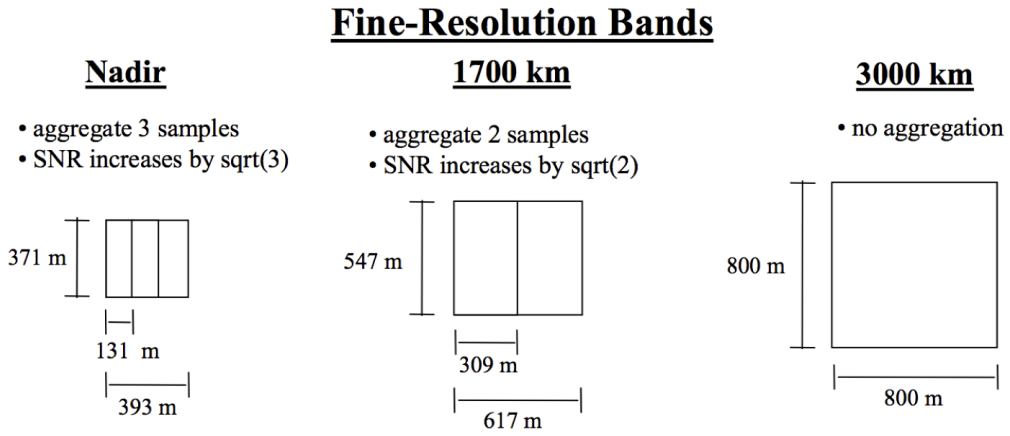
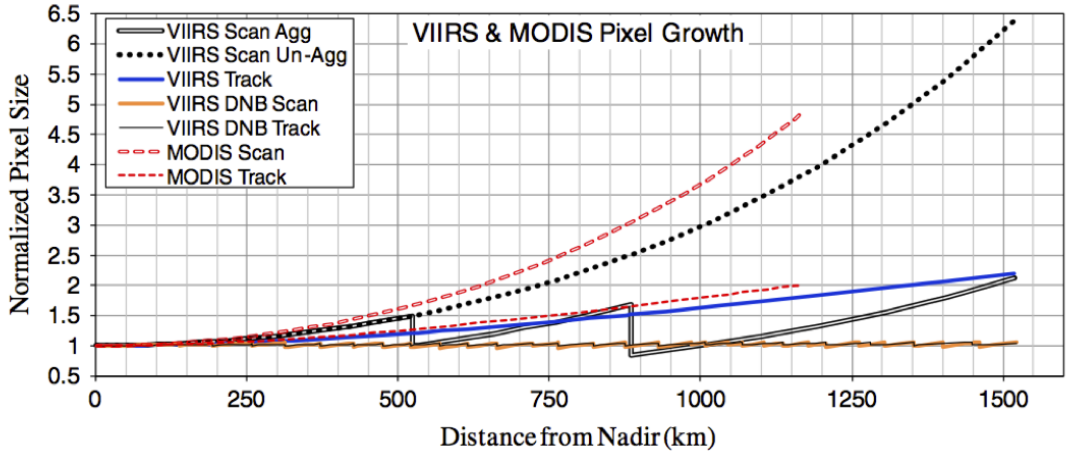



**Figure 2.2 VIIRS detector footprint aggregation scheme for building "pixels." Dimensions shown are approximate.**

(a) [Baker, 2014, Figure 2.2] illustrates the rectangular detector and along-scan aggregation scheme on the VIIRS instrument. Fewer detectors are needed to cover the same space as the distance from nadir increases, and pixel growth is constrained to ≈800m by the aligning the short edge of the rectangle along scan.




(b) [Wolfe et al., 2013, Figure 4] illustrates aggregation zone impact on pixel size growth. MODIS data was co-plotted as a reference. The overall pixel growth is constrained, particularly compared to MODIS. However, there remains growth in pixel size to more than 2 pixels by the edge of scan.

et al., 2013]. On-orbit analysis of VIIRS data geolocation accuracy indicate that the location accuracy of these data is reported to be $\approx$ 70m [Wolfe et al., 2013], which is consistent with published MODIS accuracies [Wolfe et al., 2002].

The geolocation process developed for MODIS introduced a method of pixel assignment by defining the proportion of integrated signal response during a sample interval overlapping a grid cell, called *obscov* [Wolfe et al., 1998, Yang and Wolfe, 2001]. Grid pixels are assigned the value of those observations with the highest *obscov* values, thereby preserving the original pixel value for each cell (similar to nearest neighbor), and accounting for the increasing sample interval size as associated with geometrically-increasing GSD. Observations are often shifted and/or assigned multiple times because of the mismatch between GSD and grid pixels. The VIIRS 375m I-Bands are assigned using *obscov* to the 463m grid and the 750m M-Bands are assigned to the 926m grid; in this way, the data are expected to achieve a better observation-to-grid assignment than MODIS data have previously achieved [Tan et al., 2006].

The methods and results of [Montaño, 2015] demonstrate the level of uncertainty associated with MODIS observation geometry and geolocation. There is a general expectation amongst the user community that the on-board improvements to the VIIRS instrument will provide a more consistent spatial signal across the range of GSD. To test this, a simulated series of VIIRS-like gridded broadband albedo were generated for the same study areas as the [Montaño, 2015] paper.

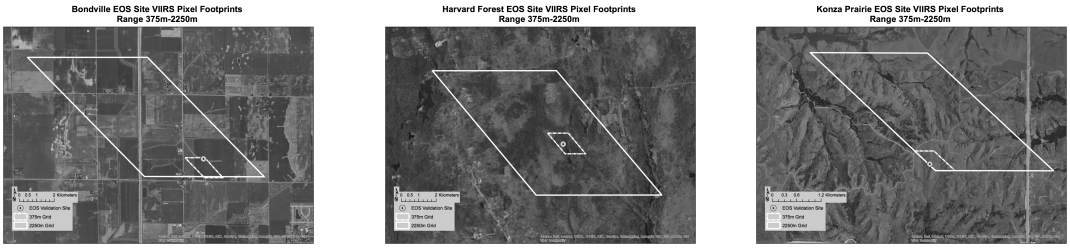
The key characteristics of the Jornada LTER semi-arid grass/scrubland site in the southwestern United States are seasonally-invariant vegetation and multi-scale

homogeneity of land cover. The key characteristics of the Konza Prairie LTER grassland site in the midwestern United States are minimally seasonally-variant vegetation and multi-scale homogeneity of land cover. The key characteristics of the Harvard Forest LTER deciduous forest site in the northeastern United States are significantly seasonally-variant vegetation and spatial heterogeneity of land cover. The key characteristics of the Lake Tahoe clear water lake site in the western United States are significantly invariant land cover type (water), nominal-scale homogeneity, and incorporation of adjacent land cover types (vegetation) at coarser scales. The key characteristics of the Bondville agricultural site in the midwestern United States are seasonally-variant vegetation and significant spatial heterogeneity of land cover type (field size, crop type, crop season).

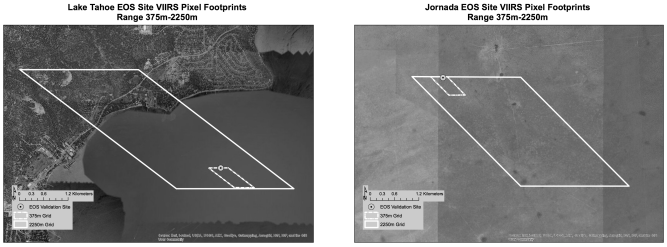
The site characteristics respective of the geolocation grid aggregation scheme from 375m to 2250m for all the sites are illustrated in Figure 4.2, and demonstrate the land features incorporated at each given stage of spatial aggregation to coarser resolutions.

## 4.3 Synthetic Broadband Albedo

The data and methods follow and extend [Montaño, 2015], thus a time series of synthetic broadband albedo data were generated from the same 11 individual Landsat TM images used in that study (Table 4.1). The use of the same scenes allows for direct comparison with the MODIS results, with the addition of VIIRS observation geometry the only new variable.




(a) Bondville, IL (b) Harvard Forest, MA (c) Konza Prairie, KS




(d) Lake Tahoe, NV (e) Jornada LTER, NM

Figure 4.2: The site characteristics respective of the geolocation grid aggregation scheme for the study areas.

The same broadband albedo was calculated following methods described by Brest and Goward [1987] for non-snow land cover using the Landsat 5 TM images calibrated to top-of-atmosphere reflectance. The Landsat broadband albedo conveniently uses a linear combination of near infrared and visible red band, which are comparable to VIIRS 375m I-Bands 1 and 2. Any difference in albedo calculations from lack of atmospheric correction of bandwidth differences can be dismissed, since absolute albedo accuracy is not the focus of this paper.

## 4.4 Simulated VIIRS Geometries

The VIIRS instrument improves on the spatial performance of the MODIS instrument by employing a rectangular detector with the short edge aligned along

Table 4.1: EOS Validation Sites and Landsat TM data used in the study

| Site Name | Lat | Lon | Path/Row | Date | Land Cover |
|---|---|---|---|---|---|
| Bondville, IL | 40.00 | -88.29 | 22/32 | 2010-04-14<br>2010-07-03<br>2010-11-08 | Broadleaf Cropland |
| Harvard Forest LTER, MA | 42.54 | -72.18 | 13/30 | 2008-04-26<br>2008-06-12<br>2006-10-13 | Broadleaf Forest |
| Konza Prairie LTER, KS | 39.08 | -96.56 | 28/33 | 2011-03-10<br>2011-07-01<br>2011-09-02 | Grassland/Cereal Crop |
| Jornada LTER, NM | 32.60 | -106.86 | 33/37 | 2011-04-30 | Shrubland/Woodland |
| Lake Tahoe, NV | 39.17 | -120.104 | 43/33 | 2009-09-21 | Needleleaf Forest/Clear Water |

scan. However, the instrument itself utilizes a scanning mirror that creates the same triangular shape of the MODIS Line Spread Function (LSF) prior to aggregation [Nishihama et al., 1997], imposing a similar neighborhood contribution (12.5% / 75% / 12.5%) to the current pixel being scanned [NOAA/STAR VIIRS SDR Team, 2013]. Therefore, the methods to simulate MODIS data described by [Montaño, 2015] can be utilized to simulate VIIRS data for this study.

The process of generating the synthetic VIIRS follows [Montaño, 2015] to replicate the triangular LSF of a VIIRS observation pixel using a triangular weight-

ing scheme. Individual VIIRS pixel footprints are determined following methods described by [Tan et al., 2006, Appendix B], using latitude, longitude and known instrument geometry including scan angle and sensor azimuth from the input geolocation file [NOAA/STAR VIIRS SDR Team, 2013].

A further improvement to spatial accuracy for VIIRS is the inclusion of native-scale geolocation data (NPP_IMFT_L1). VIIRS geolocation data are produced and distributed in both 375m and 750m resolutions, unlike MODIS MOD03 geolocation files that are only produced at 1km resolution and require interpolation for application in finer resolution data; the new geolocation data remove potential location error from interpolation.

The along-scan pixel aggregation scheme described previously occurs after the instrument finishes a scan. Thus, the simulated pixels from the previous step are aggregated and averaged according to the scan angle (SA) breakpoints described earlier using Equation 4.1. Again, note the corresponding aggregation zones using VZA are nadir to 35.75° (3-pixel), 35.75° to 51.75° (2-pixel) and 51.75° to 67.48° (1-pixel).

$$Obs_{agg} = \begin{cases} Obs_{i,i+2}/3 & \text{if } SA \leq 31.72^\circ \\ Obs_{i,i+1}/2 & \text{if } 31.72^\circ < SA \leq 44.86^\circ \\ Obs_i & \text{if } SA > 44.86^\circ \end{cases} \tag{4.1}$$

VIIRS I-Band geometries were simulated for a hypothetical $\approx$ 30 day period using daily NPP_IMFT_L1 geolocation data intersecting the site (Table 4.2). The

decision was made to generate simulated data at the finest available pixel size using 375m VIIRS I-Band geolocation data to remain consistent with the 250m MODIS data used in [Montaño, 2015]. Further, the M-Bands follow the same geometries and aggregation at exactly twice the size as I-Bands [Baker, 2014, NOAA/STAR VIIRS SDR Team, 2013], thus the results presented in this work can be expected to retain a similar structure for M-Bands.

The daily VIIRS gridded surface reflectance product from the NASA Land-PEATE is expected to follow the MODIS Surface Reflectance (MOD09GQ) processing steam, which selects the highest quality individual pixels from multiple orbits based on quality criteria including overall observation quality, maximum *obscov* and minimum VZA [Vermote et al., 2011]. However, the simulated data in this study are considered to have equal observation quality and *obscov* was not generated. Therefore, only the minimum VZA criteria was applied to select individual daily observations.

In this paper, the reprojection of swath data, maintained in a WGS84 geographic projection, uses a nearest neighbor pixel assignment to approximate the *obscov* method used in actual VIIRS product generation. The assignment difference between the resampling methods is expected to be small and is consistent with tools like the LPDAAC MODIS Reprojection Tool Swath [LPDAAC User Services, 2015]. Two swath-to-grid resolutions were generated using the nearest neighbor approach, a nominal 375m grid and a 500m grid, both using the existing MODIS geolocation sinusoidal grid to determine which is more representative of the swath pixel. The 375m grid resolution is meant to be representative of the expected detector GSD

Table 4.2: NPP_IMFT_L1 time periods used for simulating data

| | Season | Begin Date | End Date |
|---|---|---|---|
| Bondville | Spring | 2013-04-14 | 2013-05-13 |
| | Summer | 2013-07-03 | 2013-08-02 |
| | Fall | 2013-11-08 | 2013-12-05 |
| Harvard Forest | Spring | 2013-04-26 | 2013-05-28 |
| | Summer | 2013-06-13 | 2013-07-15 |
| | Fall | 2013-10-13 | 2013-11-13 |
| Konza Prairie | Spring | 2013-03-11 | 2013-04-07 |
| | Fall | 2013-09-02 | 2011-09-28 |
| | Summer | 2013-06-30 | 2013-08-02 |
| Jornada LTER | Spring | 2013-04-30 | 2013-06-02 |
| Lake Tahoe | Fall | 2013-09-21 | 2013-10-25 |

resolution of the VIIRS instrument.

VIIRS I-Band data used in coarser resolution products are aggregated using the weighted average of all observations intersecting a given coarse resolution cell determined by the *obscov* values for the input pixels. For this study, a simpler method was applied assuming the 375m grid pixel is an equal subdivision of any given coarse pixel between 750m-2250m. Each resolution is calculated using integer factors from the original 375m grid data, with the simple mean value calculated from all pixels falling into the new, coarser pixel. Finally, the pixel value for the grid cell containing the site was extracted for each gridded scene in the series.

Although the process is not completely consistent with the production code used for VIIRS land products, the process does provide an approximation of the variability. The results of the simple method demonstrate a similar effect of data smoothing through the averaging of data to coarser resolution even though the method of aggregation differs.

### 4.4.1 VIIRS Aggregation Zone Efficacy Results

Montaño [2015] reported small, though potentially significant, geometric influence for spatially homogeneous land covers like Harvard Forest or Konza Prairie. Therefore, the geometric influence using VIIRS should be likewise small, and show some improvement. However, the spatially heterogeneous Bondville demonstrated high levels of uncertainty, and will therefore be the main focus.

An example of simulated VIIRS swath data from the Bondville, IL site during

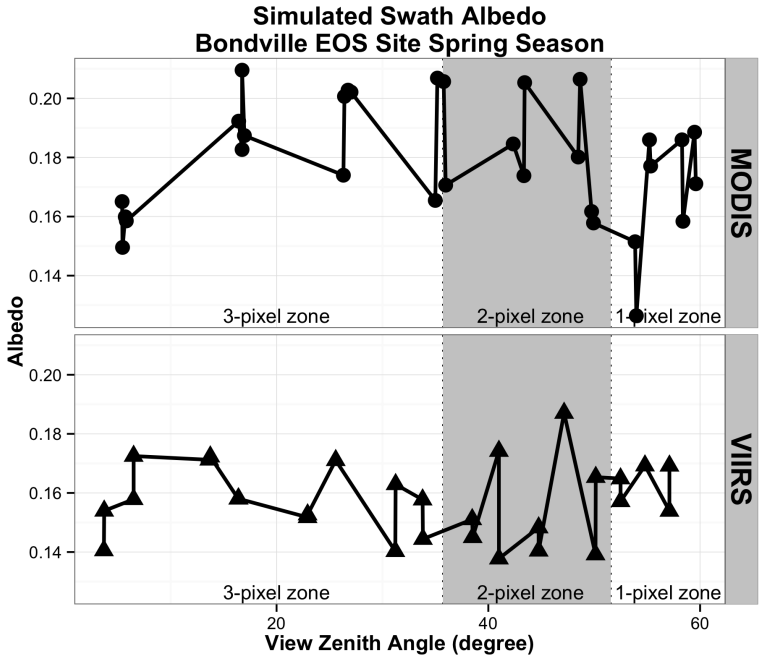



Figure 4.3: MODIS (top) and VIIRS (lower) albedo plotted as a function of VZA for the Bondville site, with the aggregation zones for reference. The range of values for VIIRS is improved over MODIS, 0.05 vs. 0.08, and the stability of the signal across the swath is somewhat smoother than MODIS.

the Spring period in Figure 4.3 (lower plot) shows the effectiveness of the aggregation scheme when compared to MODIS data (upper plot) using the same Landsat scene. The pattern of variability in the VIIRS albedo across the range of VZA suggests the aggregation scheme is helpful to stabilize the GSD. The overall range of albedo is reduced to 0.05 for VIIRS compared to 0.08 for MODIS. The difference in albedo resulting from the GSD for each of three aggregation zones, illustrated in Figure 4.4, indicate a difference in albedo $\approx 0.02$.

The introduction of the pixel aggregation scheme for VIIRS as a means to achieve a more consistent GSD along scan is an important improvement to the sensor design. Schueler et al. [2013] describe the benefits of the "constant-resolution"

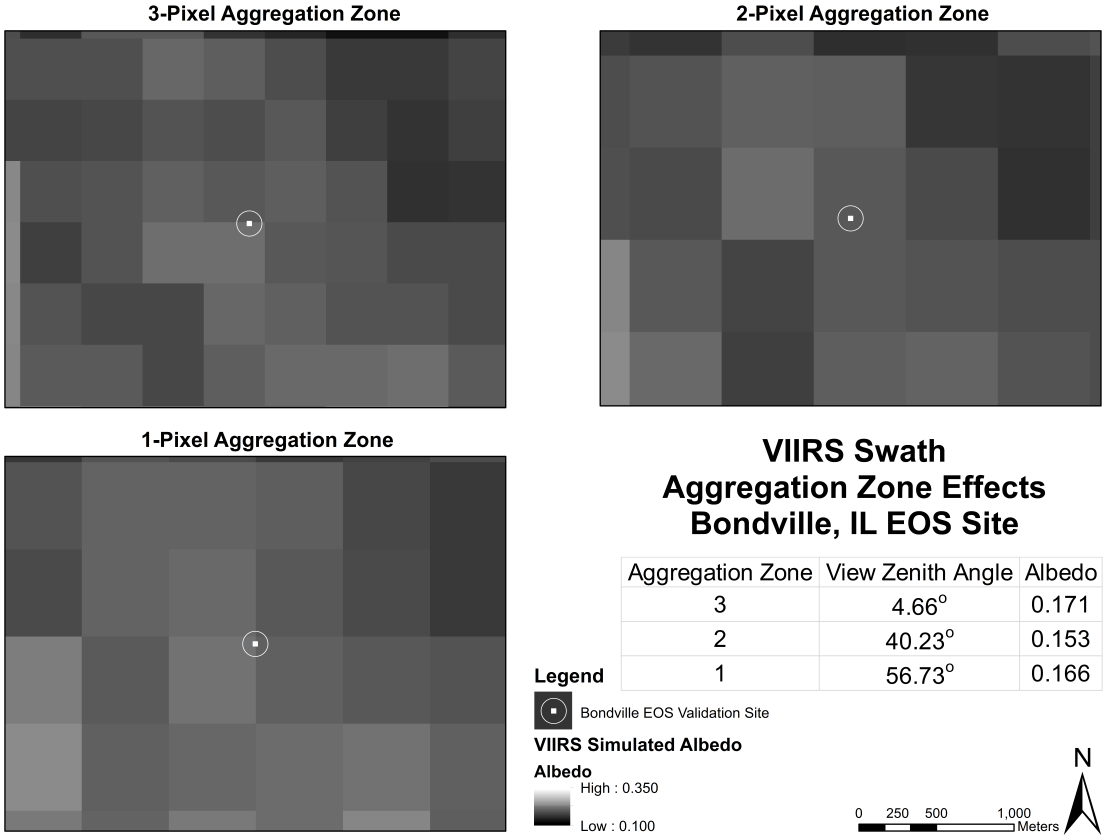



Figure 4.4: The effect of aggregation zones on simulated albedo. The three scenes illustrate the aggregation of 3-pixels, 2-pixels, and 1-pixel aggregation from a range of view zenith angles from ≈4 to ≈50 degrees on the same sample location. The scenes indicate a change in albedo of ≈0.02 from just the geometry and aggregation prior to the geolocation process.

pixels implemented in VIIRS, and provide several visual examples of improvements. The VIIRS pixel aggregation scheme, like any mathematical average, is sensitive to outlier values in the sequence being averaged. Polivka et al. [2015] investigate the effects of the VIIRS pixel aggregation on occurrence of radiance saturation in thermal M-Bands. The authors conclude saturation occurs at significantly lower levels than published limits for the 2 and 3-pixel aggregation zones resulting from the averaging of increased radiance in the smaller IFOV near-nadir pixels into a single pixel when compared to unaggregated 1-pixel results.

A series of one-way ANOVA tests were conducted to test for any significant difference between the mean albedo measurements for each aggregation zone. The tests are only conducted for the swath, 375m and 500m grid data, as these are the nominal grid resolutions derived from the original swath; data aggregated to coarser resolutions are not considered for this analysis. One-way ANOVA assumes data being tested are independent both within and between each aggregation zone (group). Data for this study are independent in that only a single pixel value is retrieved for each geometry, and each observation can only be in a single aggregation zone.

One-way ANOVA further assumes groups are normally distributed and have equal within-group variance. Group data for each site/season combination were first visually inspected for normality using a series of Quantile-Quantile normal (QQnorm) plots. Data for those groups indicating non-normal distributions were transformed to approximate normal distributions as necessary to meet model assumptions, using either log or power transformations accordingly. Finally, a Bartlett

test [Bartlett, 1937] was run for each group to test for equal variance between the groups being compared. Further transformations (either power or log functions) were applied to those few groups demonstrating unequal variance until model assumptions were met. Assumptions of normality were again tested for those transformed groups in case the underlying distributions changed, though no further transformations were required.

The results for the homogeneous sites, Harvard Forest, Konza Prairie, and Jornada, demonstrated no significant difference between aggregation zones. This result is not surprising since a homogeneous land cover presents a more spatially continuous signal for any given GSD regardless of aggregation zone. The results from the spatially heterogeneous Bondville site did demonstrate a marginal difference in the mean albedo between zones for the Spring and Summer periods, though the Bondville Fall period does not show any significant differences. The differing incorporation of land cover for each GSD that is clearly evident in the swath results above offers an explanation for the differing means between aggregation zones. The Lake Tahoe site was omitted due to the extremely low albedo measurement from the water feature at nominal resolution.

The results for the Bondville Spring period reveal a significant difference in aggregation zones at the 90% confidence level or better ($p < 0.1$). The results for the Bondville Summer period also reveal significant differences in aggregation zone, but at increased confidence ($p < 0.05$). The need to meet model assumptions was particularly important considering the unequal number of observations for each aggregation zone in the simulation (an unbalanced design) consequent of the orbit,

which the end user has no control over. The results presented here have both small sample sizes and unbalanced design, which, when combined, can lead to falsely significant differences. Thus, a Tukey HSD (honest significant difference) post-hoc test was applied for comparing multiple groups that remains useful under an unbalanced design [Tukey, 1949]. Table 4.3 summarizes the results of the pairwise Tukey test.

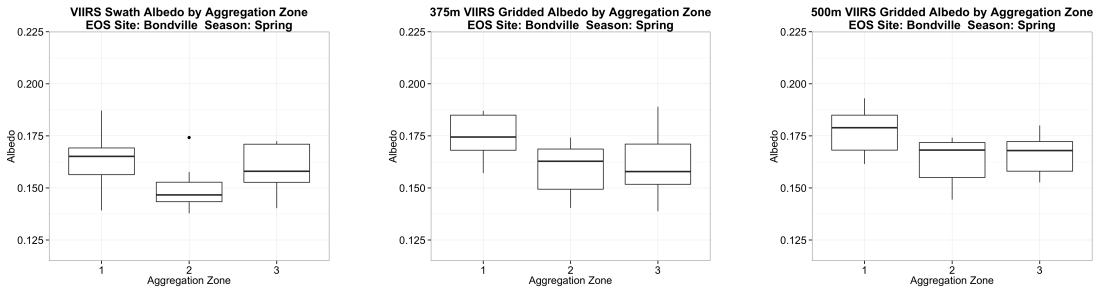
The Bondville Spring period showed significant differences between aggregation zones 1 - 2 for the swath data, and aggregation zones 1 - 2 and 1 - 3 were significantly different in both the 375m and 500m grid data. Analysis of coincident MODIS data using the same aggregation zones for comparison did not reveal any significant differences between zones for either the swath or 250m grid data.

Post-hoc analysis for the Bondville Summer period indicate differences lie between zones 1 - 3 and 2 - 3 for both swath data and 375m grid. However, only aggregation zones 1 - 3 indicated a difference in the 500m data. MODIS results reveal significant differences for both swath and 250m data, with post-hoc analysis indicating differences in swath zones 1 - 2 and 1 - 3 and 250m zones 1 - 3.

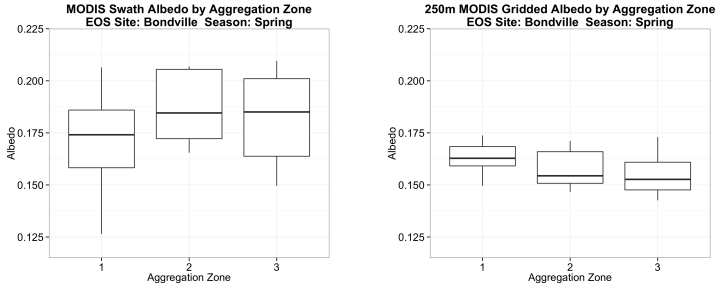
Figure 4.5 is a series of boxplots with the distribution of observed albedo measurements for the Bondville site in each aggregation zone of the swath, 375m grid and 500m grid results. Visual analysis confirms the results from the ANOVA with apparent difference between the distributions of zone 1 and 2 data, though visually the difference should carry between zones 2 and 3 as well. The difference between zones 1-2 and 1-3 for both 375m and 500m grid data is visually confirmed. MODIS swath data for the same site show larger distributions than the VIIRS data,

Table 4.3: Summary of ANOVA Post-hoc Pairwise Results

| | Agg Zone | p-value |
|---|---|---|
| VIIRS Spring Swath | 1 - 2 | 0.0857 |
| VIIRS Spring 375m | 1 - 2 | 0.0687 |
| VIIRS Spring 375m | 1 - 3 | 0.0316 |
| VIIRS Spring 500m | 1 - 2 | 0.0232 |
| VIIRS Spring 500m | 1 - 3 | 0.0787 |
| VIIRS Summer Swath | 1 - 2 | 0.0577 |
| VIIRS Summer Swath | 2 - 3 | 0.0515 |
| VIIRS Summer 375m | 1 - 3 | 0.0064 |
| VIIRS Summer 375m | 2 - 3 | 0.0867 |
| VIIRS Summer 500m | 1 - 3 | 0.0044 |
| MODIS Summer Swath | 1 - 2 | 0.0260 |
| MODIS Summer Swath | 1 - 3 | 0.0213 |
| MODIS Summer 250m | 1 - 3 | 0.0310 |




(a) Spring VIIRS Swath (b) Spring VIIRS 375m (c) Spring VIIRS 500m



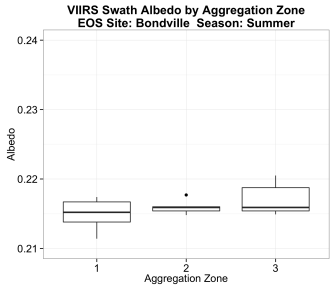

(d) Spring Swath MODIS (e) Spring MODIS 250m

Figure 4.5: Distribution of albedo by aggregation zone for the Bondville site Spring period. MODIS results are added for comparison.

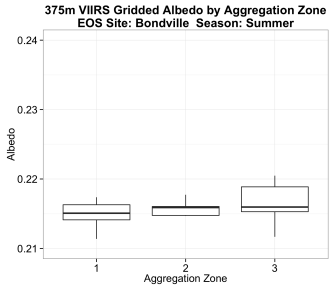
indicating the spatial performance of the VIIRS instrument is helping constrain the GSD to a smaller footprint.

Figure 4.6 is a series of boxplots for the Bondville Summer period. The distributions in the 3-pixel aggregation zone are larger than the other two in all series, which visually confirms the differences found in the ANOVA. The visually significant differences between the 1-pixel and other aggregation zones in the MODIS results clearly demonstrate that the VIIRS spatial performance improvements are particularly effective in constraining the GSD far off-nadir.

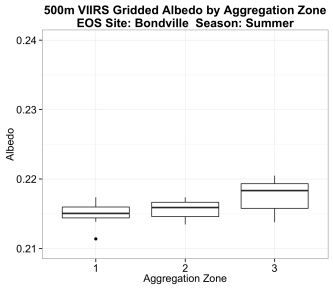
The different distributions between swath and grid datasets, particularly in the Spring, is likely a result of an inconsistent swath-to-grid pixel assignment. The



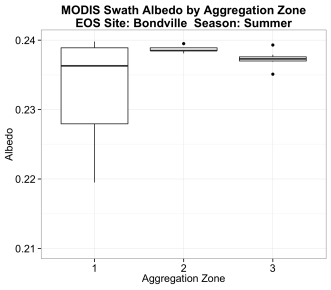

(a) Summer VIIRS Swath



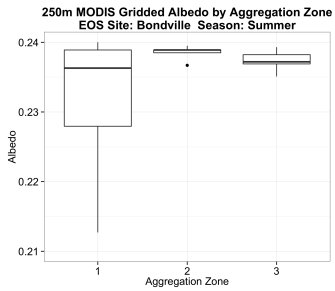

(b) Summer VIIRS 375m




(c) Summer VIIRS 500m




(d) Summer Swath MODIS




(e) Summer MODIS 250m

Figure 4.6: Distribution of albedo by aggregation zone for the Bondville site Summer period. MODIS results are added for comparison.

method used to extract swath pixel values used a nearest neighbor approach to select the swath pixel center closest to the site coordinates in Table 4.1. The grid pixel value is extracted using the same method, but in this case the grid pixel is the result of a separate nearest neighbor assignment of swath pixels prior to extracting the value, which is not necessarily the same swath pixel used for the swath analysis. For example, if two swath pixels intersect a grid cell, only the swath pixel closest to the center of the grid pixel will be chosen, but this is not necessarily the swath pixel closest to the site coordinates.

The results indicate the choice of grid resolution is not necessarily a factor in the results. The grid results are the same for Spring, but different for Summer, which indicates the underlying land cover is more influential than the choice of nominal grid resolution. However, the argument can be made for using the 500m grid in removal of difference between aggregation zones 2–3 for Summer. The persistent difference between 1-pixel and 3-pixel, and to a lesser extent the 1-pixel and 2-pixel aggregation zones in the grid results suggest the averaging that occurs in the 2 and 3-pixel is effective in smoothing the GSD, making them more consistent between these zones, but the unaggregated zone behaves differently from the lack of additional smoothing.

## 4.5 VIIRS Uncertainty Gridded Results

Montaño [2015] groups results into discrete 8-day periods to replicate what could be expected from a discrete temporal composite product using only the best

per-pixel observations (e.g. cloud-free with minimum VZA) for a given composite period. Following [Montaño, 2015], this study continues the use of range of measured albedo ($maximum\ value - minimum\ value$) as a means to assess influence from geometric effects of the changing GSD. Holding albedo constant, the range of values in any given group of VIIRS geometries remains related to the changing observation footprint as the GSD and pixel composition changes from orbit-to-orbit; observation footprints at the same location made from different VZA will necessarily be composed of differing signal contribution from adjacent space. The pixel averaging in different aggregation zones further influences the final observation. The results from replicating this grouping for the Jornada and Bondville sites using simulated VIIRS data illustrate the contrast between spatially homogeneous and heterogeneous land covers.

The Jornada site is both spatially homogeneous and seasonally-invariant, and the results in Figure 4.7 indicate a relatively constant range of albedo uncertainty, $< 0.02$, for all resolutions. The discrete analysis of the Jornada data demonstrates some minor variability between groups at 375m, and indicates aggregation of 2x to 4x (750m-1150m) can mitigate the variability to $\approx 0.01$. However, the coarsest 2250m resolution demonstrates an *increase* in inter-group variability, similar to that seen in the 375m data, indicating a limit to the aggregation before some contrasting adjacent land cover influences the result. These results along with similar results from Konza Praire suggest a level of uncertainty from geometric effects can only be mitigated to a certain extent, with some residual level retained and possibly increasing despite the smoothing of the signal.

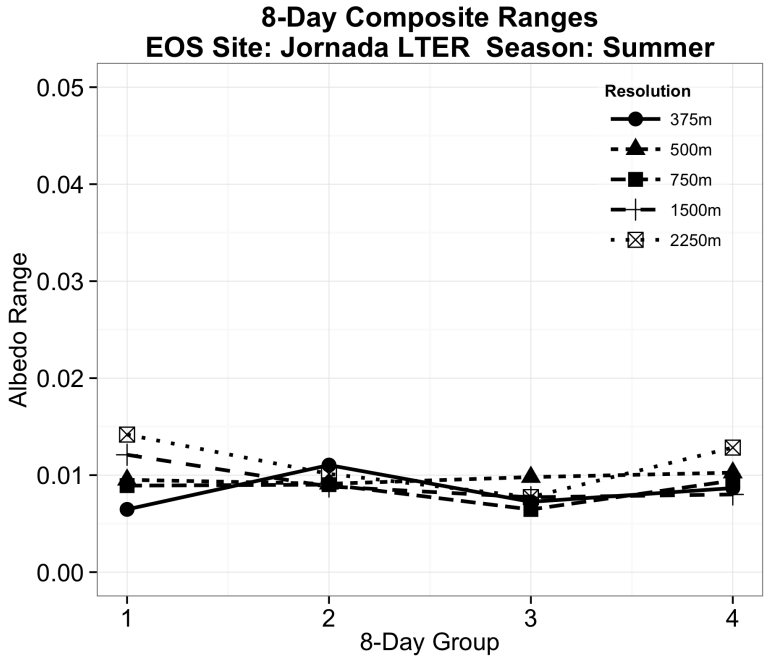



Figure 4.7: Discrete 8-day groups of simulated albedo for the Jornada site.

The series of plots in Figure 4.8 replicate the discrete analysis for all seasons at the Bondville site, and demonstrate the influence of a seasonally-variant vegetation and spatially heterogeneous land cover type. The 375m and 500m data in the Spring periods demonstrate the highest levels of uncertainty, and sizable reduction is achieved through aggregation to coarser resolutions. The high levels and trends of uncertainty in the Spring period suggest the pixel aggregation scheme only stabilizes the signal to a limited extent for spatially heterogeneous land covers. The results suggest care is needed when using observations from different geometries during a given composite period, particularly when using nominal resolution data.

The Summer period shows similar patterns and levels of uncertainty for all resolutions, which suggests the matured canopy is creating a more spatially homo-

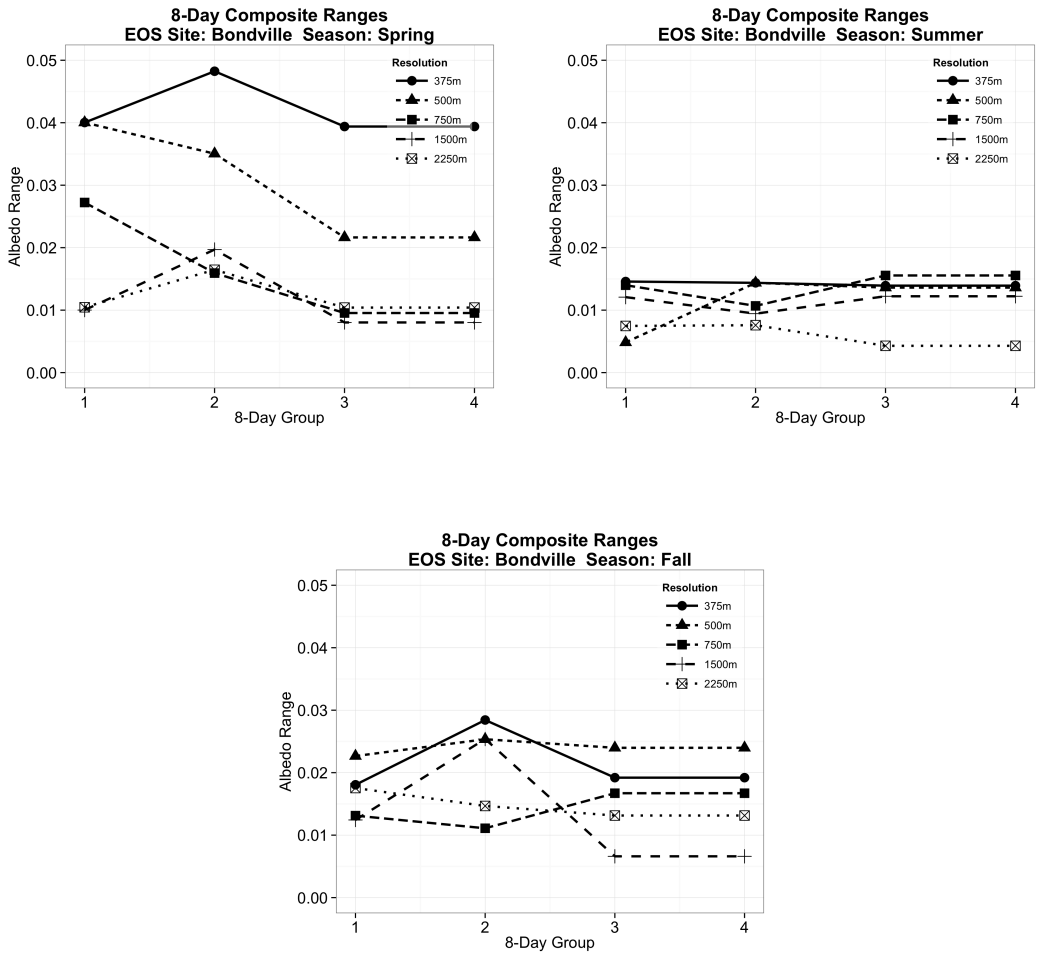



Figure 4.8: Discrete 8-day groups of simulated albedo for the Bondville site.

geneous land cover, stabilizing the signal, though some overall reduction is seen in the 1500m and 2250m resolution. The Fall period shows patterns of reduction at inconsistent resolutions; some finer resolution simulations demonstrate better reduction than coarser resolutions, indicating field size is a factor in choosing the best scale to represent spatially heterogeneous land cover.

## 4.6 Discussion and Conclusions

The objective of this research was to investigate the important improvements to the spatial performance of the VIIRS instrument. The findings inform users of VIIRS land products of the level of uncertainty added by the distortion inherent in VIIRS data due to off-nadir viewing, using albedo as a meaningful physical variable. This paper is in no way a validation of model accuracy of the VIIRS albedo or any other standard VIIRS product.

The efficacy of the pixel aggregation scheme was examined as part of this study. A series of ANOVA tests were conducted to test for significant differences between albedo measurements in each aggregation zone. No significant difference was found between aggregation zones for homogeneous land covers (i.e. Harvard Forest, Konza Prairie, Jornada). However, the results from the spatially heterogeneous Bondville site did demonstrate a marginal difference in the mean albedo between zones for several seasonal periods. The primary difference lies between the 1-pixel (unaggregated) and 3-pixel zones. The difference can be attributed to the sensitivity of mathematical averages to outlier values, which in this case is the differing incorporation of land cover for each GSD. Alternatively, averaging can also introduce a smoothing effect to heterogeneous data that is not present in the 1-pixel zone.

Like MODIS, many VIIRS science products are generated as discrete temporal composites, using only the single best quality per-pixel observation (e.g. cloud-free and/or minimum view zenith angles) for a given composite period. The results

from discrete 8-groups of data simulating temporal compositing, demonstrate the influence of geometric artifacts through differing levels of uncertainty between periods, and suggest gridded data from different GSDs are not necessarily comparable without accounting for the uncertainty between the GSDs, regardless of the pixel aggregation scheme. That is, a grid pixel from an observation near-nadir is not comparable with the same grid pixel (e.g. from a different temporal composite period) with an observation from a far off-nadir GSD without first accounting for the uncertainty from the different observation geometries.

Spatially homogeneous land cover provides a relatively stable, though potentially significant, level of uncertainty for all resolutions ($\approx$ 0.01 albedo), which in effect presents a "noise floor" for the degree of accuracy possible. Spatial heterogeneous land covers can be smoothed through aggregation of grid pixels to similar levels of uncertainty to accommodate the spatial variability. However, aggregating nominal resolutions to increasingly coarse resolutions has limits as was shown in the Jornada example when a similar level of variability was seen in the nominal 375m and 2250m results. The results suggest a grid resolution of 1150m (4x nominal) provides a consistent reduction of uncertainty for both spatially homogeneous and heterogeneous land covers.

The range of errors presented in these simulations, $\approx$0.01–0.05, offer a significant contribution to error when satisfactory levels of error in various albedo models and products range from $\pm$0.05 to $\pm$0.02 [Baker, 2014, Liu et al., 2009, Oleson et al., 2003]; as the entire error can be explained by the geometric variability informing the model. Using the same benchmark climate model sensitivity of +0.01

albedo corresponding to $-0.05Wm^{-2}$, and $-0.03$°C cooling [Matthews et al., 2003, Myhre and Myhre, 2003] from [Montaño, 2015], the VIIRS results contribute between $\approx -0.05Wm^{-2}$ and $-0.25Wm^{-2}$ to radiative forcing estimates, and between $-0.03$°C and $-0.15$°C to model uncertainty when using nominal resolution data.

An important consideration in this analysis is the consistency of pixel assignment. Sites like Bondville, with agricultural fields smaller than a given GSD, will present different portions of those fields with every orbit at finer resolutions. The pixel assignment, whether using *obscov* or nearest neighbor like this paper, will vary based on the location and size of the of the swath pixel relative to the grid from orbit-to-orbit; pixel assignment is improved by the along-scan aggregation scheme, with a more consistent GSD. However, off-nadir GSD growth is still a factor, since Figure 4.1b indicates a pixel growth of up to 2 times nominal pixel size by the end of scan, including pixel aggregation. Thus these land cover types benefit from a coarser scale to smooth the signal through aggregation.

Some applications or products might tolerate levels of uncertainty when using nominal resolution VIIRS data. Specifically, applications that do not require a high level of precision, binary pixel classification (e.g. crop/non-crop) for example, can still utilize the finest resolution data without significant reduction of result accuracy. Products incorporating multiple GSD to estimate a single measurement around the mean value might also utilize nominal resolution data and achieve desired product accuracy. Physical variables or analysis assuming a stable spatial signal through time, per-pixel comparison from two different 8-day composites for example, and requiring a high level of accuracy or measurement sensitivity can benefit from adding

a geometric uncertainty factor based on the known variability of the GSD to increase the precision.

The use of the sinusoidal grid for geolocation is a carry-over from the MODIS era. Campagnolo and Montaño [2014] report the effective resolution of 250m gridded MODIS data across a range of VZA actually varies from 344m and 835m, and behaving in a swath-like relationship to VZA. The authors further conclude the difference in alignment of the observation space and the sinusoidal grid leads to a location-dependent spatial resolution in locations farther from the central meridian of the projection at at increasing latitude. The application of an alternate projection is recommended for those regions, and can be extended to the application of the sinusoid grid to VIIRS as well since the orbits and spatial performance are similar. The need for a consistent global dataset and the level of acceptable uncertainty should be weighed prior to application of these data using this projection. Future research of alternate data processing methods should also investigate the efficacy of other equal area projections that can improve the swath-to-grid alignment over the sinusoidal projection and investigate a more representative resolution.

The MODIS instrument has proven to be a valuable contribution to the on-going monitoring and quantification of climate change, and the VIIRS instrument and follow on missions are expected to continue this tradition. As with MODIS data, the problem illustrated here is the difference between the sampling of variable observation space by the imaging system and resampling, or gridding, that image space to a fixed space with known resolution. While the efforts employed to mitigate geometric distortions with VIIRS data, primarily along-scan, on-board pixel

aggregation, lead to improved published spatial accuracies of the *imaging system*, the same improvements do not necessarily translate to improved spatial performance when the observations are assigned to a fixed grid space. The results presented here and in conjunction with similar work on data from MODIS provide the framework to assess the continuity of measurements between the two instruments.

# Chapter 5: Summary and Conclusions

## 5.1 Synthesis of Research

The role and importance of Earth observations in the form of satellite imagery will continue to grow as research seeks to better understand how the Earth is changing, for example as a result of land use or climate change. Polar-orbiting satellite instruments have played a role in making Earth observations and providing critical evidence of how and where change is manifesting. Wide field-of-view instruments like MODIS and VIIRS continue the important legacy in assessing change by providing global data with high temporal (near daily) frequency. However, in the satellite system design there is a tradeoff between high temporal frequency and spatial resolution and accuracy.

The work presented in this dissertation answers the question, "What is the actual resolution of MODIS data?" The as-launched MODIS mission configurations (Aqua and Terra) were the result of a compromise in requirements for the sake of cost and reduced complexity [Barnes et al., 1998]. In an early concept stage, two instruments were considered: a tilting scanning mirror for atmospheric and ocean observations, where the geometric distortions are of limited influence in continuously variable conditions (MODIS-T), and a nadir push-broom sensor for land observa-

tions (MODIS-N).

A push-broom sensor relies on the motion of the platform in the track direction and has no moving scanning mechanism. The primary spatial benefit of a push-broom sensor is having a fixed pixel size as a result of the optical camera rather than the pixel size determined by the GSD of the scanning system. However, the optical system in a push-broom sensor is more complex than those of the scanning mirror, which partially led to the final decision of merging the MODIS mission requirements into a single converged instrument.

Final configuration and actual resolution notwithstanding, the MODIS instruments (onboard the Aqua and Terra satellites) have provided invaluable information for the land sciences community for more than a decade. The VIIRS instrument, part of the recently launched Suomi NPP mission, will continue this important legacy into the future [Justice et al., 2011, 2013]. This dissertation addressed three fundamental aspects of the MODIS/VIIRS continuum with respect to spatial resolution.

An effective resolution of MODIS L2G data was characterized across a variety of geometries for the full image system, based on the amount of spatial signal contributing to the grid pixel in Chapter 2. An understanding of the effective resolution of these gridded data is crucial to their informed use in higher-order science products. The results clearly indicate the data from variable GSD assigned to a uniform grid continue to include the same radiometric impact of pixel growth as swath data, bringing into question the spatial consistency of the data.

Chapter 3 built on the findings of Chapter 2 by expressly evaluating the spatial

consistency of simulated measurements through the range of VZA. The findings of Chapter 3 illustrate a level of variability in the geometric distortion directly related to the variable view geometry of the MODIS instrument, that is sufficient to influence the derivation of albedo, an important input to climate models. Reducing the resolution of the data and aggregating the gridded data can achieve some level of mitigation of the noise associated with geometric distortions.

Chapter 4 was similar in approach to Chapter 3, but investigated important improvements to spatial performance of the VIIRS instrument. The results in Chapter 4 suggest there are improvements in the representativeness of the data compared to the MODIS results presented in Chapter 3. However, there remain problems in the consistency of measurements despite the improvements, which also require reduction of resolution to achieve a stable spatial representation.

### 5.1.1 Establish an Effective Resolution of MODIS L2G Data

Chapter 2 presented models to estimate the ground space actually contributing to a given geolocated grid cell in both along-scan and along-track directions, thus describing the full 2-Dimensional Point-Spread-Function, for the full range of VZA. The results model the MODIS instrument on the Aqua and Terra satellites independently, since the sensors have opposite orbits (ascending vs. descending) at different times of day (AM vs. PM).

The results indicate a much coarser spatial resolution ($>$ 300m along scan) than those assumed by the 250m-product description. Native swath observations

demonstrate a geometric growth of pixel size at increasing VZA, resulting in a pixel 4.8 times the nadir size by the end of scan. Level 2 gridded data (L2G) data continue to present off-nadir observations composed of signal contributions for a much larger area of ground surface, despite being resampled into a fixed grid pixel space. The crux of the problem lies in the misleading notion of reporting a nominal ***detector*** size at nadir, which inadequately describes the geographic space observed with increasingly off-nadir view angles. Simply put, a convenient 250m global grid, approximating the sensor resolution at nadir, is not representative of an observation GSD that exceeds 1km off-nadir.

The results further uncovered an artifact in the gridding process causing a variation in resolution as data progress away from the central meridian of the sinusoidal projection. The artifact is the result of inconsistent alignment of swath observations to grid cells as a function of instrument alignment (azimuth) to the grid at increasing distance from the central meridian and changes with latitude, which causes the misrepresentation of the PSF orientation in the grid.

The primary site used for analysis in Chapter 2 is in Holland, near the central meridian of the projection, which forms an angle with the grid aligned mostly along scan with the swath. The primary site thus provides a meaningful general estimate of the resolution for the grid for a range of VZA. Chapter 2 provides two examples, one for a Barrow, Alaska site and the other a site in Changbaishan, China, showing the opposite extremes of the problem. The Alaska site demonstrates a near perpendicular alignment of the sensor along scan to the grid, leading to an increased blurring and therefore a decreased resolution. The site in China, however, forms a

shallow 10.96° angle between the scan direction and the grid column, aligning the grid with the ***track*** direction. The effect of the different alignment is an underestimation of ground space contribution to the signal as the short side of the ground space in the track direction is aligned with the along scan grid. Ultimately, effective pixel resolutions are not consistent globally despite having the same VZA, a fact that end users of gridded data should be aware of when using L2G MODIS data that are far from the central meridian of the sinusoidal projection.

Two examples illustrate the outcome for end users. Users of data for Oregon in North America, a location far off center from the central meridian, can expect effective along-scan resolutions between 460m-1169m for nadir and 55° VZA respectively. Whereas users of data for St. Petersburg in Russia, closer to the center of the projection, can expect effective resolutions between 317m-690m for the same VZA range. User applications requiring or assuming consistent spatial accuracy across a range of VZA for multiple locations will suffer from the effective resolution misrepresented by a nominal 250m product description.

### 5.1.2 Estimate the Level of Measurement Variability Associated with Variable View Geometries

The level of geometric distortion resulting solely from the variable GSD inherent in MODIS data is quantified using multiple land cover types and seasonality. The results from Chapter 2 introduced an inconsistent, location-dependent effective GSD that is far different than the nominal grid resolutions suggested by standard

product descriptions (i.e. 250m, 500m and 1km). Chapter 3 uses this information to generate a series of gridded broadband albedos for several sites to test the spatial stability of measurements at a given site through changing geometries. Albedo was selected for this study as it is a physical quantity that can be estimated from satellite observations. The overall results show a range of albedos between $\approx 0.01$ to $\approx 0.045$ attributed solely to the variable GSD. Many MODIS standard products are generated as ***discrete*** temporal composites, using only the single best quality per-pixel observation (e.g. cloud-free and/or minimum view zenith angles) for a given composite period - usually 8-16 days. Results from discretely grouping simulated data into 8-day groups to replicate a temporal composite product, indicate geometric variability persists between groups, particularly at finer resolutions (i.e. 250m and 500m).

Simulated results indicate error from geometric distortion alone that is sufficient to influence radiative forcing in climate models. Specifically, results from several studies using simulated land cover class albedos [Matthews et al., 2003, Myhre and Myhre, 2003] describe the dependency between albedo accuracy and model uncertainty. Together, the studies describe a general relationship +0.01 albedo corresponds to $-0.05 W m^{-2}$, and $-0.03^{\circ} C$ cooling, which would equate to a range of negative radiative forcing between $\approx -0.05 W m^{-2}$ and $-0.2 W m^{-2}$, and decreased temperatures between $-0.03^{\circ} C$ and $-0.12^{\circ} C$ for the MODIS results. The variability is also significant when compared to the reported $\pm 0.05$ accuracy of the MODIS MCD43 albedo product when validated using ground measurements [Liu et al., 2009]. The authors attribute a decrease in accuracy to a mismatch in spatial

scale between ground measurements and MODIS observation footprints and also to site heterogeneity, supporting the conclusions in Chapter 3. However, the variability reported in Chapter 3 is independent of error from correlation to ground validation data and solar zenith effects, indicating the geometric variability a large contributor to the stated accuracy.

Mitigation of the geometric variability is (partially) achieved through smoothing of data by aggregation of pixels to increasingly coarse resolutions. Pixel resolutions of at least 1000m, 4x the nominal 250m detector size, are required to mitigate the geometric influence from most land cover types tested. However, complete removal of the noise from variable geometries is not achieved, even at 2000m or 8x nominal resolution. The need for aggregation to mitigate artifacts resulting from the gridding process is consistent with previous findings using gridded data [Tan et al., 2006]. An alternative method of averaging all swath pixels informing a grid space for any given grid resolution showed a significant reduction in albedo variability at all resolutions. The average resampling demonstrated a similar minimum variability at 4x the nominal 250m, but the difference between finer resolutions is smaller than that of the standard grid aggregation.

The current MCD43 BRDF/Albedo is generated at 500m-grid resolution rather than the 250m geometries used for the simulation [Schaaf et al., 2014]. The applicability of the Chapter 3 results is significant despite this difference for several reasons. First, the Red and NIR bands in the MCD43 product are nominal 250m MODIS bands that are aggregated to 500m. The results for 250m data clearly demonstrate a resolution at least twice that of the MCD43 product (i.e. 1km) is necessary to

stabilize the geometric error.

Further, the geometries for ***all*** MODIS detector resolutions function the same, but at different scales. Therefore, even the simple assumption of doubling of the effective GSD for 250m grid data (e.g. $250m \approxeq 340m \therefore 500m \approxeq 680m$) suggests a 500m-grid product is not representative of the variable space being observed through geometry. While MODIS products are designed to accurately present the portion of the space observed in a given grid cell, there is no specific benefit from using a smaller grid cell size than a cell size more representative of the maximum variable observation footprint for example; a resolution increasing the average *obscov* to greater than 30% [Tan et al., 2006] would be beneficial to stabilize the spatial signal. Users only considering nadir observations might find 250m grid data sufficient to represent the surface being observed. However, users of multi-temporal, and therefore multi-geometry, data would benefit from the data being generated at the more representative resolution of a coarser product.

Spatial homogeneity is an important factor when considering observation stability across the range of GSDs. Relatively homogeneous land cover provides a stable, though potentially significant, level of uncertainty for all resolutions. In contrast, spatially heterogeneous land covers can be smoothed through aggregation to similar levels of uncertainty to accommodate the spatial variability. The spatial configuration of the surface, influence from field size for example, is also a factor when considering the most appropriate scale for MODIS data. Sites like Bondville, with agricultural fields smaller than a given GSD, will present different portions of those fields with every orbit at finer resolutions. Thus representation of these land

cover types benefit from a coarser scale to stabilize the signal. Conversely, some care should be taken not to aggregate the data to an extent that no longer represents the land cover of interest, as the Lake Tahoe results demonstrated with two disparate and contrasting adjacent land cover types (forest and water). Ultimately, all land cover types would benefit from standard products being generated at a 1km resolution.

### 5.1.3 Evaluate the Effectiveness of Methods to Improve Spatial Accuracy

The spatial performance of VIIRS data was tested in the same manner as MODIS data, which showed some improvement for the VIIRS gridded product, though not as much as expected, given the design adopted to mitigate pixel growth across the swath. The VIIRS instrument employs two primary mechanisms in an effort to constrain the known geometric growth of pixels off-nadir, rectangular detectors and onboard along-scan pixel aggregation.

VIIRS uses rectangular detectors in place of the square detectors used for MODIS, with the smaller dimension aligned in the along-scan direction [Baker, 2014, Wolfe et al., 2013]. Added to the improvement in detector shape is a pixel aggregation scheme meant to present a similar effective GSD across the range of VZA. Three pixels are averaged for along-scan regions from nadir to 31.72°, two pixels are used between 31.73° and 44.86°, and native observations are used for the remaining data (44.87° to 56.28°). Each detector is ≈131m in the short dimension to

accommodate known pixel growth long scan, thus providing a footprint of ≈393m, 617m, and 800m for each of the 3x1, 2x1 and 1x1 aggregation zones respectively.

The result of the pixel aggregation scheme represents a significant improvement in constraining the spatial footprint along scan as compared to MODIS, but the pixel size at the edge of scan stills grows to more than twice the nominal detector GSD. The aggregation is performed onboard the satellite [Wolfe et al., 2013], thus an understanding of the impact and effect is integral to the use of data from VIIRS. Visual interpretation of data might not easily uncover any inconsistency between aggregation zones.

The efficacy of the pixel aggregation scheme was examined as part of this study. A series of ANOVA tests were conducted to test for significant differences between albedo measurements in each aggregation zone. No significant difference was found between aggregation zones for homogeneous land covers (i.e. Harvard Forest, Konza Prairie, Jornada). However, the results from the spatially heterogeneous Bondville site did demonstrate a marginal difference in the mean albedo values ($< 0.015$) between zones for several seasonal periods. The primary difference lies between the 1-pixel (unaggregated) and 3-pixel zones. The difference can be attributed to the sensitivity of mathematical averages to outlier values, which in this case is the differing incorporation of land cover for each GSD. Alternatively, averaging can also introduce a smoothing effect to heterogeneous data that is not present in the 1-pixel zone.

Like MODIS, many VIIRS science products are generated as discrete temporal composites, using only the single best quality per-pixel observation (e.g. cloud-free

and/or minimum view zenith angles) for a given composite period. The results from discrete 8-groups of data, simulating (8 day) temporal composite data, demonstrate the influence of geometric artifacts through differing levels of uncertainty between periods. The results suggest gridded data from different GSDs are not necessarily comparable without accounting for the uncertainty between the GSDs, regardless of the pixel aggregation scheme. That is, a grid pixel from an observation near-nadir is not comparable with the same grid pixel (e.g. from a different temporal composite period) with an observation from a far off-nadir GSD without first fitting an error term to the measurements accounting for the uncertainty from the different observation geometries. Alternatively, both observations can be aggregated to a comparable, coarse grid resolution.

The range of errors presented in these simulations, ≈0.01–0.05, offer a significant contribution to error when satisfactory levels of error in various albedo models and products range from ±0.05 to ±0.02 [Baker, 2014, Liu et al., 2009, Oleson et al., 2003]; as the entire error can be explained by the geometric variability from the data informing the model. Using the same benchmark climate model sensitivity of +0.01 albedo corresponding to $-0.05 W m^{-2}$, and $-0.03^{\circ} C$ cooling [Matthews et al., 2003, Myhre and Myhre, 2003] from Chapter 3, the VIIRS results contribute between $\approx -0.05 W m^{-2}$ and $-0.25 W m^{-2}$ to radiative forcing estimates, and between $-0.03^{\circ} C$ and $-0.15^{\circ} C$ to model uncertainty when using nominal resolution data. Further, the VIIRS Albedo EDR is expected to provide a threshold albedo accuracy of ±0.05, but has a specification of ±0.025 [Baker, 2014] for the 1km product.

Spatially homogeneous land cover provides a relatively stable, though poten-

tially significant, level of uncertainty for all resolutions ($\approx$ 0.01 albedo), which in effect presents a "noise floor" for the degree of accuracy possible. Aggregating data to coarse resolutions can result in a more consistent uncertainty especially for spatial heterogeneous land covers. However, limitations to aggregating nominal resolutions to increasingly coarse resolutions was shown in the Jornada example, when a similar level of variability was seen in the nominal 375m and 2250m results. The results suggest a grid resolution of 1150m (4x nominal) provides a consistent reduction of uncertainty for both spatially homogeneous and heterogeneous land covers.

## 5.2 Future Research

The findings from the research presented in this thesis would indicate several avenues for further research. The first is to explore alternatives to the continued use of the sinusoidal grid for MODIS and VIIRS data. Chapter 2 reports the effective resolution of 250m gridded MODIS data across a range of VZA actually varies between 344m and 835m along scan. The chapter further concludes the difference in alignment of the observation space and the sinusoidal grid leads to a location-dependent spatial resolution in locations farther from the central meridian of the projection. The inconsistent alignment of data will carry forward with VIIRS since the NPP satellite is on the same ascending orbit as the Aqua satellite.

The exploration of an alternate map projection, which would reduce the alignment artifacts reported in Chapter 2, is an important first step in future research. Specifically, equal area projections such as the Lambert Azimuthal or Albers Equal

Area projections offer possible alternatives for regional projections that define a local center meridian, and the longitude range can be constrained to maintain an optimal swath-to-grid alignment angle. For example, the results in Chapter 2 suggest regions can be defined at a continental scale with a central meridian defined for each, while still retaining a consistent alignment angle. The switch to a regionally defined projection should be both consistent in projection (e.g. use only Lambert Azimuthal) and pixel size (e.g. 1km) such that comparisons can be made between data from different regions, and also allow for systematic processing in a similar manner to the current MODIS/VIIRS processing stream. The existing sinusoidal projection should be retired and replaced once a suitable replacement is identified and improvements measured and demonstrated.

The current L2G product, so-called L2G-lite, no longer contains the data layers necessary to reverse the sinusoidal projection back to swath [Vermote et al., 2011]. Thus, any alternative projection must occur at the initial swath-to-grid stage, since simply reprojecting existing sinusoidal grid data will necessarily incorporate the uncertainty demonstrated throughout this thesis. A tool like the existing LPDAAC MODIS Reprojection Tool Swath [LPDAAC User Services, 2015] offers alternate projections and is useful for the investigation of alternate projections or for end users interested in processing their own data for a specific study area. The MODIS Science Team recognized the utility of global datasets for the modeling community by generating products using the Climate Model Grid (CMG) at 0.05 degree (5600-meter) spatial resolution [Justice et al., 1998]. The coarser spatial resolution of the CMG reflects limitations to data complexity climate model applications are capable

of handling, and is also an indication of the need to spatially degrade the data for interoperability between instruments like AVHRR [Pedelty et al., 2007].

Coincident with the exploration of alternative projections is the application of a representative grid size for these data. Results from all chapters of this dissertation demonstrate disconnect between the sinusoidal grid resolutions and the physical space described by the variable GSD of the instrument. The number 4x the nominal detector size has been demonstrated as an effective scale factor for both MODIS and VIIRS data. Therefore, local projections combined with a representative grid cell 4x the nominal detector size (i.e. 1000m and 1500m for MODIS and VIIRS respectively) are potentially useful for products using off-nadir views, such as albedo.

A consistent method of mitigation through aggregation is particularly important when considering data continuity between sensors. Continuity of measurements requires integration of multiple characteristics including viewing geometries and spatial resolutions, which require a common factor such as aggregated pixel size to achieve a common representative spatial footprint between sensor data. NASA's Land Long Term Data Record (LTDR) is a multi-sensor data continuity product that already integrates data from the AVHRR sensor with MODIS to produce a long term data set at 0.05° spatial resolution, illustrating the need to coarsen data to accommodate differences in spatial resolutions [Pedelty et al., 2007].

The results from Chapter 3 also suggest an alternative average resampling technique is more effective than the maximum *obscov* for at least MODIS data. The approach was attempted for VIIRS data in Chapter 4, but presented poor results and was omitted. One possible explanation is the pixel aggregation scheme is already

averaging the data with only a marginal new contribution to the final LSF, thus the average resampling is simply averaging the average.

Further, the need to provide a 1:1 swath observation to geolocation grid pixel assignment (e.g. nearest neighbor or maximum *obscov* resampling) is not necessary when one considers that both the MODIS and VIIRS instruments already present a variable GSD as part of the imaging system. That is, no one "nadir" pixel is comprised exclusively of nadir view signal and necessarily integrates signal from neighboring space (i.e. 12.5% / 75% / 12.5%) and the GSD grows along scan. Averaging all the observations informing a projected grid space (possibly weighted using *obscov* or other relative contribution weight) would be more informative than attempting to achieve a 1:1 swath-to-grid assignment and should be further explored.

Many MODIS and VIIRS products are generated as *discrete* temporal composites, using only the single best quality per-pixel observation (e.g. cloud-free and/or minimum view zenith angles) for a given composite period - usually 8-16 days. Data in these composites ultimately suffer from the same geometric distortion problem, as daily data in that a seemingly spatially continuous scene nonetheless represents a spatially discontinuous surface comprised of observations from any number of disparate GSD.

Figure 5.1 illustrates the disparate days of data and VZA used in a typical discrete composite MOD13 NDVI product. The composite day of year seen in the figure is an illustration of the false assumption of spatial continuity of data within a composite scene. There are 8 days worth of data in the scene that are not necessarily spatially contiguous. Further, and of particular import to this study, the range of

VZA in the composite is far larger than one might expect. Specifically, the MOD13 product series applies a minimum VZA criterion in the selection of best quality pixels [Solano et al., 2010]. However, there are clearly circumstances overriding this criteria (e.g. a limited number of cloud-free days) making the use of less desirable observation days (e.g. those with large VZA) a necessity. Stabilization of geometric influence for the range of GSD present in temporal composite scenes is important to improve accuracy of analyses such as per-pixel classification or per-pixel comparison between periods, and should be the focus of future work.

## 5.3 Conclusion

The research presented in this dissertation illustrates the false assumption that gridded data from wide field-of-view instruments offer a consistent spatial representation of land features. Many tradeoffs are required in using remotely sensed data to understand an environmental system. In the case of MODIS and VIIRS, the value of high temporal frequency observations and the wealth of temporal detail they provide come at the cost of spatial accuracy and precision. The significance of this tradeoff for MODIS and VIIRS is understanding the space being observed, or more specifically, the contribution of signal to a given space, is unstable through changing geometry, thus negating an assumption of a consistent target in any spatial sense. The recent trend to develop finer spatial resolution gridded products from MODIS/VIIRS (e.g. the MCD43 originally produced at 1km resolution is now produced at 500m) misrepresents the data by attempting to achieve finer scale mea-

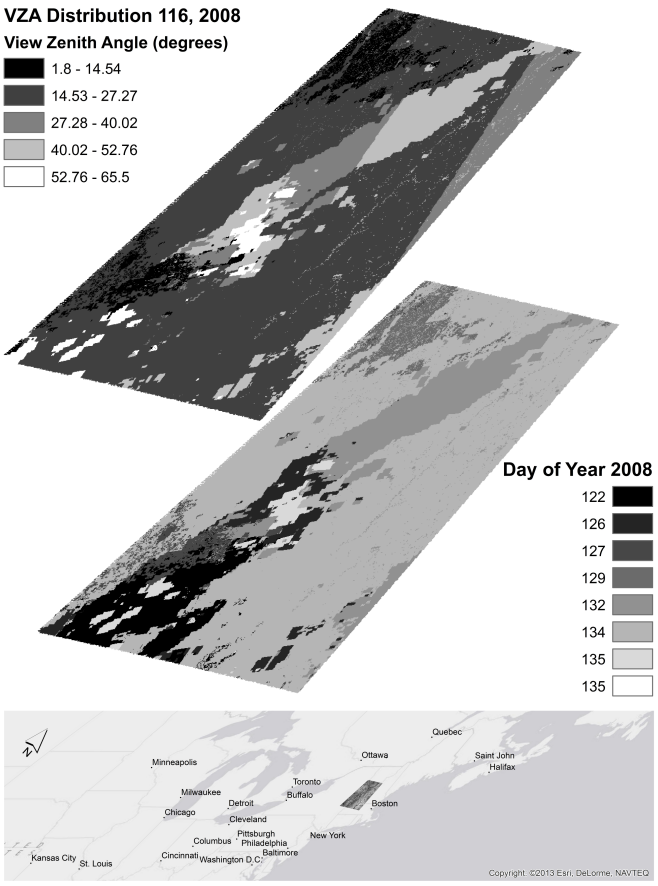



Figure 5.1: Illustration of days used in compositing a MODIS MYD13Q1 NDVI product from the April 27, 2008 (day 116) time period. The top right figure shows the view zenith angles from the day used. The lower left figure shows the day within the period that was used in the composite. The MOD13 product uses a minimum VZA as part of the quality selection, but some large angles are used despite this as a result of coincident factors such as persistent cloud contamination.

surements than the observations warrant. The results presented in this dissertation demonstrate that data from a nominal 250m detector do not produce a consistent 250m GSD, and the variable GSD is not represented by a 250m grid pixel.

While the wealth of published results in the literature remain valid despite these results, care should be taken in interpreting and extending their utility. The literature largely ignores the impact of artifacts arising from geometric distortion, or fails to breakout the uncertainty explicitly related to these artifacts when describing results. Research using data or models derived from MODIS L2G of any scale is subject to the errors reported in this dissertation.

The fact some applications or products might tolerate levels of uncertainty when using nominal resolution MODIS or VIIRS data should be noted. Specifically, applications that do not require a high level of precision, binary pixel classification (e.g. snow/non-snow) for example, can still utilize the finest resolution data without significant reduction of result accuracy. Products incorporating multiple GSD to estimate a single measurement around the mean value might also utilize nominal resolution data and achieve desired product accuracy. Physical variables or analysis assuming a stable spatial signal through time, per-pixel comparison from two different 8-day composites for example, and requiring a high level of accuracy or measurement sensitivity, can benefit from adding a geometric uncertainty factor based on the known variability of the GSD to increase the precision.

The ultimate outcome of this work is attaining an understanding of what surface features are being resolved in a given space and time by wide field-of-view instruments, and quantifying the impact of placing an inherently variable observa-

tion into a fixed space. The importance of achieving this fundamental understanding is bridging the span between actual observation and convenient, consistent packaging. The highest level of utility for data from MODIS and VIIRS can be achieved through a new projection with a representative pixel size using a resampling method that better describes the continuous overlapping observations.

# Bibliography


N. Baker. Joint Polar Satellite System (JPSS) VIIRS surface albedo algorithm theoretical basis document. Technical report, NASA, 2014.

W. L. Barnes, T. S. Pagano, and V. V. Salomonson. Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1. *IEEE Transactions on Geoscience & Remote Sensing*, 36(4):1088–1100, 1998. ISSN 0196-2892.

M. S. Bartlett. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 160(901):268–282, 1937. ISSN 00804630. URL `http://www.jstor.org/stable/96803`.

C. L. Brest and S. N. Goward. Deriving surface albedo measurements from narrow band satellite data. *International Journal of Remote Sensing*, 8(3):351–367, 1987.

J. J. Butler and J. Xiong, editors. *MODIS on-orbit spatial characterization results using ground measurements*, volume 6296 of *Earth Observing Systems XI*, 2006. Proc. of SPIE.

M. L. Campagnolo and E. L. Montaño. Estimation of effective resolution for daily MODIS gridded surface reflectance products. *Geoscience and Remote Sensing, IEEE Transactions on*, 52(9):5622–5632, Sept 2014. ISSN 0196-2892. doi: 10.1109/TGRS.2013.2291496.

K. W. Dixon, T. L. Delworth, T. R. Knutson, M. J. Spelman, and R. J. Stouffer. A comparison of climate change simulations produced by two GFDL coupled climate models. *Global and Planetary Change*, 37:81–102, 2003.

M. R. B. Forshaw, A. Haskell, P. F. Miller, D. J. Stanley, and J. R. G. Townshend. Spatial resolution of remotely sensed imagery a review paper. *International Journal of Remote Sensing*, 4(3):497–520, 1983. ISSN 0143-1161. URL `http://www.informaworld.com/10.1080/01431168308948568`.

L. Giglio. MODIS Collection 5 active fire product user's guide version 2.4. University of Maryland, College Park, 2008.

A. A. Gitelson, Y. J. Kaufman, and M. N. Merzlyak. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58(3):289–298, 1996. ISSN 0034-4257. URL <GotoISI>://A1996VV51300006.

S. N. Goward. *Encyclopedia of World Climatology*, chapter Albedo and Reflectivity, pages 32–35. Springer, 2005.

J. Hansen, M. Sato, and R. Ruedy. Radiative forcing and climate response. *Journal of Geophysical Research*, 102(D6):6831–6864, 1997.

D. Helder, T. Choi, and M. Rangaswamy. In-flight characterization of spatial quality using point spread functions. In S. A. Morain and A. M. Budge, editors, *Post-Launch Calibration of Satellite Sensors, The International Workshop on Radiometric and Geometric Calibration*, Gulfport, MS, 2003. A.A. Balkema Publishers.

C. Huang, J. R. Townshend, S. Liang, S. N. V. Kalluri, and R. DeFries. Impact of sensor's point spread function on land cover characterization: assessment and deconvolution. *Remote Sensing of Environment*, 80:203–212, 2002.

A. Huete, K. Didan, W. van Leeuwen, and E. Vermote. Global-scale analysis of vegetation indices for moderate resolution monitoring of terrestrial vegetation. *Remote Sensing for Earth Science, Ocean, and Sea Ice Applications*, 3868:141–151, 1999. ISSN 0277-786X. URL <GotoISI>://000084986600015.

L. Ji, K. Gallo, J. C. Eidenshink, and J. Dwyer. Agreement evaluation of AVHRR and MODIS 16-day composite NDVI data sets. *International Journal of Remote Sensing*, 29(16):4839–4861, 2008. ISSN 0143-1161. URL http://www.informaworld.com/10.1080/01431160801927194.

G. Joseph. How well do we understand earth observation electro-optical sensor parameters? *ISPRS Journal of Photogrammetry and Remote Sensing*, 55(1): 9–12, 2000. ISSN 0924-2716. URL http://www.sciencedirect.com/science/article/B6VF4-400X1WM-1/2/de81476af72f4a505e7d1b1fd4db873e.

C. O. Justice, E. Vermote, J. R. G. Townshend, R. Defries, D. P. Roy, D. K. Hall, V. V. Salomonson, J. L. Privette, G. Riggs, A. Strahler, W. Lucht, R. P. Myneni, Y. Knyazikhin, S. W. Running, R. R. Nemani, Z. Wan, A. R. Huete, W. van Leeuwen, R. E. Wolfe, L. Giglio, J-P. Muller, P. Lewis, and M. J. Barnsley. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience & Remote Sensing*, 36(4):1228–1249, 1998.

C. O. Justice, J. R. G. Townshend, E. F. Vermote, E. Masuoka, R. E. Wolfe, N. Saleous, D. P. Roy, and J. T. Morisette. An overview of MODIS land data processing and product status. *Remote Sensing of Environment*, 83(1-2): 3–15, 2002. ISSN 0034-4257. URL http://www.sciencedirect.com/science/article/B6V6V-46P1SXS-2/2/9deb80a73b79082c10e5a4670d78f305.

C. O. Justice, E. Vermote, J. L. Privette, and A. Sei. The evolution of U.S. moderate resolution optical land remote sensing from AVHRR to VIIRS. In B. Ramachandran, C. O. Justice, and M. J. Abrams, editors, *Land Remote Sensing and Global Environmental Change: NASA's Earth Observing System and the Science of ASTER and MODIS*, volume 11 of *Remote Sensing and Digital Image Processing*. Springer, 2011.

C.O. Justice, M. O. Román, I. Csiszar, E. Vermote, R. E. Wolfe, S. J. Hook, M. A. Friedl, C. B. Schaaf, T. Miura, M. Tschudi, G. Riggs, D. K. Hall, A. I. Lyapustin, S. Devadiga, C. Davidson, and E. J. Masuoka. Land and cryosphere products from Suomi NPP VIIRS: Overview and status. *Journal of Geophysical Research*, 118(17):9753–9765, 2013.

A. A. Lacis, G. A. Schmidt, D. Rind, and R. A. Ruedy. Atmospheric CO2: Principal control knob governing earth's temperature. *Science*, 330:356–359, 2010.

S. Liang, A. H. Strahler, and C. Walthall. Retrieval of land surface albedo from satellite observations: A simulation study. *Journal of Applied Meteorology*, 38: 712–725, 1999.

G. Lin, J. C. Tilton, R. E. Wolfe, K. P. Tewari, and M. Nishihama. SNPP VIIRS spectral bands co-registration and spatial response characterization. In J. J. Butler, X. Xiong, and X. Gu, editors, *Earth Observing Systems XVIII*, volume 8866. SPIE, 2013.

J. Liu, C. Schaaf, A. Strahler, Z. Jiao, Y. Shuai, Q. Zhang, M. Roman, J. A. Augustine, and E. G. Dutton. Validation of Moderate Resolution Imaging Spectroradiometer (MODIS) albedo retrieval algorithm: Dependence of albedo on solar zenith angle. *Journal of Geophysical Research*, 114, 2009.

LPDAAC User Services. MODIS Reprojection Tool Swath. `https://lpdaac.usgs.gov/tools/modis_reprojection_tool_swath`, February 2015.

H. D. Matthews, A. J. Weaver, M. Eby, and K. J. Meissner. Radiative forcing of climate by historical land cover change. *Geophysical Research Letters*, 30:1055–1059, 2003.

E. L. Montaño. Demonstration of uncertainty resulting from MODIS-Like geometries: an albedo case study. In Review: International Journal of Digital Earth, 2015.

J. L. Monteith and M. H. Unsworth. *Principles of Environmental Physics*. Elsevier, 2008.

A. Montenegro, M. Eby, Q. Z. Mu, M. Mulligan, A. J. Weaver, E. C. Wiebe, and M. S. Zhao. The net carbon drawdown of small scale afforestation from satellite observations. *Global and Planetary Change*, 69:195–204, 2009.

J. Morisette, J. L. Privette, and C. O. Justice. A framework for the validation of MODIS land products. *Remote Sensing of Environment*, 83:77–96, 2002.

G. Myhre. Consistency between satellite-derived and modeled estimates of the direct aerosol effect. *Science*, 325(5937):187–190, Jul 2009. ISSN 1095-9203 (Electronic); 0036-8075 (Linking). doi: 10.1126/science.1174461.

G. Myhre and A. Myhre. Uncertainties in radiative forcing due to surface albedo changes caused by land-use changes. *Journal of Climate*, 16(10):1511–1524, 2013/11/20 2003. doi: 10.1175/1520-0442-16.10.1511. URL `http://dx.doi.org/10.1175/1520-0442-16.10.1511`.

G. Myhre, D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang. Anthropogenic and natural radiative forcing. In T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 8. Cambridge University Press, Cambridge, U. K. and New York, USA, 2013.

NASA. NASA's Earth Observing System. `http://eospso.gsfc.nasa.gov/mission-category/3`, October 2014.

M. Nishihama, R. E. Wolfe, D. Solomon, F. S. Patt, J. Blanchette, A. J. Fleig, and E. Masuoka. MODIS L1A earth location algorithm theoretical basis document. Technical report, 1997.

NOAA/STAR VIIRS SDR Team. Joint Polar Satellite System (JPSS) Visible Infrared Imaging Radiometer Suite (VIIRS) Sensor Data Record (SDR) Geolocation Algorithm Theoretical Basis Document (ATBD). Technical Report Rev. A, NOAA, 2013.

K. W. Oleson, G. B. Bonan, C. Schaaf, F. Gao, Y. Jin, and A. Strahler. Assessment of global climate model land surface albedo using MODIS data. *Geophysical Research Letters*, 30(8), 2003.

R. K. Pachauri and A. Reisinger. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. In *IPCC Fourth Assessment Report: Climate Change 2007*, Geneva, Switzerland, 2007.

M. Pagnutti, S. Blonski, Cramer M., D. Helder, K. Holekamp, E. Honkavaara, and Ryan R. Targets, methods, and sites for assessing the in-flight spatial resolution of electro-optical data products. *Canadian Journal of Remote Sensing*, 36(5): 583–601, 2010.

S. K. Park, R. A. Schowengerdt, and M. Kaczynski. Modulation-transfer-function analysis for sampled image systems. *Applied Optics*, 23(15):2572–2582, 1984.

J. Pedelty, S. Devadiga, E. Masuoka, M. Brown, J. Pinzon, C. Tucker, D. Roy, Junchang Ju, E. Vermote, S. Prince, J. Nagol, C. Justice, C. Schaaf, Jicheng Liu, J. Privette, and A. Pinheiro. Generating a long-term land data record from the AVHRR and MODIS instruments. In *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*, pages 1021–1025, 2007.

T. N. Polivka, E. J. Hyer, J. Wang, and D. A. Peterson. First global analysis of saturation artifacts in the VIIRS infrared channels and the effects of sample aggregation. *IEEE Geoscience and Remote Sensing Letters*, 12(6):1262 – 1266, 2015.

F. Rojas, R. A. Schowengerdt, and S. F. Biggar. Early results on the characterization of the Terra MODIS spatial response. *Remote Sensing of Environment*, 83:50–61, 2002.

D. P. Roy, J. S. Borak, S. Devadiga, R. E. Wolfe, M. Zheng, and J. Descloitres. The MODIS land product quality assessment approach. *Remote Sensing of Environment*, 83:62–76, 2002.

C. P. Ruiz and F. J. A. Lopez. Restoring SPOT images using PSF-derived deconvolution filters. *International Journal of Remote Sensing*, 23(12):2379–2391, 2002.

C. Sagan, O.B. Toon, and J. Pollack. Anthropogenic albedo changes and the earth's climate. *Science*, 206:1363–1368, 1979.

V. V. Salomonson, B. Guenther, W. L. Barnes, N. J. Therrien, and R. E. Murphy. Early instrument performance results from the Terra/Moderate Resolution Imaging Spectroradiometer (MODIS). In *Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International*, volume 3, pages 943 –946 vol.3, 2000.

C. Schaaf, J. Liu, F. Gao, Z. Jiao, Y. Shuai, and A. Strahler. Collection 005 change summary for MODIS BRDF/Albedo (MCD43) algorithms. `http://landweb.nascom.nasa.gov/QA_WWW/forPage/C005_Change_BRDF.pdf`, 2014.

C. B. Schaaf, F. Gao, A. H. Strahler, W. Lucht, X. W. Li, T. Tsang, N. C. Strugnell, X. Y. Zhang, Y. F. Jin, J. P. Muller, P. Lewis, M. Barnsley, P. Hobson, M. Disney, G. Roberts, M. Dunderdale, C. Doll, R. P. d'Entremont, B. X. Hu, S. L. Liang, J. L. Privette, and D. Roy. First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, 83:135–148, 2002.

R. A. Schowengerdt. *Remote Sensing: Models and Methods for Image Processing*. Elsevier, Amsterdam, 3rd edition, 2007. ISBN 978-0-12-369407-2.

R. A. Schowengerdt, C. Archwamety, and R. C. Wrigley. Landsat thematic mapper image-derived MTF. *Photogrammetric Engineering & Remote Sensing*, 51(9): 1395–1406, 1985.

C. F. Schueler, T. F. Lee, and S. D. Miller. VIIRS constant spatial-resolution advantages. *International Journal of Remote Sensing*, 34(16):5761–5777, 2013.

G. Schurgers, U. Mikolajewicz, M. Groger, E. Maier-Reimer, M. Vizcaino, and A. Winguth. Long-term effects of biogeophysical and biogeochemical interactions between terrestrial biosphere and climate under anthropogenic climate change. *Global and Planetary Change*, 64:26–37, 2008.

R. Solano, K. Didan, A. Jacobson, and A. Huete. MODIS vegetation index user's guide. Technical report, The University of Arizona, 2010.

B. Tan, C.E. Woodcock, J. Hu, P. Zhang, M. Ozdogan, D. Huang, W. Yang, Y. Knyazikhin, and R.B. Myneni. The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions. *Remote Sensing of Environment*, 105(2):98–114, 2006. ISSN 0034-4257.

J. R. G. Townshend and C. O. Justice. Towards operational monitoring of terrestrial systems by moderate-resolution remote sensing. *Remote Sensing of Environment*, 83(1-2):351–359, 2002.

J. R. G. Townshend, C. Huang, S. N. V. Kalluri, R. S. Defries, S. Liang, and K. Yang. Beware of per-pixel characterization of land cover. *International Journal of Remote Sensing*, 21(4):839–843, 2000. ISSN 0143-1161. URL `http://pdfserve.informaworld.com/655302_915428398_713860620.pdf`.

J. Tukey. Comparing individual means in the analysis of variance. *Biometrics*, 5 (2):99–114, 1949.

E. F. Vermote, S. Y. Kotchenova, and J. P. Ray. MODIS surface reflectance user's guide. Technical report, 2011. URL `http://modis-sr.ltdri.org/products/MOD09_UserGuide_v1_3.pdf`.

D. G. Waggoner and I. N. Sokolik. Seasonal dynamics and regional features of MODIS-derived land surface characteristics in dust source regions of East Asia. *Remote Sensing of Environment*, 114:2126–2136, 2010.

R. E. Wolfe, D. P. Roy, and E. Vermote. MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *IEEE Transactions on Geoscience & Remote Sensing*, pages 1324–1338, 1998. ISSN 0196-2892.

R. E. Wolfe, M. Nishihama, A. J. Fleig, J. A. Kuyper, D. P. Roy, J. C. Storey, and F. S. Patt. Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sensing of Environment*, 83(1-2):31–49, 2002. ISSN 0034-4257. URL `<GotoISI>://000179160200004`.

R. E. Wolfe, G. Lin, M. Nishihama, K. P. Tewari, and E. L. Montaño. NPP VIIRS early on-orbit geometric performance. In J. J. Butler, X. Xiong, and X. Gu, editors, *Earth Observing Systems XVII*, volume 8510. SPIE, 2012.

R. E. Wolfe, G. Lin, M. Nishihama, K. P. Tewari, J. C. Tilton, and A. R. Isaacman. Suomi NPP VIIRS prelaunch and on-orbit geometric calibration and characterization. *Journal of Geophysical Research: Atomospheres*, 118(11):508–511, 2013.

X. Xiong and W. Barnes. An overview of MODIS radiometric calibration and characterizations. *Advances in Atmospheric Sciences*, 23(1):69–79, 2006.

X. Xiong, N. Che, W. Barnes, Y. Xie, L. Wang, and J. Qu. Status of Aqua MODIS spatial characterization and performance. In H. Shimoda R. Meynart, S.P. Neeck, editor, *Sensors, Systems, and Next-Generation Satellites X*, volume 6361. SPIE, 2006.

K. Yang and R.E. Wolfe. MODIS level 2 grid with the ISIN map projection. In *IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS 2001)*, volume 7. IEEE, 2001.