Retrieval of Arctic Vegetation Biophysical and Biochemical Properties from CHRIS/PROBA Multi-Angle Imagery Using Empirical and Physical Modelling

Blair E. Kennedy 1,* , Doug J. King 2 and Jason Duffe 1

Abstract: Mapping and monitoring of Arctic vegetation biochemical and biophysical properties is gaining importance as global climate change is disproportionately affecting this region. Previous studies using remote sensing to model Arctic vegetation biochemical and biophysical properties have generally involved empirical modelling with nadir looking broadband sensors and have typically been conducted at the field scale in one study area. Satellite hyperspectral remote sensing has not been previously investigated for retrieving leaf and canopy biochemical and biophysical properties of Arctic vegetation across multiple sites using either empirical or physically-based modelling approaches. Furthermore, multi-angle hyperspectral sensors (CHRIS/PROBA), which can provide insight into vegetation reflectance anisotropy and potentially improve vegetation parameter estimation, have also not been investigated for this purpose. In this study, three modelling approaches previously investigated with field spectroscopy data (Kennedy et al., 2020) were used with CHRIS Mode-1 imagery to predict leaf chlorophyll content, plant area index and canopy chlorophyll content across a bioclimatic gradient in the Western Canadian Arctic. Modelling approaches included: parametric linear regression based on vegetation indices (VI), non-parametric machine learning Gaussian processes regression (GPR) and inversion of the PROSAIL radiative transfer model using a look-up table approach (LUT). CHRIS imagery was acquired with −55°, −36°, 0°, +36°, +55° view zenith angles (VZA) between 2011 and 2014 over three field sites extending from the Richardson Mountains in central Yukon, Canada to the north end of Banks Island, Northwest Territories, Canada. Field measurements were acquired within several weeks of satellite acquisitions. GPR had the best model fit (mean cross-validated (cv) coefficient of determination, $r^2_{cv} = 0.61$ across all vegetation variables, sites and VZAs vs. 0.59 for the simple ratio, SR) and predictive performance (normalized root mean square error, NRMSE$_{cv} = 0.13$ vs. 0.14 for SR). The revised optimized soil adjusted VI (ROSAVI) performance was slightly poorer ($r^2_{cv} = 0.51$; NRMSE$_{cv} = 0.15$). The physically-based PROSAIL model performance poorer than all empirical models ($r^2 = 0.50$; NRMSE = 0.18). This ranking of model performance is similar to that found in the previous field spectroscopy study, where empirical model fits and predictive performance were only slightly worse. With respect to view angle performance, NRMSE varied only slightly, indicating no distinct advantage for any one VZA. Overall, strong potential has been demonstrated for empirical modelling of Arctic vegetation chlorophyll and plant area index using hyperspectral data combined with band selection/optimization procedures in the Arctic. Recently launched and future hyperspectral satellites, including next generation airborne sensors, will likely provide improvements to the model performance reported here.

Keywords: Arctic ecosystems; CHRIS/PROBA; multi-angle satellite spectroscopy; vegetation biochemical and biophysical properties; plant traits; empirical modelling; physical modelling; inversion; bioclimatic gradient

1 Landscape Science and Technology Division, Environment and Climate Change Canada, 1125 Colonel BY Drive, Ottawa, ON K1A 0H3, Canada; Jason.Duffe@canada.ca
2 Department of Geography and Environmental Studies, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1A 0H3, Canada; Doug.King@carleton.ca

* Correspondence: Blair.Kennedy3@Canada.ca
1. Introduction

The Arctic is experiencing environmental changes that are expected to have profound impacts on its vegetation-based ecosystems [1,2]. Arctic temperatures have risen by ~2 °C since the 1950s, resulting in changes in vegetation composition and species dominance [3–5], with short tundra vegetation being replaced and overgrown by shrubs and trees [3,6–10]. Satellite observations have shown that longer growing seasons (earlier onset) and enhanced greening are occurring in various Arctic regions [11–13]. However, to date there have been no studies that attempt to quantify the biophysical (e.g., plant area index) and biochemical (e.g., leaf and canopy chlorophyll content) variables (plant traits) that are related to vegetation productivity across a bioclimatic gradient using spaceborne hyperspectral sensors [14].

Currently, monitoring Arctic vegetation change poses two major challenges. (1) Arctic landscapes are characterized by multiple scales of vegetation spatial heterogeneity ranging from centimeters to kilometers [14–17]. Spatial heterogeneity and species diversity of tundra environments make it difficult to apply universal modelling and retrieval approaches for monitoring. (2) Even though small-scale plot-based vegetation studies may provide highly detailed local information (e.g., [18]), the climate and remote location make such studies logistically constrained and impractical for long-term, systematic investigation and monitoring [19,20] at ecosystem or landscape scales [21–23]. Remote sensing can provide extensive spatial and temporal coverage and has the potential to improve vegetation attribute estimation precision and accuracy over larger areas, specifically through leaf and canopy biochemical and biophysical variable retrievals [14,24–32].

Since many natural materials have diagnostic absorption and reflectance features that are only 10 nm to 20 nm wide, hyperspectral sensors (or spectroscopic sensors) often have sufficient spectral resolution for quantification of those materials based on the shape and magnitude of their associated spectra [33]. Advances in moderate resolution hyperspectral satellites, such as CHRIS/PROBA (Compact High Resolution Imaging Spectrometer on the Project for On Board Autonomy satellite) [34], the recently launched PRISMA (Hyperspectral Precursor and Application Mission) [35], or near-future sensors such as EnMAP (Environmental Mapping and Analysis Program) [36], present a significant opportunity to improve retrievals of biochemical (both pigment and non-pigment, depending upon sensor spectral range; e.g., [25]) and biophysical vegetation properties using empirical and physically-based modelling approaches.

To date, Arctic vegetation spectroscopic investigations have mostly been conducted at the field level and not with satellite data e.g., [14,16,37–47]. Very few studies have measured foliar biochemical quantities in the field (e.g., vegetation chlorophyll content [14,48] but none have integrated such data with satellite remote sensing-based modelling and retrieval. Most of the above field studies have used some form of empirical modelling typically involving parametric regression, but none have compared empirical (e.g., machine learning) and physical retrieval methods.

Arctic remote sensing studies have shown that the optical properties of tundra vegetation are complicated by numerous factors, including illumination conditions and sensor viewing geometry (e.g., view zenith angles, VZAs) [14,15,18,37–47,49–53]. Surface conditions that influence the optical properties of Arctic vegetation include: (1) the highly variable, spatial-temporal patterns of soil and surface moisture [15,18]; (2) the complex spatial patterns of vegetation species assemblages (e.g., lichens, bryophytes, graminoids and shrubs) as they relate to microtopographic variability and broader scale environmental gradients (e.g., temperature and precipitation/moisture) [15,18,54]; (3) the percent cover of photosynthetic and non-photosynthetic vegetation components (e.g., senescent vegetation, lichens, etc.) [14,15,37–39,55]; (4) shadowing due to microtopographic variation and plant leaf-canopy structure [14,18,56]; and (5) the prevalence of canopy gap fractions and the influence of soil, rock and sand on reflectance [14,37]. Even though surface features and spatial heterogeneity can confound remote sensing retrievals [15,18], knowledge of sun-sensor geometry is particularly important for retrieving consistent and repeatable estimates of vegetation parameters [53], especially in the Arctic where illumination condi-
tions are characterized by low sun angles and reported in the literature as a challenge to high-latitude vegetation cover retrievals [15,18].

The Compact High Resolution Imaging Spectrometer (CHRIS) is a sun-synchronous push-broom hyperspectral sensor that is capable of acquiring five along-track co-registered images at five separate view angles in one overpass (−55°, −36°, 0°, +36°, +55°) [34,57]. CHRIS operating in Mode-1 (M1) provides 62 band imagery covering wavelength regions between ~410 nm and ~1000 nm with spectral sampling intervals of ~6 nm to ~20 nm (depending on wavelength) [34,57]. The spectrodirectional [58] capabilities of CHRIS provide a unique platform for investigating the angular distribution of spectral reflectance (i.e., reflectance anisotropy) in relation to observation geometry and localized lighting conditions [59]. Specifically, CHRIS provides a means for investigating: (1) the angular variability of vegetation/surface reflectance as it relates to ecosystem spatial variability and sun angles e.g., [14]; (2) the performance of empirically-based vegetation biophysical and biochemical retrieval methods, such as iterative optimization-based machine learning algorithms, in relation to observation geometry e.g., [14]; (3) wavelength sensitivity/response to leaf and canopy vegetation parameters in relation to sun and sensor geometry e.g., [14,60]; and (4) the performance of directionally-based leaf and canopy physical radiative transfer models (i.e., PROSAIL; [61]). Previous studies using CHRIS have demonstrated that multi-angle hyperspectral remote sensing can improve the quality and reliability of vegetation variable/trait estimates in non-tundra ecosystem while providing greater insights into the factors affecting reflectance, such as the influence of observation-illumination geometry (e.g., shadowing) and ecophysiological parameters (e.g., pigments and leaf area) [29,30,59,60].

Multi-angle studies conducted in Arctic tundra environments are not well reported in the remote sensing literature. Using spectrodirectional data, Vierling et al. [62] showed that off-nadir spectral measurements can be used to discriminate between tussock tundra sites possessing varying degrees of woody material and that vegetation index (VI) values for tussock tundra differ depending on view angle (i.e., forward and backscatter directions). Buchhorn et al. [53] showed that VI-based estimates of biomass in Arctic tundra ecosystems are influenced by sun-object-sensor geometries, and that spectrodirectional effects are more pronounced over the rough surfaces at extreme sun-sensor geometries. However, no studies have examined the extraction of Arctic biophysical and biochemical vegetation variables from multi-angle satellite spectroscopic data. In addition, most studies have been conducted in a single study area; if such methods are to be applied over large areas, there is a need for incorporation of multiple sites representing different Arctic ecosystems and their illumination conditions into vegetation modelling studies. This paper addresses all of the above stated gaps in remote sensing of Arctic vegetation biochemical and biophysical properties.

A companion paper to this study [14], addressed these research gaps in field-based Arctic vegetation biochemical/biophysical modelling. Three modelling approaches were compared using multi-angle field spectroscopy data for predicting leaf chlorophyll content (LCC, µg/cm²), plant area index (PAI, m²/m²) and canopy chlorophyll content ([CCC = LCC × PAI], g/m²) at four field sites across a bioclimatic gradient in the Western Canadian Arctic. Both full spectra data and simulated CHRIS Mode-1 spectral bands were evaluated as preparation for the study presented here using CHRIS imagery. The modelling approaches were: (1) parametric linear regression based on two-band VIs with band selection optimization as well as the use of predefined narrowband vegetation indices; (2) non-parametric machine learning Gaussian processes regression (GPR) [63]; and (3) inversion of the PROSAIL [37] radiative transfer model (RTM) using Levenberg-Marquardt (LM) iterative numerical optimization [64] and a look-up table (LUT) approach. The modelling results obtained from the field study were used to inform several aspects of the modelling of this study. For example, GPR performed the best with the simulated and spectrally reduced CHRIS data. Of 11 two-band VIs, each with wavelengths selected through an optimization procedure, the simple ratio (SR) performed best in parametric
regression and about equally as well as GPR. Of 65 pre-defined narrowband VIs evaluated, the revised optimized soil adjusted vegetation index (ROSAVI) was determined to be the best, with slightly poorer performance than SR and GPR. Thus, these VIs and methods were retained for this remote sensing study. PROSAIL inversions did not perform as well in the field study, but were retained for this study to assess their performance at a remote sensing scale. Inversion regularization strategies obtained from the field study used to inform the physical remote sensing modelling included: the implementation of an appropriate/high-performing cost function; the addition of noise to account for sensor and environmental uncertainty; the use of the mean of multiple best solutions to account for sensor uncertainty and inversion ill-posedness; and the use of LUTs with sufficient dimensionality to characterize the measured/assumed variable parameter space of Arctic tundra canopies.

The primary objective of this study was to estimate and compare models of LCC, PAI and CCC in Arctic tundra environments representing a latitudinal gradient using empirical and physical modelling approaches implemented with multi-angle CHRIS full spatial and spectral resolution satellite data (Mode-1, M1). Models were tested at both local and global scales (i.e., individual field sites vs. all field sites pooled together).

2. Materials and Methods

The study areas and field data are described fully in Kennedy et al. [14] and Kennedy [65] and briefly summarized here.

2.1. Study Sites

Field sites included Herschel Island, Yukon (center of the nadir CHRIS acquisition, 69.58° N, −139.05° W), Banks Island, Northwest Territories (74.19° N, −120.64° W) and the Richardson Mountains, Yukon/Northwest Territories (67.05° N, −136.21° W). Due to the lack of CHRIS image acquisitions for the Richards Island, Northwest Territories field site used in Kennedy et al. [14], it was excluded from this analysis. All field sites were visited and sampled between 10 July and 4 August 2011 (Herschel Island), 2012 (Banks Island), 2013 and 2014 (Richardson Mountains) during the peak of the growing season for the Western Arctic (i.e., determined through peak NDVI values, [18]). At the regional scale, plant species diversity and morphology are related to the summer mean air temperatures found along latitudinal and altitudinal gradients [2,66]. As such, species diversity (and size/stature) decreases with increasing latitude and altitude [67], where altitudinal gradients are additionally affected by slope and aspect, duration of snow cover and sunlight hours [66,68]. Consequently, the most northerly Banks Island site had the lowest species diversity and plant height, and ground cover was more variable than the most southerly Richardson Mountains site, where canopy gaps were less frequent. Herschel Island vegetation was intermediate between the other two sites’ conditions. At each of the field sites, vegetation distribution and vegetation assemblage/community transitions were evident along moisture gradients (e.g., [66]). Spatial variation of vegetation was also observed to be related to pattered ground features, such as ice-wedge polygons, that are characteristic of periglacial regions/environments [69]. Fine spatial scale variations of vegetation, soil moisture, pH and temperature in patterned ground have been found, and in turn, vegetation biomass and thickness have been shown to affect the hydrologic, thermal and nutrient properties of soils across the study region [66,70,71].

2.2. Ground Data and Spectral Measurements

The field sampling design and measurements are given in Kennedy et al. [14] and briefly summarized here. Field-measurements were used as reference data in this study to train and validate models of reflectance, LCC, PAI and CCC derived from CHRIS reflectance data. They were taken in a systematic manner within 90 m × 90 m plots, representing approximately 3 × 3 CHRIS M1 pixels, and in subplots of 1 m × 1 m located along transects. Spectral measurements were taken with an ASD FieldSpec handheld spectroradiometer.
(Analytical Spectral Devices, Inc. Boulder, CO, USA) with a 10° fore optic at nadir, +/−36° and +/−55°, resulting in a field of view between about 26 cm and 50 cm (depending on VZA), which was much smaller than the subplot size. Five spectral measurements were taken across each subplot and then averaged into one value. For the Herschel Island site measurements were only taken at nadir as this was the first field season and logistics were being worked out. The ASD full spectrum data were convolved \[37,72\] to the 62 ~10 nm (6 nm to 20 nm actual) bandwidth CHRIS M1 spectral bands within the range of ~410 nm to ~1000 nm.

LCC was measured using a Minolta SPAD-161 502 chlorophyll meter (Minolta Camera Co., Osaka, Japan) at 10 locations in each subplot and averaged to one value per subplot. A linear calibration function (Equation (1)), or consensus equation \[73\], was used to convert the proprietary SPAD values to measurements of chlorophyll in \(\mu g/cm^2\). The Richardson Mountains SPAD data were analyzed against lab-based LCC measurements using destructive samples and analysis of variance (ANOVA); no significant differences were found, thereby providing confidence that SPAD measurements for all sites were well representative of LCC \[14\]. These destructive vegetation samples were also used for lab measurements of foliar and canopy carotenoid content, equivalent water thickness, and dry matter content (leaf mass) for modelling with PROSAIL.

\[
\text{Chlorophyll} \left(\mu g/cm^2\right) = \frac{99 \times \text{SPAD}}{144 \times \text{SPAD}} \quad (1)
\]

PAI, where only green, non-woody vegetation components are considered \[74,75\] was measured using downward hemispherical photography (DHP) and modelled with CAN-EYE software \[76,77\] at 21 locations in each plot. As for LCC, Richardson Mountains DHP PAI data were compared against lab measurements made by classification of photos of destructive vegetation samples laid out on a white background; very strong model fit \((r^2 = 0.86)\) provided confidence in the DHP measurements for the other sites. CAN-EYE was also used to calculate mean leaf inclination angle (ALA) for PROSAIL modelling. CCC was calculated by multiplying LCC X PAI (e.g., \[26,78–80\]. LCC, PAI and CCC values all showed a north-south gradient, where lowest overall values for each occurred at the northernmost Banks Island site \[14,65\].

As the image bounds of the CHRIS sensor are different for each view angle, in some instances, ground plots were not useable as they were outside of the field of view of the sensor. The following sample sizes were obtained for each VZA at each field site (−55°, −36°, 0°, +36°, +55°, respectively): Herschel Island \(n = 25, 33, 25, 33, 25\); Banks Island \(n = 25, 27, 31, 31, 27\); Richardson Mountains (2013) \(n = 19, 21, 25, 23, 19\); Richardson Mountains (2014) \(n = 11, 19, 20, 20, 14\).

2.3. Satellite Image Acquisition and Pre-Processing

Multi-angle CHRIS M1 images were acquired over the three study sites in the same years and within several weeks of field work. Table 1 shows the acquisition dates and times and the associated sun target-sensor geometry for each CHRIS image. As an example, Figure 2 shows the multi-angle image acquisitions for Banks Island along with the outlines of the off-nadir images.

The CHRIS data had significant issues such as missing metadata which resulted in processing errors encountered with the proprietary processing software (i.e., BEAM VISAT). VZA and VAA, where missing or deemed incorrect, were estimated using a 1st order derivation based on a tool created by White and Alonso \[81\]. The updated CHRIS metadata values are shown in Table 1 (these values have been bolded for clarity). A polar plot demonstrating the typical CHRIS acquisition angles and sun position encountered at each of the field sites is shown in Figure 1 (Banks Island shown as an example). Positive CHRIS VZAs characteristically capture forward scatter, whereas negative VZAs capture reflectance backscatter. It is possible that missing and erroneous metadata values are the result of orbit degradation of the CHRIS/PROBA satellite platform e.g., \[57\].
Table 1. CHRIS imagery timing and solar/viewing angles. Bolded values were estimated or corrected from the metadata delivered with the images. Acronyms are: view zenith angle (VZA), view azimuth angle (VAA), sun zenith angle (SZA), sun azimuth angle (SAA) and relative azimuth angle (RAA; the angular azimuth difference between the sun and sensor). It should be noted that the theoretical VZA and actual VZA (denoted with a *) differ. Acquisitions are in Coordinated Universal Time (UTC).

| Image Location                   | Acquisition Date | Time (UTC) | VZA | VZA * | VAA | SZA | SAA | RAA |
|----------------------------------|------------------|------------|-----|-------|-----|-----|-----|-----|
| Herschel Island, Yukon           | 20 July 2011     | 18:36:55   | +55°| 51.82 | 27.89| 55.00| 130.83| 102.94|
|                                 |                  | 18:37:44   | +36°| 32.22 | 35.14| 55.00| 131.05| 95.91 |
|                                 |                  | 18:38:33   | 0°   | 15.12 | 138.46| 55.00| 131.28| 7.18  |
|                                 |                  | 18:39:23   | −36°| 34.94 | 176.77| 55.00| 131.51| 45.26 |
|                                 |                  | 18:40:12   | −55°| 53.23 | 187.49| 55.00| 131.73| 55.76 |
| Banks Island, Northwest Territories | 1 August 2012  | 17:25:27   | +55°| 55.22 | 18.28| 61.00| 134.40| 116.12|
|                                 |                  | 17:26:16   | +36°| 36.69 | 12.53| 61.00| 134.61| 122.08|
|                                 |                  | 17:27:05   | 0°   | 9.00  | 300.05| 61.00| 134.83| 165.22|
|                                 |                  | 17:27:54   | −36°| 36.69 | 222.57| 61.00| 135.05| 87.52 |
|                                 |                  | 17:28:44   | −55°| 55.22 | 216.83| 61.00| 135.27| 81.56 |
| Richardson Mountains, Yukon/Northwest Territories | 2 August 2013 | 16:18:06   | +55°| 53.65 | 28.82| 67.00| 98.42 | 69.60 |
|                                 |                  | 16:19:00   | +36°| 39.67 | 48.62| 67.00| 98.42 | 49.80 |
|                                 |                  | 16:19:43   | 0°   | 21.00 | 137.52| 67.00| 98.81 | 38.71 |
|                                 |                  | 16:20:32   | −36°| 38.20 | 167.46| 67.00| 99.00 | 68.46 |
|                                 |                  | 16:21:21   | −55°| 55.29 | 180.44| 67.00| 99.19 | 81.25 |
| Richardson Mountains, Yukon/Northwest Territories | 6 July 2014   | 00:59:02   | +55°| 55.72 | 145.77| 56.81| 248.41| 102.64|
|                                 |                  | 00:59:51   | +36°| 38.06 | 140.10| 56.81| 248.62| 108.52|
|                                 |                  | 01:00:40   | 0°   | 21.48 | 42.68| 56.81| 248.83| 153.85|
|                                 |                  | 01:01:29   | −36°| 36.51 | 6.93 | 56.81| 249.03| 117.90|
|                                 |                  | 01:02:18   | −55°| 54.32 | 352.12| 56.81| 249.24| 118.64|

Figure 1. Polar plot showing the acquisition angles (i.e., forward and backscatter) for Banks Island (1 August 2012). The acquisitions represented here were typical of each site.
Figure 1. Multi-angle CHRIS Mode-1 (M1) acquisitions for Banks Island, Northwest Territories (NT) on 1 August 2012. The images are colour composites of bands 23 (red, 651 nm), 13 (green, 551 nm) and 3 (blue, 452 nm). Field plots (90 m × 90 m) are indicated with a red cross. Map created using Atlas of Canada National Scale Data 1:1,000,000 [82] and Northern Geodatabase [83].

Additional image pre-processing was conducted using the CHRIS/PROBA Toolbox software developed as an extension for the VISAT BEAM version 5.0 (Brockmann Consult, https://www.brockmann-consult.de/beam accessed on 7 April 2021). De-striping and noise reduction of raw imagery involved correction of dropouts in image rows, which is common in channel 2, and correction of vertical striping in image columns by interpolating from neighboring pixel values and spectral bands [84,85]. Cloud detection and masking was conducted using the algorithm developed by Gómez-Chova et al. [86], which provides probability and cloud abundance for each pixel. Atmospheric correction was implemented using the method described in Guanter et al. [87]. The CHRIS/PROBA atmospheric correction module (ACM) is based on the MODerate resolution TRANsmittance (MODTRAN4)
atmospheric radiative transfer code \cite{87,88} and generates a look-up table (LUT) of atmospheric parameters used in the inversion of the CHRIS top-of-atmosphere (TOA) radiance measurements for retrieval of surface reflectance values. The process requires six input parameters: VZA, sun zenith angle (SZA), relative azimuth angle (RAA), surface elevation, aerosol optical thickness (AOT) at 550 nm and columnar water vapor (CWV). The atmospheric profile is defined by a mid-latitude summer atmosphere, the aerosol type is fixed to the continental model, and the ozone concentration is fixed to 7.08 g m$^{-2}$.

Geometric correction (orthorectification) was implemented in Geomatica OrthoEngine \cite{PCI Geomatics, 2016} using Toutin’s rigorous model \cite{89–92}. Nadir imagery was the most accurately rectified with an average x, y positional error of approximately ~13.05 m. The most extreme VZAs produced the greatest degree of positional error; for example -55° had a mean positional error of ~21.62 m. Imagery for Banks Island, which is relatively flat, produced a mean positional error of 10.58 m, whereas the Richardson Mountains (2013 and 2014) images produced positional accuracies of ~22.46 m and ~21.44 m, respectively. Despite these variations, all orthorectified images had sub-pixel accuracy. Finally, spectra from the processed surface reflectance images were extracted at each field plot and compared to the ground-level spectral measurements to determine the spectral fidelity and accuracy of the spaceborne measurements using cross-correlation analysis.

2.4. Parametric and Non-Parametric Regression Modelling

Vegetation indices (VIs) were computed from the CHRIS surface reflectance data for each view angle and modelled using parametric linear regression against the field-based vegetation measurements (i.e., LCC, PAI and CCC). The VIs used were the strongest in vegetation modelling in the previous multi-angle field spectroscopy analysis \cite{14}: (1) the simple ratio \cite{93} with CHRIS M1 wavelength selection implemented using an optimization process, and (2), the revised optimized soil adjusted vegetation index (ROSAVI; \cite{94}), which is given in Equation (2) (where $R$ is the measured reflectance for wavelengths 705 nm and 750 nm). For the CHRIS red edge bands, the closest wavelengths were used (703 nm for 705 nm and 748 nm for 750 nm).

\[ \text{ROSAVI} = 1.16 \times \left( \frac{R_{750\text{nm}} - R_{705\text{nm}}}{R_{750\text{nm}} + R_{705\text{nm}} + 0.16} \right) \]  

Gaussian processes regression (GPR), a non-parametric machine learning technique was implemented using the Band Analysis Tool (GPR-BAT) algorithm \cite{32}. It was also selected because of strong performance in the field-based study \cite{14} and because of its successful implementation in leaf area index (LAI) modelling in another recent study \cite{32}. GPR assumes that an a priori Gaussian process governs the set of all possible unobserved latent functions that are consistent with the observed/known data; i.e., a Gaussian process defines the distribution over the explanatory functions and inferences about the known data are made within that Gaussian function space \cite{32,63}. Gaussian processes regression also allows for probabilistic estimates of the predicted variables \cite{32}. Here, GPR was used to identify the optimally predictive spectral bands from the CHRIS data by iteratively removing the worst performing band until only one remained.

Validation of all models was based upon 10-fold ($k = 10$) cross-validation implemented in the ARTMO (Automated Radiative Transfer Models Operator) software package \cite{95–97}. Models were assessed based on cross validated $r^2_{cv}$, RMSE$_{cv}$ and normalized RMSE (NRMSE$_{cv}$ = RMSE$_{cv}$/data range).

2.5. Physically Based Radiative Transfer Modelling and Inversion

The PROSAIL RTM \cite{61} combines the PROSPECT-5 leaf model \cite{51,98} and the SAILH canopy model \cite{99–101}. It was inverted using two methods: (1) the Levenberg-Marquardt iterative numerical optimization algorithm (LM) \cite{64} to simulate top-of-canopy (TOC) spectral reflectance and to retrieve LCC, PAI and CCC; and (2) a look-up table (LUT) approach to retrieve LCC, PAI and CCC. The selection of input parameter values and the LM and LUT inversion algorithms are described and justified in Kennedy et al. \cite{14} and
Kennedy [65]. Initial input parameter values and ranges were based on the mean and range of values over all sites of each measured field variable and, for unknown parameters, accepted values from the literature (e.g., [26–28]). The LUT inversion/regularization strategies included: (1) the use of the RMSE divergence measure (i.e., the cost function comparing measured and simulated spectra); (2) the addition of 10% Gaussian noise added to the simulated leaf-canopy spectra to account for sensor and environmental uncertainty; (3) the use of the mean of the top 10% best solutions to account for sensor uncertainty and inversion ill-posedness; and (4) the use of look-up tables with sufficient dimensionality to represent the range of measured and theoretical leaf and canopy conditions generated to match the CHRIS M1 sensor configurations and its 5 VZAs (+55°, +36°, 0°, −36° and −55°), resulting in a combined LUT with approximately $3.9 \times 10^6$ unique spectral/parameter values. The inverted PROSAIL TOC spectral reflectance and leaf and canopy parameters were directly compared to measured field variables (spectra, LCC, PAI and CCC) based on model fit ($r^2$) and predictive performance (RMSE and NRMSE).

LM was implemented through the MPFIT package for ENVI [102], whereas the LUT approach was implemented through the ARTMO software package.

3. Results

3.1. Comparison of Satellite Derived Surface Reflectance and Ground Measured Reflectance

CHRIS-derived surface reflectance matched field-based ASD reflectance closely (Figure 3) with an overall mean RMSE of 0.033 (min = 0.01, max = 0.06). The discrepancies between the two datasets are likely due to: (1) timing offsets between field work and image acquisition – maximum time difference was ~20 days (Richardson Mountains, 2014), and ~15° equivalent SZA difference (manifesting as differences to vegetation phenology and plant/soil moisture content); (2) ground-based spectral measurements taken at different VZAs than actual CHRIS VZAs (e.g., 0° vs. 9° for Banks Island, 2012; 0° vs. 21.48° for Richardson Mountains, 2013); (3) CHRIS estimated surface reflectance is affected by the MODTRAN assumptions regarding atmospheric properties/conditions; (4) CHRIS pixels are much larger than the field of view of the field measurements (i.e., issues of scaling and spectral mixing); and (5) the discrepancies in CHRIS metadata values that required their correction or estimation (look angles and atmospheric correction parameters).

Similarities between the field and CHRIS reflectance spectra included: (a) a gradual increase in visible spectrum reflectance with increasing wavelength, particularly noticeable for the high Arctic Banks Island site, where significant amounts of dry bare soil and senescent vegetation occur (i.e., red wavelengths show greater reflectance intensities than green); (b) increased chlorophyll absorption at the more southerly Herschel Island and Richardson Mountains sites, where vegetation density (i.e., percent cover) is higher; (c) backscatter reflectance (e.g., −55° and −36°) was generally greater than at nadir (0°) or in the forward scatter direction (+55° and +36°), as expected; and (d) the red edge slope was greatest at lower latitudes where vegetation density (biomass) was greater. Like the field-based spectra of the previous study, the CHRIS reflectance spectra did not have a typical green vegetation NIR reflectance plateau; this has been found in other studies and may be attributed to the low vegetation height/structure and low green leaf area of these Arctic sites (e.g., [37–40,47]).
Figure 3. Comparisons of CHRIS Mode-1 (M1) surface reflectance measurements with ground-based multi-angle ASD (Analytical Spectral Devices) spectra acquired at each of the field sites. Herschel Island field spectra were only acquired at nadir.
Differences between the field and CHRIS reflectance spectra included the following: (a) greater reflectance values were generally observed in the NIR wavelengths for the CHRIS spectra than for the field spectra for Richardson Mountains 2013, the opposite occurred in 2014; (b) field spectra showed deeper absorption in the red wavelengths (centered around ~680 nm) and greater peak green reflectance (centered around ~550 nm) for all sites except Banks Island, whereas these features were not as pronounced or were absent from the CHRIS spectra at all field sites; (c) reflectance in the visible wavelengths (~400 nm to ~700 nm) was almost always greater for the field spectra than for the CHRIS spectra.

3.2. Parametric Regression: SR and ROSAVI

Over all variables, sites and VZAs, SR ($r^2_{cv} = 0.60$; NRMSE$_{cv} = 0.14$) outperformed ROSAVI ($r^2_{cv} = 0.51$; NRMSE$_{cv} = 0.15$). It produced higher mean $r^2_{cv}$ values over all sites and VZAs for LCC ($r^2_{cv} = 0.61$, NRMSE$_{cv} = 0.13$) than for PAI ($r^2_{cv} = 0.60$, NRMSE$_{cv} = 0.14$) or CCC ($r^2_{cv} = 0.58$, NRMSE$_{cv} = 0.14$). These results differed from the field-scale analysis [14] where variable rankings were best for PAI, CCC and then LCC. Similar variable ranking occurred for ROSAVI and for the field-scale analysis, where PAI produced the best results followed by LCC and CCC. Possible reasons for ROSAVI-based PAI results include the detectability of plant greenness/structure as compared to pigment quantities occurring at finer spatial scales; however, the better results obtained for SR-based LCC may be due to the range of values included into the modelling approach. Better results for PAI (or LAI) and CCC as compared to LCC have been reported in other studies (e.g., [26,27]).

The VZA that produced the best mean model fit for all three vegetation variables was $0^\circ$ for SR ($r^2_{cv} = 0.64$, NRMSE$_{cv} = 0.13$) and $+55^\circ$ for ROSAVI ($r^2_{cv} = 0.54$, NRMSE$_{cv} = 0.14$); however, the results for the other VZAs were generally within the same range (e.g., $r^2_{cv} = \pm 0.03$, NRMSE$_{cv} = \pm 0.01$) except for the $+36^\circ$ VZA which produced the lowest overall scores for both SR and ROSAVI ($r^2_{cv} = 0.50$, NRMSE$_{cv} = 0.14$ and $r^2_{cv} = 0.44$, NRMSE$_{cv} = 0.15$, respectively). Similar angular performance was observed in the field-scale analysis (i.e., $0^\circ$ and $+55^\circ$ VZA producing the best overall results). Table 2 shows the results for SR and ROSAVI at each view angle for the global analysis.

When VI results were stratified by geographic location (local models), similar results were found as for the previous field-scale analysis [14]; i.e., Banks Island data produced the best overall predictions for all vegetation variables (e.g., LCC, $r^2_{cv} = 0.74$, NRMSE$_{cv} = 0.13$ for the $+55^\circ$ VZA; PAI, $r^2_{cv} = 0.87$, NRMSE$_{cv} = 0.10$ and CCC, $r^2_{cv} = 0.85$, NRMSE$_{cv} = 0.10$ for the $0^\circ$ VZA) due to its larger variation in LCC and PAI as vegetation density varied significantly between ground plots. Figure 4 shows the best local SR models for each vegetation variable and VZA for Banks Island; they are only slightly poorer than those of the previous field study. The local model predictions for Banks Island (Figure 4) demonstrated the following: (1) LCC models showed over predictions in the lower value ranges (below ~15 $\mu g/cm^2$) with moderate prediction variability (error) observed across the full measurement range; and (2) both PAI and CCC predictions showed variably proportional biases (positive in the lower end and negative in the upper end) with relatively low prediction variability (error) across all measurement ranges. It is likely that the proportion of photosynthetic to non-photosynthetic materials is driving the CCC modelling outcomes in Banks Island, since LCC is only likely to be detected when green vegetation cover (i.e., PAI) reaches a threshold of detectability within a pixel.
Table 2. Global simple ratio (SR) and revised optimized soil adjusted vegetation index (ROSAVI) model performance for combined field sites. Root mean squared error (RMSE) units: leaf chlorophyll content, LCC (µg/cm²); plant area index, PAI (m²/m²); canopy chlorophyll content, CCC (g/m²). Normalized root mean squared error (NRMSE) = RMSE/range of measured data. The best results are bolded and highlighted.

| Variable | Angle | $r^2_{cv}$ | RMSE$_{cv}$ | NRMSE$_{cv}$ | $r^2_{cv}$ | RMSE$_{cv}$ | NRMSE$_{cv}$ |
|----------|-------|------------|-------------|--------------|------------|-------------|--------------|
| LCC      | All   | 0.61       | 5.39        | 0.13         | 0.50       | 6.12        | 0.14         |
| PAI      | All   | 0.60       | 0.25        | 0.14         | 0.54       | 0.27        | 0.15         |
| CCC      | All   | 0.58       | 0.11        | 0.14         | 0.50       | 0.12        | 0.15         |
| LCC      | +55°  | 0.64       | 5.67        | 0.13         | 0.56       | 6.28        | 0.14         |
| LCC      | +36°  | 0.43       | 5.58        | 0.13         | 0.36       | 5.89        | 0.13         |
| LCC      | 0°    | 0.68       | 4.90        | 0.11         | 0.53       | 5.90        | 0.13         |
| LCC      | −36°  | 0.68       | 4.95        | 0.15         | 0.55       | 5.89        | 0.13         |
| LCC      | −55°  | 0.61       | 5.87        | 0.13         | 0.50       | 6.66        | 0.15         |
| PAI      | +55°  | 0.61       | 0.25        | 0.14         | 0.58       | 0.26        | 0.14         |
| PAI      | +36°  | 0.54       | 0.26        | 0.14         | 0.49       | 0.28        | 0.15         |
| PAI      | 0°    | 0.63       | 0.24        | 0.13         | 0.56       | 0.26        | 0.14         |
| PAI      | −36°  | 0.58       | 0.25        | 0.14         | 0.52       | 0.27        | 0.15         |
| PAI      | −55°  | 0.63       | 0.25        | 0.14         | 0.54       | 0.28        | 0.15         |
| CCC      | +55°  | 0.58       | 0.11        | 0.14         | 0.55       | 0.12        | 0.15         |
| CCC      | +36°  | 0.54       | 0.11        | 0.15         | 0.46       | 0.12        | 0.16         |
| CCC      | 0°    | 0.61       | 0.10        | 0.14         | 0.52       | 0.12        | 0.15         |
| CCC      | −36°  | 0.57       | 0.11        | 0.14         | 0.49       | 0.12        | 0.15         |
| CCC      | −55°  | 0.59       | 0.11        | 0.14         | 0.48       | 0.13        | 0.16         |
| All      | +55°  | 0.61       | -           | 0.14         | 0.56       | -           | 0.14         |
| All      | +36°  | 0.50       | -           | 0.14         | 0.44       | -           | 0.15         |
| All      | 0°    | 0.64       | -           | 0.13         | 0.54       | -           | 0.14         |
| All      | −36°  | 0.61       | -           | 0.14         | 0.52       | -           | 0.14         |
| All      | −55°  | 0.61       | -           | 0.14         | 0.51       | -           | 0.15         |

The simple ratio wavelength optimization procedure produced a wide range of optimal spectral bands for the various datasets. For LCC over all VZAs and field sites (i.e., global models), optimal bands were clustered around the visible pigment absorption regions for chlorophyll (B1 = ~539 to 681 nm) and within the red edge and NIR regions typical of healthy vegetation (B2 = ~717 to 847 nm). Optimal weighted bands for PAI were more overlapping with B2 centered more in the visible than the red edge (B1 = 627–711 nm; B2 = 626–678 nm). Optimal bands for CCC were broader than for both LCC and PAI, with B1 extending from the blue-green to near the top of the red edge and B2 from the green to the red edge (B1 = 519–756 nm; B2 = 576–738 nm). Band selections for LCC, PAI and CCC match very closely those bands reported at the field scale, and other studies that have measured similar green based vegetation variables, such as percent green vegetation cover (~680–872 nm) [37] and biomass (~680–872 nm) [39]. Band selections for local models were similar to the global models. However, there were instances in both local and global modelling where band selections were either wide-ranging/separated (e.g., occurring in the ~480 nm and ~900 nm ranges) or selected within close proximity to each other (e.g., occurring within ~20 nm). Figure 5 shows the best local and global models for CCC applied to the Banks Island imagery (both models are for the 0° VZA and show the optimal band selections). Global models provided greater upper range predictions while suppressing low-level spatial variability (Figure 5).
Figure 4. Scatterplots showing measured vs. predicted relationships for the best locally developed SR modelling outputs for Banks Island, 2012. Root mean squared error (RMSE) units: (a) leaf chlorophyll content, LCC (µg/cm²); (b) plant area index, PAI (m²/m²); (c) canopy chlorophyll content, CCC (g/m²). Normalized root mean squared error (NRMSE) = RMSE/range of measured data. The 1:1 identify line (red) is shown in each scatterplot to provide a comparative reference.

With respect to view angle, a very weak trend of decreasing wavelength of the optimal bands was evident as VZA progresses from backscatter to forward scatter. For example, the optimal B2 wavelength varied between −55° and +55° as follows: LCC: ~850 nm to 720 nm; PAI: ~690 nm to 620 nm; CCC: 710 nm to 580 nm. Given the variability observed in the band selections, it was uncertain if this may simply be the result of natural variability or noise in the measured reflectance signals; more controlled testing would be needed to confirm this trend.
Figure 4. Scatterplots showing measured vs. predicted relationships for the best locally developed SR modelling outputs for Banks Island, 2012. Root mean squared error (RMSE) units: (a) leaf chlorophyll content, LCC (μg/cm²); (b) plant area index, PAI (m²/m²); (c) canopy chlorophyll content, CCC (g/m²). Normalized root mean squared error (NRMSE) = RMSE/range of measured data. The 1:1 identify line (red) is shown in each scatterplot to provide a comparative reference.

Figure 5. Simple ratio (SR) -based global model (a) and local model (b) predictions for canopy chlorophyll content (CCC) for Banks Island, 2012. Both CCC models were developed and applied to the 0° view zenith angle (VZA). Wavelengths (optimal bands) used for each retrieval are shown above the maps. Predictive maps are shown raster up for visualization purposes (see Figure 2 for geographic orientation of the image).

3.3. Non-Parametric Regression: Gaussian Processes Regression (GPR)

GPR produced the best model fit for all vegetation variables regardless of VZA and geographic location (mean $r^2_{cv} = 0.61$). Predictive performance was marginally better than for SR (NRMSE$_{cv}$ = 0.13 vs. NRMSE$_{cv}$ = 0.14, respectively). The best single predictor (i.e., the last remaining spectral band) produced a mean $r^2_{cv} = -0.30$. As variables were removed, $r^2_{cv}$ values generally increased and RMSE$_{cv}$ decreased until about two to four variables remained (averaged across variables and VZAs), when models became less stable and predictive performance dropped suddenly (as shown in Figure 6). Over all VZAs and sites, models were generally strongest for PAI (mean $r^2_{cv} = 0.62$, NRMSE$_{cv}$ = 0.14) and LCC (mean $r^2_{cv} = 0.62$, NRMSE$_{cv}$ = 0.12), followed by CCC (mean $r^2_{cv} = 0.59$, NRMSE$_{cv}$ = 0.14). However, the band removal (wavelength optimization) results demonstrated the following: (1) PAI showed more consistent and better results (based on model fits) than LCC and CCC when comparing the equivalent number of predictor variables through each of the successive model iterations (Figure 6a); (2) LCC showed the largest range of results between VZAs, with the +36° VZA producing the lowest overall scores and the −36° VZA producing the highest; (3) PAI and CCC models were more consistent than LCC between VZAs as bands were removed from the models; and (4) the −55° VZA generally produced the strongest model gains in $r^2_{cv}$ as variables were removed, whereas the nadir VZA (0°) produced amongst the lowest NRMSE$_{cv}$ values and most consistent (highest/best) results for each of the variables across each stage of the band optimization routines. This last point contrasts the field scale analyses, where −55° produced the best GPR models [14]. Table 3 shows the global GPR results at each view angle and for each of the predicted variables. Figure 7 shows the overall best global GPR model fits for LCC (−36° VZA), PAI (+55° VZA) and CCC (0° VZA).
Similar to SR and ROSAVI, and to the previous field-based study, Banks Island produced the best local GPR models over all vegetation variables and VZAs (mean $r^2_{cv} = 0.68$; mean NRMSE$_{cv} = 0.17$). The Richardson Mountains (2014) site was very comparable (mean $r^2_{cv} = 0.68$; mean NRMSE$_{cv} = 0.24$). However, models may have been over-fit at this location as the number of predictor variables was high with respect to the number of independent observations on the ground, especially for $+55^\circ$ where only 14 plots were within the CHRIS imagery. The global GPR model predictions shown in Figure 7 demonstrated the following: (1) a higher degree of prediction variability (error) was observed in LCC values below $26\mu g/cm^2$, however low measurements (e.g., $0\mu g/cm^2$) helped boost model statistics; (2) like the local SR models (Figure 4), PAI predictions showed a variably proportional bias (positive bias in the lower end and negative bias in the upper end), however prediction error was overall greater than local PAI models for Banks Island; and (3) a greater degree of error was observed for the upper ranges of PAI and CCC (in values $\geq 1.2 m^2/m^2$ and $\geq 0.5 g/m^2$, respectively) than for the lower ranges.

Figure 6. Gaussian processes regression (GPR) mean $r^2_{cv}$ values for: (a) all sites and VZAs combined; (b) leaf chlorophyll content (LCC) over all sites and for each view angle; (c) plant area index (PAI) over all sites and for each view angle; and (d) canopy chlorophyll content (CCC) over all sites and for each view angle. Graphs demonstrate modelling results as spectral bands are removed from the optimized GPR model. Results improved as bands were removed; optimal models typically contained 3 to 4 spectral features (i.e., predictor variables). The $0^\circ$ view zenith angle (VZA) produced the best overall modelling results when variables were averaged. PAI model fits were typically highest and most consistent as predictor variables were iteratively removed from the GPR models.
26 µg/cm², however low measurements (e.g., 0 µg/cm²) helped boost model statistics; (2) like the local SR models (Figure 4), PAI predictions showed a variably proportional bias (positive bias in the lower end and negative bias in the upper end), however prediction error was overall greater than local PAI models for Banks Island; and (3) a greater degree of error was observed for the upper ranges of PAI and CCC (in values ≥1.2 m²/m² and ≥0.5 g/m², respectively) than for the lower ranges.

Table 3. Global Gaussian processes regression (GPR) model performance for combined field sites. Root mean squared error (RMSE) units: leaf chlorophyll content, LCC (µg/cm²); plant area index, PAI (m²/m²); canopy chlorophyll content, CCC (g/m²). Normalized root mean squared error (NRMSE) = RMSE/range of measured data. The best results are shown highlighted and bolded.

| Variable | Angle | $r^2_{cv}$ | RMSE$_{cv}$ | NRMSE$_{cv}$ |
|----------|-------|------------|-------------|--------------|
| LCC      | All   | 0.62       | 5.49        | 0.12         |
| PAI      | All   | 0.62       | 0.25        | 0.14         |
| CCC      | All   | 0.59       | 0.11        | 0.14         |
| LCC      | +55°  | 0.58       | 6.19        | 0.14         |
| LCC      | +36°  | 0.51       | 6.03        | 0.14         |
| LCC      | 0°    | 0.66       | 5.04        | 0.11         |
| LCC      | −36°  | 0.70       | 4.76        | 0.11         |
| LCC      | −55°  | 0.67       | 5.42        | 0.12         |
| PAI      | +55°  | 0.66       | 0.24        | 0.13         |
| PAI      | +36°  | 0.58       | 0.25        | 0.14         |
| PAI      | 0°    | 0.64       | 0.23        | 0.13         |
| PAI      | −36°  | 0.60       | 0.25        | 0.14         |
| PAI      | −55°  | 0.61       | 0.26        | 0.14         |
| CCC      | +55°  | 0.61       | 0.11        | 0.14         |
| CCC      | +36°  | 0.56       | 0.11        | 0.14         |
| CCC      | 0°    | 0.64       | 0.10        | 0.13         |
| CCC      | −36°  | 0.56       | 0.11        | 0.14         |
| CCC      | −55°  | 0.60       | 0.11        | 0.14         |
| All      | +55°  | 0.62       | -           | 0.14         |
| All      | +36°  | 0.50       | -           | 0.14         |
| All      | 0°    | 0.65       | -           | 0.12         |
| All      | −36°  | 0.62       | -           | 0.13         |
| All      | −55°  | 0.63       | -           | 0.13         |

Optimal bands selected by GPR were generally in the pigment absorption and red edge spectral regions with LCC optimal bands extending into the NIR (LCC: 553–818 nm; PAI: 514–772 nm; CCC: 515–772 nm). Similar to the SR band selections, overall GPR band selections showed a very weak downward trend in wavelength from backscatter through nadir to forward scatter angles.

3.4. Physical Modelling: PROSAIL Inversion Using LM and LUT
3.4.1. Spectral Curve Fitting Validation Using LM

LM inversion of the PROSAIL model reproduced TOC directional-reflectance very closely for all value ranges of LCC, PAI and CCC, (i.e., low to high values) and at all VZAs. (mean $r^2 = 0.99$, mean RMSE = 0.005 ± 0.005). These results were only slightly poorer than the field-based study (mean RMSE = 0.004 ± 0.002). Figure 8 shows example spectra from Richardson Mountains (2013) and Banks Island, which represent more dense and sparse vegetation cover, respectively. Inverted modelled parameters (LCC, PAI and CCC) are shown for each curve.
PAI: 514–772 nm; CCC: 515–772 nm). Similar to the SR band selections, overall GPR band selections showed a very weak downward trend in wavelength from backscatter through nadir to forward scatter angles.

Figure 7. Comparison of Gaussian processes regression (GPR) modelling retrievals for (a) leaf chlorophyll content (LCC, μg/cm²), (b) plant area index (PAI, m²/m²), and (c) canopy chlorophyll content (CCC, g/m²) for combined field sites (global model) using the best fitting models. GPR retrievals for LCC were best for the −36° view zenith angle (VZA), +55° VZA for PAI, and 0° VZA for CCC. The 1:1 identify line (red) is shown in each scatterplot to provide a comparative reference.

Observation spectral matching errors may be attributable to the following: (1) the CHRIS/PROBA ACM (atmospheric correction) may have produced erroneous surface reflectance values due to metadata parameters (e.g., view angle data) that were missing and had to be estimated for some acquisitions; (2) spatial averaging of subplot leaf and canopy field measurements to the plot scale (i.e., scaling of field data to CHRIS pixel sizes); and (3) value ranges for variables (e.g., carotenoids) measured in the lab only for the Richardson Mountains (2014) site may not have been correct for Herschel Island or Banks Island.

Figure 9 shows the differences between the modelled and measured reflectance values for two example sites and VZAs. Error patterns appeared to be systematic (i.e., non-random) and observable for all sites and VZAs. The reflectance values in these spectral regions are either not well measured, with respect to inherent noise associated with spectroradiometers [32], related to the atmospheric correction code/values used to correct the
imagery, or, more likely, not well modelled by the PROSAIL RTM (e.g., the parameter values may be off/incorrect, the PROSAIL model requires recalibration to match Arctic conditions, or the model is not able to account for spectral variability encountered at each of the field sites).

**Figure 8.** Examples representing the range of measured and modelled top-of-canopy (TOC) spectral-directional reflectance curves for: (a) Richardson Mountains (2013) with more dense vegetation cover (view zenith angle, VZA = −55°); and (b) Banks Island with sparse vegetation cover (VZA = +36°).

**Figure 9.** Difference ($\Delta$) between the PROSAIL modelled and measured spectral-directional reflectance values for full spectrum CHRIS Mode-1 (M1) datasets for two example field sites and view zenith angles (VZAs): (a) Richardson Mountains (2013), −55° VZA; and (b) Banks Island, +36° VZA. Each plot shows all modelled spectra for that location and VZA. Similar error patterns between sites and VZAs were evident. The range of differences between the modelled and measured spectra were generally about ±0.03, with slightly larger error evident in the blue wavelengths and lower in the near-infrared (NIR). Error patterns may be related to noise associated with the CHRIS sensor (e.g., <~500 nm).
3.4.2. Vegetation Modelling: PROSAIL Inversion Using LM and LUT

Similar to the field-based study, the performance of PROSAIL LM inversions were not as good as empirical modelling of vegetation parameters. However, the average model fit over all vegetation variables, sites and VZAs was the same as for the field data using simulated CHRIS bands ($r^2 = 0.28$; NRMSE = 0.27). On the other hand, the LUT-based inversion performed much better than the LM inversion ($r^2 = 0.50$, NRMSE = 0.18), were within range of ROSAVI results ($r^2_{cv} = 0.51$, NRMSE$_{cv} = 0.15$) of this study, and performed slightly better than the field-scale LUT inversion results ($r^2 = 0.48$; NRMSE = 0.22). Table 4 shows the LM and LUT PROSAIL inversion results at each view angle and across the three field sites (i.e., global results).

Table 4. Global PROSAIL model performance for Levenberg–Marquardt (LM) and look-up table (LUT) inversions for combined field sites. Root mean square error (RMSE) units: leaf chlorophyll content, LCC ($\mu g/cm^2$); plant area index, PAI ($m^2/m^2$); canopy chlorophyll content, CCC (g/m$^2$). Normalized root mean squared error (NRMSE) = RMSE/range of measured data. The best results are shown highlighted and bolded.

| Variable | Angle | $r^2_{cv}$ | RMSE$_{cv}$ | NRMSE$_{cv}$ | $r^2_{cv}$ | RMSE$_{cv}$ | NRMSE$_{cv}$ |
|----------|-------|------------|-------------|--------------|------------|-------------|--------------|
| LCC      | All   | 0.30       | 7.61        | 0.25         | 0.55       | 7.60        | 0.17         |
| PAI      | All   | 0.21       | 0.39        | 0.38         | 0.46       | 0.34        | 0.19         |
| CCC      | All   | 0.32       | 0.13        | 0.18         | 0.48       | 0.14        | 0.18         |
| LCC      | $+55^\circ$ | 0.30    | 7.13        | 0.24         | 0.35       | 7.43        | 0.17         |
| LCC      | $+36^\circ$ | 0.32    | 7.18        | 0.23         | 0.41       | 7.70        | 0.17         |
| LCC      | $0^\circ$  | 0.34      | 7.64        | 0.24         | 0.62       | 7.46        | 0.17         |
| LCC      | $-36^\circ$ | 0.25    | 8.70        | 0.27         | 0.63       | 8.13        | 0.18         |
| LCC      | $-55^\circ$ | 0.31    | 7.40        | 0.24         | 0.35       | 7.26        | 0.16         |
| PAI      | $+55^\circ$ | 0.21    | 0.40        | 0.39         | 0.40       | 0.32        | 0.18         |
| PAI      | $+36^\circ$ | 0.21    | 0.40        | 0.39         | 0.41       | 0.30        | 0.17         |
| PAI      | $0^\circ$   | 0.18      | 0.40        | 0.39         | 0.50       | 0.29        | 0.16         |
| PAI      | $-36^\circ$ | 0.26    | 0.40        | 0.39         | 0.51       | 0.29        | 0.16         |
| PAI      | $-55^\circ$ | 0.21    | 0.36        | 0.35         | 0.47       | 0.50        | 0.28         |
| CCC      | $+55^\circ$ | 0.35    | 0.13        | 0.18         | 0.46       | 0.13        | 0.17         |
| CCC      | $+36^\circ$ | 0.30    | 0.15        | 0.22         | 0.48       | 0.14        | 0.18         |
| CCC      | $0^\circ$   | 0.42      | 0.11        | 0.16         | 0.55       | 0.15        | 0.19         |
| CCC      | $-36^\circ$ | 0.32    | 0.12        | 0.17         | 0.40       | 0.15        | 0.19         |
| CCC      | $-55^\circ$ | 0.22    | 0.12        | 0.16         | 0.51       | 0.13        | 0.17         |
| All      | $+55^\circ$ | 0.29    | -           | 0.27         | 0.47       | -           | 0.17         |
| All      | $+36^\circ$ | 0.28    | -           | 0.28         | 0.43       | -           | 0.17         |
| All      | $0^\circ$   | 0.31      | -           | 0.26         | 0.56       | -           | 0.17         |
| All      | $-36^\circ$ | 0.28    | -           | 0.28         | 0.51       | -           | 0.18         |
| All      | $-55^\circ$ | 0.25    | -           | 0.25         | 0.51       | -           | 0.20         |

LM models for CCC ($r^2 = 0.32$, NRMSE = 0.18) were stronger than for LCC ($r^2 = 0.30$, NRMSE = 0.25) and PAI ($r^2 = 0.21$, NRMSE = 0.38), whereas LUT models were strongest for LCC ($r^2 = 0.55$, NRMSE = 0.17) than for PAI ($r^2 = 0.46$, NRMSE = 0.19) or CCC ($r^2 = 0.48$, NRMSE = 0.18). With respect to VZA, nadir ($0^\circ$) produced the best models for both LM and LUT inversions ($r^2 = 0.31$, NRMSE = 0.26 and $r^2 = 0.56$, NRMSE = 0.17, respectively). This is the same as the empirical results achieved in this study (e.g., for SR, ROSAVI and GPR), but differs slightly to the field study [14] where LUT-based inversion results were best for the $+55^\circ$ VZA. Finally, as for the empirical analysis and the previous field study, Banks Island produced the best models for all vegetation variables, for example: (1) for the LM parameter set at $0^\circ$ VZA: LCC, $r^2 = 0.61$, NRMSE = 0.25; PAI, $r^2 = 0.65$, NRMSE = 0.05; CCC, $r^2 = 0.63$, NRMSE = 0.29; and (2) for the LUT parameter set at $0^\circ$ VZA: LCC, $r^2 = 0.73$, NRMSE = 0.19; PAI, $r^2 = 0.79$, NRMSE = 0.28; CCC, $r^2 = 0.80$, NRMSE = 0.35. Figure 10 shows the LUT-based modelling results of CCC for both global and local
models (Banks Island shown as an example). The model prediction scatterplots shown in Figure 10 demonstrated the following: (1) an overall positive bias was observed for both local and global CCC predictions, where values from 0.0 to 0.4 g.m² were generally over predicted by ~0.1 g/m²; (2) a higher degree of prediction variability and underprediction (error) was observed in the global model upper end (CCC values above ~0.4 g/m²). Over prediction in both global and local models was also observed in the field study [14] and may attributable to parametrization ranges of the LUT (i.e., the step sequence of parameterized values). Greater variability and underpredictions observed above ~0.4 g/m² may be due to reflectance saturation effects (e.g., spectral similarity between plots due to lack of canopy gap fractions and similar biomass quantities) and the resulting model ill-posedness that would be encountered as vegetation spectra become more alike when progressing southward.

![Figure 10. The best look-up table (LUT) inversion results for canopy chlorophyll content (CCC) for all sites (a) and Banks Island (b) using only CHRIS Mode-1 (M1) nadir (view zenith angle, VZA = 0°) acquisitions. Root mean square error (RMSE) units: CCC (g/m²). The 1:1 identify line (red) is shown in each scatterplot to provide a comparative reference.](image)

The vegetation variable that was best predicted with LM across all sites was CCC (mean value across all VZAs, datasets and models; NRMSE = 0.14), followed by LCC (NRMSE = 0.17) and PAI (NRMSE = 0.24). The superior prediction of CCC follows that of Darvishzadeh et al. [27]; however, the relatively poor prediction of PAI in comparison to both CCC and LCC contrasts with other studies (e.g., [26,103,104]). Reasons for this are discussed in Kennedy et al. [14] and point mainly towards poor signal propagation from leaf to canopy scales (e.g., [49,105]); theoretically, PAI should produce better models than LCC. In comparison, the global LUT inversion approach best predicted LCC (NRMSE = 0.17) followed by CCC (NRMSE = 0.18) and then PAI (NRMSE = 0.19), which follows the result reported for the field scale LUT-based inversions (e.g., LCC, NRMSE = 0.14; CCC, NRMSE = 0.23; PAI, NRMSE = 0.26).

3.5. VZA Synthesis

The combined mean prediction results (from both empirical and physical models, excluding LM inversion) showed that all VZAs produced similar NRMSEcv and NRMSE values (e.g., mean values for all combined datasets NRMSE = 0.15 ± 0.05), with 0° producing the best overall mean predictions across all modelling types (mean NRMSE = 0.14 ± 0.01). In contrast, results from the field-based study showed comparable empirical results for most VZAs (i.e., 0°, ±36°, ±55°) (mean NRMSEcv = 0.15) with 0° and −55° having the
best overall performance, but $0^\circ$, $\pm 36^\circ$, $+55^\circ$ were better than $-55^\circ$ in physical modelling (mean NRMSE = 0.20 + 0.03) [14]. Finally, the magnitude of NRMSE$_{cv}$ and NRMSE values (when stratified according to modelling method) are similar to the field-based study [14] and suggest negligible loss of model quality when the field-based model approaches are translated to satellite-based modelling. It also suggests that the field-sampling or scaling approach (i.e., spatial averaging across the subplots) was appropriate for the spatial resolution of the CHRIS sensor.

4. Discussion

In this study, a satellite-based, multi-angle, multi-method spectroscopic analysis of Arctic vegetation was conducted across a bioclimatic gradient in the Western Canadian Arctic between the years 2011 to 2014. To date there have been no studies using satellite-based imaging spectroscopy to retrieve LCC, PAI and CCC from Arctic tundra vegetation, nor that compare empirical and physically-based modelling approaches to retrieve these variables. Overall, the results demonstrate the importance of methodological testing across multiple sites within large ecosystems.

4.1. CHRIS M1 TOC Reflectance Spectra

Analysis of the multi-angle CHRIS M1 TOC reflectance spectra revealed sun-sensor angular dependencies across all field sites. Backscatter VZAs ($-55^\circ$ and $-36^\circ$) typically produced the greatest overall reflectance values due to a lack of shadows when the angular difference between the sensor and the sun was relatively small, as forward scattering observations contain more signal from shaded leaves and branches [14,29]. Nadir spectra typically produced the lowest reflectance across all sites and wavelengths due to greater visible canopy gap fractions [29,106,107]. However, Herschel Island VZAs of $+36^\circ$ and $+55^\circ$ produced reflectance values that were slightly lower than the nadir spectra. At this study site, hummocks were prevalent and possibly contributed shadows at nadir and in the forward scatter VZAs (i.e., shadowing caused by microtopographic variation, tussocks and shrubs). Some plots were also on gradual slopes facing away from the sun. The overall low sun angles, in combination with the greater degree of visible micro-shadows caused by hummocks, may have contributed to larger visible shadowing in the positive VZAs. For the Richardson Mountains (2014), $+55^\circ$ and $-55^\circ$ VZAs produced very similar surface reflectance values. This may be because this site was flatter (in a valley) and imaged with relatively high sun angles ($\sim 57^\circ$), which can produce more volume scattering and a reflectance response that is parabolic as reflectance changes with VZA from backscatter to forward scatter (e.g., [108]).

The differences between the Richardson Mountains 2013 and 2014 CHRIS spectral data, especially in the $-36^\circ$ and $-55^\circ$ (backscatter) VZAs, are likely the result of sun position and sun-sensor relative azimuth angles, as imaging times were quite different. The Richardson Mountains 2013 backscatter acquisitions had a very low sun angle of $-67^\circ$ and relative azimuth angles of $-68$ to $-81^\circ$ (for the $-36^\circ$ and $-55^\circ$ VZAs, respectively), whereas the 2014 backscatter acquisitions had a higher sun angle of $-57^\circ$ and relative azimuth angles between $-118^\circ$ and $-119^\circ$ (for the $-36^\circ$ and $-55^\circ$ VZAs, respectively). The observed spectral differences demonstrate the influence that sun zenith and azimuth angles have on vegetation reflectance (e.g., the greater degree of backscatter reflectance in the 2013 imagery due to smaller relative azimuth angles) but they also highlight the potential confounding influence of additional factors such as vegetation phenology and surface/soil moisture on the reflectance signal e.g., [18]. For example, the 2014 imagery acquired earlier in the season (6 July) may have captured the ecosystem in a state with a greater quantity of residual surface and soil moisture from snowmelt, which may have resulted in the observed attenuated NIR reflectance values. Conversely, the 2013 imagery acquired later in the season (2 August) captured the canopy in a more advanced phenological state, whereby later season vegetation pigment and structure development (i.e., leaf-canopy growth) would be expressed, respectively, as deeper absorption in the red wavelengths.
(≈680 nm) and greater reflectance in the NIR wavelengths (≥750 nm)—both of which are observable across all VZAs in the 2013 CHRIS spectra. These observations demonstrate the need for: (a) repeat spectroscopic measurements in one study site across the growing season so that environmental and phenological variables (e.g., sun geometry variability, surface moisture, senescent vegetation) are better characterized; (b) a temporal match between field and satellite data acquisitions (when possible); and (c) in situ optical sampling to evaluate/validate satellite imaging spectroscopy data.

CHRIS spectra were similar to the previous field-based spectral analysis [14] and other field-based studies (e.g., [37–39]), where non-vegetation components contributed significantly to reflectance (when spectra were stratified and compared by VZA). At all sites, visible substrates (i.e., soils, sand, rocks), variability of photosynthetic vegetation components (i.e., pigment pools and species assemblages), non-photosynthetic vegetation components (i.e., senescent vegetation, woody materials and lichens) and the presence of moisture influenced reflectance and absorption across visible and NIR wavelength ranges. This was most evident at the northern Banks Island site where vegetation was sparser, and spectra resembled a combination of vegetation and soil spectra with low visible reflectance increasing gradually through the NIR. For the other two sites, visible reflectance had a distinct green reflectance peak but often reflectance remained flat into the red spectrum. The NIR plateau was also generally sloped positively as wavelength increased.

Discrepancies between the field and CHRIS spectra were particularly evident across the NIR wavelengths for the Richardson Mountains 2013 and 2014 sites. Field sampling locations between years were approximately the same; however, timing differences between field spectral measurements and CHRIS acquisitions existed for both years. In 2013, field spectra were collected approximately 11 days before the CHRIS acquisitions and 2014 field spectra were collected approximately 20 days after the CHRIS acquisitions. Gamon et al. [18] showed that peak NDVI timing (i.e., reflectance) can vary from year to year (e.g., day of year 216 vs. 222) in coastal Alaskan tundra ecosystems and that vegetation productivity is largely a function of microtopographic effects, precipitation, soil moisture and growing degree days. It is likely that differences in field sampling locations coupled with timing differences between the field and CHRIS acquisitions (as they relate to spatial heterogeneity, vegetation species compositions, vegetation growth based on microtopographic effects, shadowing, phenology, temperature and precipitation/moisture, etc.) contributed to the observed discrepancies between the field and CHRIS reflectance values. Finally, it should be noted that field spectra were measured (integrated) across multiple hours/days as sites were sampled throughout each field season, compared to the instantaneous satellite acquisitions which captured (almost) identical sun positions at each of the respective field sites (Table 1). The integration of spectra over the solar day and across multiple sampling days (as it relates to differing illumination conditions) may have acted to reduce the apparent influence that SZA and SAA have on surface/vegetation reflectance, as can be observed in the comparison between the Richardson Mountains 2013 and 2014 CHRIS acquisitions.

4.2. Empirical Modelling of LCC, PAI and CCC

Models of LCC, CCC and PAI were stronger using empirical techniques than physical-based modelling with PROSAIL and all models were only slightly poorer than in the previous field-based study [14]. GPR produced the best model fits based on cross-validated coefficient of determination ($r^2_{cv}$) and the best predictive capability based on the cross-validated normalized RMSE (NRMSE$_{cv}$), with the simple ratio (SR) producing the 2nd best overall results. Both GPR and SR incorporated a spectral band optimization routine and the selected optimal wavelength regions were generally in well-known pigment absorption regions and along the red edge into the NIR as found in the previous field-scale study and in other studies examining green vegetation components (e.g., [37–39]). However, in some cases optimal bands were outside these regions. For example, for SR, some LCC data sets resulted in optimal band selections as low as 481 nm and optimal band 2 selections as
high as 915 nm. For GPR, which started the optimization routine with all 62 bands in a model, optimal results were generally achieved with 3–4 variables; i.e., removal of the least contributing variable at each iteration increased model fit and prediction capability but only up to a point where models started to deteriorate. Darvishzadeh [26] also showed that vegetation variable retrievals using partial least squares regression can benefit from cautiously selected spectral subsets. The selection of bands outside of expected ranges can occur by chance or through indirect reflectance responses to biochemical constituents, canopy structural and non-vegetated background components [29,109,110]. In this research, it is possible that canopy gaps (i.e., exposed non-green vegetation and non-vegetation materials) were responsible for the cases where optimal wavelengths were selected outside of expected regions of vegetation reflectance. It is also possible that errors are related to incorrect atmospheric correction parameters, to geolocation errors, and to spectral noise (systematic and/or random).

ROSAVI (3rd in overall performance) was utilized in parametric regression with the CHRIS bands closest to the published ROSAVI wavelengths. The benefit of using fixed wavelengths in a predefined VI such as ROSAVI is that the universality of the formulation and the wavelength positions can be explored across measurement scenarios (e.g., geographic locations and VZAs as in this study), as opposed to reselecting optimal bands for each image condition as for the SR and GPR approaches. It is likely that ROSAVI would benefit from an optimization routine whereby all possible band combinations are utilized/iterated with its formulation.

For all empirical model types, the best model fits and predictions were most often achieved for LCC while CCC and PAI were the poorest, respectively. This result differs from the field-based study where PAI produced the best model fits followed by CCC and then LCC. However, predictive capability for PAI was generally poorer than for CCC and LCC. This is similar to the field-based study, but contrasts with other studies showing LAI to be easier to predict than LCC [26,27,111] and to studies that hypothesize poor signal propagation from leaf to canopy, concluding that canopy scale variables (i.e., those that are physically larger) are easier to estimate [49,105,112]. The poor estimation of PAI in this study is likely related to the low stature and complexity of Arctic canopies (i.e., number of species present, species morphology, overtopping vegetation, etc.) and/or the presence/absence of gap fractions and non-photosynthetic vegetation components (e.g., lichens, senescent vegetation, moisture and soil/rock).

The VZA that produced the best models was nadir (0°). This differs from the field-based analysis, which produced the best overall results for −55°. However, model NRMSE values were very similar for all VZAs, indicating no distinct advantage of off-nadir observations. This contrasts with other studies in agricultural crops that have achieved better models with off-nadir view angles due to reduced gap fraction and shadow effects (i.e., when using backscatter observations) [113–115]. However, other multi-angle studies have concluded that shadowing is not the driving factor regarding the reliability of the reflectance signal, but instead the proportion of photosynthetic, non-photosynthetic and background components is more important [29,60]. This may hold true in Arctic environments as shadowing is relatively minimal compared to the degree of exposed soils/backgrounds (i.e., gap fractions) and non-photosynthetic materials (e.g., lichens and woody materials, etc.). In particular, ROSAVI was designed to minimize background effects when retrieving vegetation biochemical and biophysical variables [94] and has not been used extensively in Arctic environments. In this study, it was shown to perform better at nadir and positive VZAs, thus providing some evidence that such a soil-adjusted index may be working as intended within Arctic environments and may provide added benefit over traditional VIs when a greater degree of non-vegetation components are visible.

It was shown that weighted selection of optimal bands decreased in wavelength as VZAs progress from backscatter to forward scatter (for both SR and GPR). To further explore this phenomenon in the estimation/retrieval of biophysical and biochemical variables, there should be a more in-depth analysis at both the field and remote sensing scales. The
conditions under which field measurements were taken (e.g., sun-target-sensor geometry, direct/diffuse lighting) differed slightly from those of the CHRIS satellite acquisitions. Furthermore, the field measurements were integrated over various sun positions as opposed to the sun position being almost constant for the CHRIS acquisition. Further, it is possible that the ground based spectral assessment is not capturing the true reflectance behavior as compared to the satellite based spectral measurements (which is dependent on scale of observation and which incorporates atmospheric conditions and effects). As indicated by Buchhorn et al. [53], correcting for angular BRDF effects is a necessary step when creating long-term time series datasets in tundra environments; however, the results obtained in this study demonstrate that VZA-dependent band selections make less difference in vegetation-based retrievals when using a robust iterative/optimization-based modelling method such as Gaussian processes regression (GPR) with small footprint spectroscopic sensors such CHRIS. It is possible that iterative band selection algorithms such as GPR are able to account and compensate for VZA-based wavelength dependencies when estimation LCC, PAI and CCC in Arctic tundra ecosystems. Nonetheless, small estimation/retrieval errors when encountered in a time series approach can propagate and lead to spurious conclusions [18,53]; BRDF corrections should be applied where appropriate.

4.3. Physically Modelling of LCC, CCC and PAI Using PROSAIL

PROSAIL models (ranked 4th) were overall weaker than the empirical models. The LM iterative numerical optimization inversion technique has not been extensively tested with PROSAIL (e.g., [116]) or with CHRIS data, and has never been tested in Arctic tundra ecosystems to retrieve vegetation biochemical and biophysical parameters and variables from TOC multi-angle spectral reflectance (i.e., satellite-based). The LM inversion method was chosen for two reasons: (1) a recent study showed that it can produce excellent results when inverting multi-angle reflectance data for biochemical variable estimation at the remote sensing scale [116], and (2) comparative studies have shown no real advantage to using a LUT over iterative optimization methods, other than mapping speed and computation time [117,118]. However, this study and the previous field-based study [14] demonstrate that a LUT-based inversion strategy can be superior to the LM inversion method for retrieving LCC, PAI and CCC in tundra environments when regularization strategies are used to account for model ill-posedness (e.g., in this study: the addition of 10% Gaussian noise added to the simulated leaf-canopy spectra to account for sensor and environmental uncertainty; the use of the mean of the top 10% best solutions; the use of look-up tables with sufficient dimensionality to represent the range of measured and theoretical leaf and canopy conditions). Like the LM method, a LUT-based inversion approach has never been tested with satellite-based spectroscopic CHRIS data for the retrieval of LCC, PAI or CCC in tundra environments.

In this study, Arctic tundra TOC spectral reflectance was very well simulated across multiple field sites and varying sun-target-sensor geometries with LM, however, the vegetation models were generally poor. LM models for CCC were the strongest in both $r^2$ and NRMSE. This may be attributable to the information content that exists within LCC and PAI when combined; they are known to conjointly control canopy reflectance [27,29,32,119]. This agrees with the LM-based PROSAIL modelling in the previous field-based study (i.e., [14]) and with similar remote sensing based PROSAIL retrieval studies (e.g., [27]). The LUT-based approach produced the strongest models for LCC and CCC, respectively, (based on both the $r^2$ and NRMSE), with PAI having the highest normalized error. This result also agrees with the LUT-based results of the previous field study [14]. It is possible that the PROSAIL model did not adequately simulate reflectance across multiple wavelength regions. Even though relatively small in magnitude, the largest TOC reflectance errors (Figure 9) were detected across wavelength regions known to be sensitive to leaf biochemistry and leaf/canopy structure in Arctic vegetation canopies [37,39]. These errors may be attributable to the amount of noise encountered in the CHRIS spectra or through improper calibration of the PROSAIL RTM for use in Arctic ecosystems; the observed systematic error patterns point towards model
calibration issues (and towards spectral-spatial heterogeneity inherent to Arctic ecosystems). Furthermore, it is also possible that TOC reflectance is not being accurately modelled with respect to sun-target-sensor geometry since the results, stratified per vegetation variable and VZA, are variable. If the physical model is indeed fully-generalizable across measurement scenarios, then lower variability would be expected.

Overall, the poorer results of the global inversions (e.g., all sites pooled) may be explained by the inability of the one-dimension PROSAIL turbid medium model to accurately describe the radiation field of complex spatially heterogeneous canopies (e.g., gap fractions, clumping effects, multiple leaf layers having different optical characteristics, standing senescent vegetation, non-vegetated components, etc.). However, if the physical modelling assumptions are violated, it is expected that the PROSAIL model will fail outright [27,120]; in such an instance, alternative modelling assumptions should be considered (e.g., [121–123]). It has been shown by Darvishzadeh et al. [28] that short, heterogeneous grassland canopies can be simulated with PROSAIL at the remote sensing level; however, at the field scale, Darvishzadeh et al. [27] showed that the stratification of plant heterogeneity must be completed to increase accuracy. In this study and the previous field study [14], LUT modelling showed that PROSAIL can be successfully applied in tundra environments when a full range of vegetation variables are sampled (i.e., when there is a greater degree of separation between min/max vegetation variable measurements/classes), as was seen in the local inversions for Banks Island. Furthermore, it is possible that PROSAIL may not be suitable for use with high spatial resolution sensors where the naturally complex/heterogeneous canopies of Arctic tundra are more apparent. In this study, the leaf and canopy features (measured at the meter scale) were considered to be sufficiently mixed in relation to the spatial resolution of the CHRIS sensor (~34 m at nadir), thus providing justification for the use of the PROSAIL one-dimensional turbid medium model e.g., [14].

5. Conclusions

Overall conclusions drawn from this study are as follows.

1. The remote sensing modelling results produced very similar results to the field-based modelling approaches described in Kennedy et al. [14], thus providing confirmation that the spatial sampling methodology employed at the field scale was appropriate for scaling to the remote sensing level.

2. The spectroscopy data (comparison of field to CHRIS spectra) demonstrated that sun and sensor geometry are factors controlling the intensity and magnitude of Arctic vegetation reflectance curves. The observed relationships between spectra and surface features (e.g., the percent cover of photosynthetic materials, PAI) highlights the importance of sampling along environmental gradients representative of the greater ecosystem in question. Iteration based empirical modelling methodologies, such as GPR, should be tested further for their ability to compensate for sun-sensor induced variability when estimating leaf and canopy variables in Arctic environments.

3. Iteration based empirical modelling methodologies (i.e., SR and GPR) produced the best overall results. Like the field-based approach, optimal weighted multi-angle spectral regions for variable retrievals (LCC, PAI and CCC) were revealed to be largely within expected pigment absorption, red edge and NIR regions.

4. The non-iterative empirical approach based on preselected wavelengths used in a soil based VI (i.e., ROSAVI) was shown to produce the least accurate fits/model predictions; however, results were not substantially different from SR and GPR. Further investigation into the use of soil-based VIs is warranted due to the prevalence of gap fractions and non-photosynthetic materials in Arctic tundra canopies. These models should be tested iteratively using the same band selection techniques used with SR in this study.

5. The PROSAIL model produced more accurate inversion results when vegetation variables ranges were more or more separable and produced very similar results to the previous field-scale inversions. The similarity between the field and remote
sensing results provides some indication that PROSAIL is capable of being scaled to the remote sensing level in the Arctic. Analysis of the error patterns for the simulated TOC reflectance showed a systematic error pattern across field sites and VZAs. Implications of these results are: (a) it is likely that Arctic tundra canopies violate some of the assumptions behind the physical model, and (b) model calibration (e.g., refractive index and absorption coefficients of leaf constituents) should be explored for Arctic plants/canopies. Future research should consider examining three-dimensional RTMs, as the assumptions behind these approaches may be able to better account for the complex optical properties of Arctic vegetation. For example, the three-dimensional DART model [123] may be able to better simulate the spatial heterogeneity of Arctic vegetation canopies as it relates to species assemblages, senescent vegetation and gap fractions/soil contributions [124].

6. VZA-based results for all modelling types demonstrated that nadir (0° VZA) acquisition produced the best predictions of LCC, PAI and CCC; however, retrieval errors varied only slightly, indicating no distinct advantage for any one VZA.

Results are expected to improve on those presented here, particularly if more recent hyperspectral sensors such as AVIRIS-NG (Airborne Visible Infrared Imaging Spectrometer - Next Generation) [125], PRISMA, the soon to be launched EnMAP, or other sensors that are of higher quality and better calibrated than the CHRIS data that was available for this study are used. Full spectrum data that extend to the short-wavelength infrared regions (e.g., 1500–2500 nm) may provide additional benefits for retrieving and mapping biochemical and biophysical variables in Arctic environments and should be explored in future research - especially for the retrieval of non-pigment vegetation-based biochemicals such as lignin, cellulose and water [25]. Finally, the results of this study contribute to knowledge of baseline performance for hyperspectral sensors and their associated retrieval methods. It is anticipated that the next step to this study will be to use imaging spectroscopy data (both airborne and spaceborne datasets) to conduct process-based ecosystem studies, whereby relationships between Arctic plant traits and their associated environments are explored across time and space.

Author Contributions: Conceptualization, B.E.K.; methodology, B.E.K.; validation, B.E.K.; formal analysis, B.E.K.; investigation, B.E.K.; resources, B.E.K., D.J.K. and J.D.; data curation, B.E.K.; writing—original draft preparation, B.E.K.; writing—review and editing, B.E.K. and D.J.K.; visualization, B.E.K.; funding acquisition, D.J.K. and J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by funding from the Natural Sciences and Engineering Research Council of Canada and by Environment and Climate Change Canada through the Government Related Initiatives Program (GRIP).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The field data used in this study are not currently available publicly due to their ongoing use. CHRIS/PROBA data are available through the European Space Agency.

Acknowledgments: We would like to thank Mélanie Carrière, Tobias Ullmann, Brendan O’Neill and Alisa Chung for their contributions to both field and lab work. We would also like to thank Tobias Ullmann, University of Würzburg, for helping with the numerical inversion of the PROSAIL models.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Arctic Climate Impact Assessment (ACIA). Impacts of a Warming Arctic: Arctic Climate Impact Assessment; Cambridge University Press: Cambridge, UK, 2005.
2. Conservation of Arctic Flora and Fauna (CAFF). Arctic Biodiversity Assessment: Status and Trends in Arctic Biodiversity; Conservation of Arctic Flora and Fauna: Akureyri, Iceland; Narayana Press: Odder, Denmark, 2013.
3. Sturm, M.; Racine, C.; Tape, K. Increasing shrub abundance in the Arctic. Nature 2001, 411, 546–547. [CrossRef]
4. Huemmrich, K.F.; Gamon, J.A.; Tweedie, C.E.; Oberbauer, S.F.; Kinoshita, G.; Houston, S.; Kuchy, A.; Hollister, R.D.; Kwon, H.; Mano. M. Remote sensing of tundra gross ecosystem productivity and light use efficiency under varying temperature and moisture conditions. *Remote Sens. Environ.* 2010, 114, 481–489. [CrossRef]

5. Huemmrich, K.F.; Kinoshita, G.; Gamon, J.A.; Houston, S.; Kwon, H.; Oechel, W.C. Tundra carbon balance under varying temperature and moisture regimes. *J. Geophys. Res.* 2010, 115. [CrossRef]

6. Post, E.; Forchhammer, M.; Bret-Harte, M.; Callaghan, T.; Christensen, T.; Elberling, B.; Fox, A.; Gilg, O.; Hik, D.; Høye, T.; et al. Ecological Dynamics Across the Arctic Associated with Recent Climate Change. *Science* 2009, 325, 1355–1358. [CrossRef] [PubMed]

7. Myers-Smith, I.H.; Hik, D.S.; Kennedy, C.; Cooley, D.; Johnstone, J.F.; Kenney, A.J.; Krebs, C.J. Expansion of Canopy-Forming Willows Over the Twentieth Century on Herschel Island, Yukon Territory, Canada. *AMBIO* 2011, 40, 610–623. [CrossRef]

8. Fraser, R.H.; Olthof, I.; Kokelj, S.V.; Sims, R.A. Warming-Induced Shrub Expansion and Lichen Decline in the Western Canadian Arctic. *Ecosystems* 2014, 17, 1151–1168. [CrossRef]

9. Fraser, R.H.; Olthof, I.; Kokelj, S.V.; Lantz, T.C.; Lacelie, D.; Brooker, A.; Wolfe, S.; Schwarz, S. Detecting landscape changes in high-latitude ecosystems using Landsat trend analysis: 1. Visualization. *Remote Sens.* 2014, 6, 11533–11557. [CrossRef]

10. Olthof, I.; Fraser, R.H. Detecting landscape changes in high latitude environments using Landsat trend analysis: 2. Classification. *Remote Sens.* 2014, 6, 11558–11578. [CrossRef]

11. Bhatt, U.S.; Walker, D.A.; Raynolds, M.K.; Comiso, J.C.; Epstein, H.E.; Jia, G.; Gens, R.; Pinzon, J.E.; Tucker, C.J.; Tweedie, C.E.; et al. Circumpolar Arctic tundra vegetation change is linked to sea ice decline. *Earth Interact.* 2010, 14, 1–20. [CrossRef]

12. Goetz, S.J.; Bunn, A.J.; Fiske, G.J.; Houghton, R.A. Satellite-observed photosynthetic trends across North America associated with climate and fire disturbance. *Proc. Natl. Acad. Sci. USA* 2005, 102, 13521–13525. [CrossRef] [PubMed]

13. Myneni, R.B.; Keeling, C.D.; Tucker, C.J.; Asrar, G.; Nemani, R.R. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* 1997, 386, 698–702. [CrossRef]

14. Kennedy, B.E.; King, D.J.; Duffe, J. Comparison of Empirical and Physical Modelling for Estimation of Biochemical and Biophysical Vegetation Properties: Field Scale Analysis across an Arctic Bioclimatic Gradient. *Remote Sens.* 2020, 12, 3073. [CrossRef]

15. Stow, D.; Hope, A.; McGuire, D.; Verbyla, D.; Gamon, J.; Huemmrich, F.; Houston, S.; Racine, C.; Sturm, M.; Tape, K.; et al. Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems. *Remote Sens. Environ.* 2004, 89, 281–308. [CrossRef]

16. Laidler, G.J.; Teitz, P.M.; Atkinson, D.M. Remote Sensing of Arctic Vegetation: Relations between the NDVI, Spatial Resolution and Vegetation Cover on Boothia Peninsula, Nunavut. *Arctic* 2008, 61, 1–13. [CrossRef]

17. McFadden, J.P.; Chapin, F.S.; Hollinger, D.Y. Sub grid-scale variability in the surface energy balance of Arctic tundra. *J. Geophys. Res.* 1998, 103, 28947–28961. [CrossRef]

18. Gamon, J.A.; Huemmrich, K.F.; Stone, R.S.; Tweedie, C.E. Spatial and temporal variation in primary productivity (NDVI) of coastal Alaskan tundra: Decreased vegetation growth following earlier snowmelt. *Remote Sens. Environ.* 2013, 129, 144–153. [CrossRef]

19. Shippert, M.M.; Walker, D.A.; Auberbach, N.A.; Lewis, B.E. Biomass and leaf-area index maps derived from SPOT images for Toolik Lake and Ilnuvik Creek areas, Alaska. *Polar Rec.* 1995, 31, 147–154. [CrossRef]

20. Jacobsen, A.; Hansen, B. Estimation of the soil heat flux/net radiation ratio based on spectral vegetation indexes in high-latitude Arctic areas. *Int. J. Remote Sens.* 1999, 20, 445–461. [CrossRef]

21. Dungan, J. Spatial prediction of vegetation quantities using ground and image data. *Int. J. Remote Sens.* 1998, 19, 267–285. [CrossRef]

22. Lobo, A.; Moloney, K.; Chic, O.; Chiariello, N. Analysis of fine-scale spatial pattern of a grassland from re-motely-sensed imagery and field collected data. *Landsc. Ecol.* 1998, 13, 111–131. [CrossRef]

23. Ostendorf, B.; Reynolds, J.F. A model of Arctic tundra vegetation derived from topographic gradients. *Landsc. Ecol.* 1998, 13, 187–201. [CrossRef]

24. Ustin, S.L.; Gitelson, A.A.; Jacquemoud, S.; Schaepman, M.E.; Asner, G.P.; Gamon, J.A.; Zarco-Tejada, P.J. Retrieval of foliar information about plant pigment systems by high resolution spectroscopy. *Remote Sens. Environ.* 2009, 113, 567–577. [CrossRef]

25. Kokaly, R.F.; Asner, G.P.; Ollinger, S.V.; Martin, M.E.; Wessman, C.A. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sens. Environ.* 2009, 113, 578–591. [CrossRef]

26. Darvishzadeh, R.; Skidmore, A.; Schlerf, M.; Atzberger, C.; Corsi, F.; Cho, M. LAI and chlorophyll estimated for a heterogeneous grassland using hyperspectral measurements. *ISPRS J. Photogramm. Remote Sens.* 2008, 63, 409–426. [CrossRef]

27. Darvishzadeh, R.; Skidmore, A.; Schlerf, M.; Atzberger, C. Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote Sens. Environ.* 2008, 112, 2592–2604. [CrossRef]

28. Darvishzadeh, R.; Atzberger, C.; Skidmore, A.; Schlerf, M. Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models. *ISPRS J. Photogramm. Remote Sens.* 2011, 66, 894–906. [CrossRef]

29. Stagakis, S.; Markos, N.; Sykioti, O.; Kyparissis, A. Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a Phlomis fruticosa Mediterranean ecosystem using multi-angular CHRIS/PROBA observations. *Remote Sens. Environ.* 2010, 114, 977–994. [CrossRef]
30. Sykioti, O.; Paronis, D.; Stagakis, S.; Kyparissis, A. Band depth analysis of CHRIS/PROBA data for the study of a Mediterranean natural ecosystem. Correlations with leaf optical properties and ecophysiological parameters. Remote Sens. Environ. 2011, 115, 752–766. [CrossRef]

31. Verrelst, J.; Rivera, J.; Veroustraete, F.; Munoz-Mari, J.; Clevenger, J.; Camps-Valls, G.; Moreno, J. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods: A comparison. ISPRS J. Photogramm. Remote Sens. 2015, 108, 260–272. [CrossRef]

32. Verrelst, J.; Rivera, J.; Gitelson, A.; Delegido, J.; Moreno, J.; Camps-Valls, G. Spectral band selection for vegetation properties retrieval using Gaussian processes regression. Int. J. Appl. Earth Obs. Geoinf. 2016, 52, 554–567. [CrossRef]

33. Jensen, J.R. Introductory Digital Image Processing; Prentice Hall, Pearson Education Inc.: Upper Saddle River, NJ, USA, 2007.

34. Barnsley, M.J.; Settle, J.J.; Cutter, M.A.; Lobb, D.R.; Teston, F. The PROBA/CHRIS mission: A low-cost small sat for hyperspectral multiangle observations of the earth surface and atmosphere. IEEE Trans. Geosci. Remote Sens. 2004, 42, 1512–1520. [CrossRef]

35. Galeazzi, C.; Sacchetti, A.; Cisbani, A.; Babini, G. The PRISMA Program. In Proceedings of the IGARSS 2008—2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008; pp. IV-105–IV-108.

36. Guanter, L.; Kaufmann, H.; Segl, K.; Foerster, S.; Rogass, C.; Chabrillat, S.; Kuester, T.; Hollstein, A.; Rossner, G.; Chlebek, C.; et al. The EnMAP spaceborne imaging spectroscopy mission for earth observation. Remote Sens. 2015, 7, 8830–8857. [CrossRef]

37. Liu, N.; Budkewitsch, P.; Treitz, P. Examining spectral reflectance features related to Arctic percent vegetation cover: Implications for hypserspectral remote sensing of Arctic tundra. Remote Sens. Environ. 2017, 192, 58–72. [CrossRef]

38. Bratsch, S.N.; Epstein, H.E.; Buchhorn, M.; Walker, D.A. Differentiating among four Arctic tundra plant communities at Ivotuk, Alaska using field spectroscopy. Remote Sens. 2016, 8, 51. [CrossRef]

39. Buchhorn, M.; Walker, D.A.; Heim, B.; Raynolds, M.K.; Epstein, H.E.; Schwieder, M. Ground-based hyperspectral characterization of Alaska tundra vegetation along environmental gradients. Remote Sens. 2013, 5, 3971–4005. [CrossRef]

40. Davidson, S.; Santos, M.; Sloan, V.; Watts, J.; Phoenix, G.; Oechel, W.; Zona, D. Mapping Arctic Tundra Vegetation Communities using Field Spectroscopy and Multispectral Satellite Data in North Alaska, USA. Remote Sens. 2016, 8, 978. [CrossRef]

41. Hope, A.S.; Kimball, J.S.; Stow, D.A. The relationship between tussock tundra spectral reflectance properties and biomass and vegetation composition. Int. J. Remote Sens. 1993, 14, 1861–1874. [CrossRef]

42. Huemmrich, K.F.; Gamon, J.A.; Tweedie, C.E.; Campbell, P.K.; Landis, D.R.; Middleton, E.M. Arctic tundra vegetation functional types based on photosynthetic physiology and optical properties. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2013, 6, 265–275. [CrossRef]

43. Kushida, K.; Kim, Y.; Tsuyuzaki, S.; Fukuda, M. Spectral vegetation indices for estimating shrub cover, green phytomass and leaf turnover in a sedgeshrub tundra. Int. J. Remote Sens. 2009, 30, 1651–1658. [CrossRef]

44. Kushida, K.; Hobara, S.; Tsuyuzaki, S.; Kim, Y.; Watanabe, M.; Setiawan, Y.; Harada, K.; Shaver, G.R.; Fukuda, M. Spectral indices for remote sensing of phytomass, deciduous shrubs, and productivity in Alaskan Arctic tundra. Int. J. Remote Sens. 2015, 36, 4344–4362. [CrossRef]

45. Riedel, S.M.; Epstein, H.E.; Walker, D.A. Biotic controls over spectral reflectance of Arctic tundra vegetation. Int. J. Remote Sens. 2005, 26, 2391–2405. [CrossRef]

46. Riedel, S.M.; Epstein, H.E.; Walker, D.A.; Richardson, D.L.; Calef, M.P.; Edwards, E.; Moody, A. Spatial and temporal heterogeneity of vegetation properties among four tundra plant communities at Ivotuk, Alaska, USA. Arct. Antarct. Alp. Res. 2005, 37, 25–33. [CrossRef]

47. Ulrich, M.; Grosse, G.; Chabrillat, S.; Schirrmeister, L. Spectral characterization of periglacial surfaces and geomorphological units in the Arctic Lena Delta using field spectrometry and remote sensing. Remote Sens. Environ. 2009, 113, 1220–1235. [CrossRef]

48. Tieszen, L.L.; Johnson, P.L. Pigment Structure of Some Arctic Tundra Communities. In Vegetation of the Circumpolar Arctic: A Natural Resource; Albritton, W.L., Ed.; University of Alaska Press: Fairbanks, AK, USA, 1998;

49. Verrelst, J.; Rivera, J.; Gitelson, A.; Delegido, J.; Moreno, J.; Camps-Valls, G. Spectral band selection for vegetation properties retrieval using Gaussian processes regression. Int. J. Appl. Earth Obs. Geoinf. 2016, 52, 445–456. [CrossRef]
114. Walter-Shea, E.A.; Privette, J.; Cornell, D.; Mesarch, M.A.; Hays, C.J. Relations between directional spectral vegetation indices and leaf area and absorbed radiation in alfalfa. *Remote Sens. Environ.* 1997, 61, 162–177. [CrossRef]

115. Galvão, L.S.; Roberts, D.A.; Formaggio, A.R.; Numata, I.; Breunig, F.M. View angle effects on the discrimination of soybean varieties and on the relationships between vegetation indices and yield using off-nadir Hyperion data. *Remote Sens. Environ.* 2009, 113, 846–856. [CrossRef]

116. Hilker, T.; Galvão, L.S.; Aragão, L.; de Moura, Y.M.; do Amaral, C.H.; Lyapustin, A.I.; Wu, J.; Albert, L.P.; Ferreira, M.J.; Anderson, L.O.; et al. Vegetation chlorophyll estimates in the Amazon from multi-angle MODIS observations and canopy reflectance model. *Int. J. Appl. Earth Obs. Geoinf.* 2017, 58, 278–287. [CrossRef]

117. Vohland, M.; Mader, S.; Dorigo, W. Applying different inversion techniques to retrieve stand variables of summer barley with PROSPECT + SAIL. *Int. J. Appl. Earth Obs. Geoinf.* 2010, 12, 71–80. [CrossRef]

118. Preidl, S.; Doktor, D. Comparison of Radiative Transfer Model Inversions to Estimate Vegetation Physiological Status Based on Hyperspectral Data. In Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lisbon, Portugal, 6–9 June 2011.

119. Ollinger, S.V. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytol.* 2011, 189, 375–394. [CrossRef] [PubMed]

120. Meroni, M.; Colombo, R.; Panigada, C. Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations. *Remote Sens. Environ.* 2004, 92, 195–206. [CrossRef]

121. Verhoef, W.; Bach, H. Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. *Remote Sens. Environ.* 2007, 109, 166–182. [CrossRef]

122. Gastellu-Etchegorry, J.P.; Demarez, V.; Pinel, V.; Zagolski, F. Modeling radiative transfer in heterogeneous 3-D vegetation canopies. *Remote Sens. Environ.* 1996, 58, 131–156. [CrossRef]

123. Gastellu-Etchegorry, J.P.; Yin, T.; Lauret, N.; Cajigfinger, T.; Gregoire, T.; Grau, E.; Feret, J.B.; Lopes, M.; Guilleux, J.; Dedieu, G.; et al. Discrete anisotropic radiative transfer (DART 5) for modeling airborne and satellite spectroradiometer and LIDAR acquisitions of natural and urban landscapes. *Remote Sens.* 2015, 7, 1667–1701. [CrossRef]

124. Juszak, I.; Iturrate-Garcia, M.; Gastellu-Etchegorry, J.P.; Schaepman, M.E.; Maximov, T.C.; Schaepman-Strub, G. Drivers of shortwave radiation fluxes in Arctic tundra across scales. *Remote Sens. Environ.* 2017, 193, 86–102. [CrossRef]

125. Chapman, J.W.; Thompson, D.R.; Helmlinger, M.C.; Bue, B.D.; Green, R.O.; Eastwood, M.L.; Geier, S.; Olson-Duvall, W.; Lundeen, S.R. Spectral and radiometric calibration of the Next Generation Airborne Visible Infrared Spectrometer (AVIRIS-NG). *Remote Sens.* 2019, 11, 2129. [CrossRef]