L-band multi-polarization radar scatterometry over global forests: modelling, analysis, and applications

Thesis by

Yu Xian Lim

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2020

Defended 6 June 2019

Yu Xian Lim
ORCID: 0000-0002-3777-7986

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr Jakob J. van Zyl, who introduced me to this field and directly trained me from its basic foundations, and for much guidance, encouragement, and inspiration during the course of this work. A great many important lessons, including those of a technical nature and beyond, were learnt from him.

I would also like to thank the members of my defense and candidacy committees: Professor Ali Hajimiri, Professor P.P. Vaidyanathan, Professor Charles Elachi, and Professor Amnon Yariv.

My gratitude also goes to staff members both at Caltech and JPL: Katey Velazquez, Tanya Owen, Connie Rodriguez, Mariko Burgin.

Acknowledgements to my funding agency, DSO National Laboratories (Singapore) for their generous support.

Much thanks to all my friends who have supported me during my time at Caltech, including members of the Yariv group who welcomed me during my first year.

Most of all, I'm indebted to my family: my parents, for their unconditional love, sacrifice, and support throughout my life in all my endeavours, and having endured my many years of absence from home on the other side of the globe; my elder brother, whom I've always looked up to but never admitted; and my loving wife, Wei Lin, who has had to suffer my untold folly all this time while working on her own PhD.

ABSTRACT

Spaceborne L-band radars have the ability to penetrate vegetation canopies over forested areas, suggesting a potential for regular and frequent global monitoring of both the vegetation state and the subcanopy soil moisture. However, L-band radar's sensitivity to both vegetation and ground also complicates the relationship between the radar observations and the ecological and geophysical parameters. Accurate yet parsimonious forward models of the radar backscatter are valuable to building an understanding of these relationships. In the first part of this thesis, a model of L-band multi-polarization radar backscatter from forests, intended for use at regional to global spatial scales, is presented. Novel developments in the model include the consideration of multiple scattering within the dense vegetation canopy, and the application of a general model of plant allometry to mitigate the need for much intensive field data for training or over-tuning towards specific sites and tree species.

Aided by our model, in the remainder and majority of the thesis, a detailed analysis and interpretation of L-band backscatter over global forests is performed, using data from the Aquarius and SMAP missions. Quantitative differences in backscatter predicted by our model due to freeze/thaw states, branch orientation, and flooding are partially verified against the data, and fitted values of aboveground-biomass and microwave vegetation optical depths are comparable to independent estimates in the literature. Polarization information is used to help distinguish vegetation and ground effects on spatial and temporal variations. We show that neither vegetation nor ground effects alone can explain spatial variations within the same land cover class. For temporal variations during unfrozen periods, soil moisture is found to often be an important factor at timescales of a week to several months, although vegetation changes remain a non-negligible factor. We report the observation of significant differences in backscatter depending on beam azimuthal angle, possibly due to plant phototropism.

We also investigated diurnal variations, which have the potential to reveal signals related to plant transpiration. SMAP data from May-July 2015 showed that globally, co-polarized backscatter was generally higher at 6PM compared to 6AM over boreal forests, which is not what one might expect based on previous studies. Based on our modelling, increased canopy extinction at 6AM is a possible cause, but this is unproven and its true underlying physical cause undetermined.

Finally, by making simplifying approximations on our forward model, we propose and explore algorithms for soil moisture retrieval under forest canopies using L-band scatterometry, with preliminary evaluations suggesting improved performance over existing algorithms.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Background and overview

The ability to model and predict feedbacks in a changing climate requires many inputs, including knowledge of where, when, and how much carbon, water, and energy are stored and exchanged between the land surface and the atmosphere. Many of these interactions take place in forests, which are major carbon sinks, and contain a majority of global plant biomass [1]. However, there is significant uncertainty in the knowledge of some of these variables and their fluxes. Importantly, soil moisture, a key element in evapotranspiration, climate state, weather and landslide prediction, and flood and drought monitoring [2], is poorly measured in forests [3, 4]. There is also significant uncertainty in the total biomass carbon stock of forests, the rate at which it is changing due to deforestation and regrowth, and their associated spatial distributions [5, 6, 7].

Spaceborne remote sensing offers a good platform to study such key elements pertinent to the understanding of our climate and environment on a global scale. In particular, microwave remote sensing enables timely monitoring without interruption by cloud cover and can measure geophysical parameters that are complementary to the visible-infrared part of the electromagnetic spectrum. Over forested land areas, long wavelength microwaves, e.g. at L-band, can penetrate the vegetation canopy and offer sensitivity to both the vegetation as well as to moisture in the ground under the canopy [8, 9], holding promise as a tool for learning about changes in their state and related processes if multiple and frequent observations over time are available.

To date, both active and passive L-band instruments have been operated by spaceborne Earth observatories. While active radars and passive radiometers each have their own uses, one difference between them is in spatial resolution. Practical constraints on antenna size place a limit on the spatial resolution obtainable by passive radiometers, whereas radars using

synthetic aperture techniques allow finer spatial resolutions to be tailored according to specific application requirements.

Amongst the spaceborne L-band radar missions that have flown, two stand out in terms of temporal coverage and polarization information: Aquarius and SMAP (Soil Moisture Active Passive). Aquarius has a 7-day repeat orbit, while SMAP has global coverage every 2-3 days. They also offer the advantage of full multi-polarization measurements, sending and receiving in both horizontal and vertical polarizations. Polarization information has long been recognized to be useful in helping to distinguish different scattering mechanisms involving the vegetation and ground [10, 11, 12, 13]. This is important because the sensitivity of L-band radar to both vegetation and ground is a double-edged sword that complicates the retrieval of their ecological and geophysical parameters from the radar observations.

Having motivated the study of forested areas and identified the opportunity provided by the L-band multi-polarization radars of Aquarius and SMAP, in this thesis we study L-band radar backscatter from forests to better understand its spatial and temporal relationships and sensitivities to the underlying physical conditions, with the aid of polarization information. It is worthwhile to mention that Aquarius was not a mission that was originally intended for observations over land, but instead for the measurement of ocean salinity. SMAP, though focused on soil moisture, was not primarily targeted at forested areas, due in part to limitations in understanding the relationships between the L-band measurements and subcanopy soil moisture in the presence of significant intervening vegetation. On the other hand, there have been previous studies dedicated to studying L-band radar backscatter from forests, but these have been over specific, localized forest stands in the context of airborne synthetic aperture radar (SAR) experiments [9, 14, 15, 16]. There is thus significant potential in exploiting the Aquarius and SMAP radar measurements over forests for analysis and interpretations at the regional to global scale.

One approach towards performing such an analysis and deepening our understanding is to model the radar backscatter, and there has been much work in the literature on developing such models [14, 17, 16, 18, 19, 20, 21]. Many of the previous models were applied and evaluated on specific forest stands where detailed measurements of the tree architecture,

dielectric constants, and ground conditions were made concurrently with airborne SAR measurements. These efforts helped to validate many of the concepts and approaches used in the models. Extending the modelling effort to a global scale, Kim et al. [22] have developed physical models of L-band radar backscatter for application to soil moisture retrieval. In particular, their model for forests was based on work by Tabatabaeenejad et al. [9] and Burgin et al. [21], which required a detailed and complete characterization of the geometry of vegetation structures as input. To reduce the number of free input parameters required, each forest land cover class was modelled with representative species, and species-specific allometric relations were applied to relate all vegetation parameter to a single free input parameter – vegetation water content (VWC). In-situ samples from field measurements were also used as training data to tune some geometric parameters within the model. Two other input parameters, soil surface root mean square (RMS) height and soil dielectric constant, describe the ground. Such parsimony was essential for model inversion and parameter estimation from limited data – only three measurement channels (HH, HV, VV polarization) from the radar of the Soil Moisture Active Passive (SMAP) mission.

In this thesis, we shall present in Chapter 2 an L-band forward model of multi-polarization radar backscatter from forests that is intended for application at regional to global spatial scales. The term "scatterometry" in the title was chosen to suggest the greater emphasis on accurate backscatter cross-section values, as opposed to the regime of imaging SARs with very high resolution at the scale of metres, but poorer radiometric accuracy. Our modelling approach builds upon many of these earlier ideas and foundations, and is similar to that by Kim et al. [22], Burgin et al. [21] and Durden et al. [14] in many respects, but we made significant novel developments. Identifying the distribution of the size of tree trunks and branches as key contributors to L-band scattering, we extend the afore-mentioned allometry strategy for model parsimony by applying a general model of plant allometry to obtain this distribution, mitigating the need for much intensive field data for training and/or over-tuning towards specific sites and tree species. Unlike some previous models, our model also takes into account the significant multiple-scattering within dense forest canopies by introducing a correction factor computed using radiative transfer. This correction factor is important in

providing accurate agreement between the model and data, especially when the dynamic range in the radar backscatter from forests is only several decibels. Further modelling details are described in Chapter 2.

Equipped with our forward model which relates radar backscatter to input parameters from the forest and ground, in Chapter 3 we apply it to the analysis of global forest L-band multi-polarization backscatter data from the Aquarius scatterometer. We show that our model appears to be consistent with the data overall, and interpret spatial and temporal variations in the radar backscatter in terms of ground and vegetation factors. Neither ground nor vegetation factors alone suffices to explain the spatial variance within the same land cover class. For temporal variations during unfrozen periods, soil moisture may be a primary factor at timescales of one week to several months. Vegetation changes remain a non-negligible factor, and for larger incidence angles over deciduous needleleaf forests may even become the primary factor at longer timescales (months). Differences in L-band radar due to freeze/thaw states, branch orientation, and flooding were also partially verified quantitatively with our model.

In Chapter 4, we analyze diurnal variations in L-band multi-polarization backscatter from forests using SMAP data. Transpiration and related plant processes follow a diurnal cycle and there is potential for monitoring vegetation water status using radar [23, 24, 25, 26, 27]. We find that the co-polarized L-band radar backscatter observed in late spring-summer over the northern boreal forests is higher at 6PM than 6AM, which is not what one might expect based on previous studies. Based on our modelling, increased canopy extinction at 6AM is a possible cause, but this is unproven and its true underlying physical cause is undetermined. Aside from the diurnal variations, we also report the observation of significant differences in backscatter due to beam azimuthal angle, possibly associated with plant phototropism.

As recognized in attempts by Kim et al. [22] and Tabatabaeenejad et al. [9], soil moisture remote sensing from under dense vegetation remains a challenge. Existing radar-based and radiometer-based algorithms by the SMAP mission primarily focus on areas with VWC<5kg/m^2, which excludes most forests. In Chapter 5, we make simplifying approximations of our forward model to obtain a linear relationship between HH-polarized

radar backscatter and subcanopy soil moisture, under certain conditions. Based on this linear relationship, we propose some algorithms for soil moisture retrieval from forests using L-band radar. Preliminary evaluations suggest improved performance over existing algorithms.

The subsequent sections in this chapter will briefly review some preliminary foundations that are essential for subsequent chapters. The content can be found in most radar textbooks, though notations and conventions vary; here we largely follow the notations and conventions adopted by Ulaby and Long [28]. In Section 1.2, we establish the scattering geometry and coordinate system, and define the 2x2 scattering matrix. Section 1.3 reviews the radar cross-section for a point target and a distributed target. Section 1.4 covers the scatterer covariance matrix, and Section 1.5 the optical theorem. We shall not derive these concepts from first principles; rather, we merely provide the definitions to set the stage for their use in the rest of the thesis.

1.2 Radar scattering geometry and scattering matrix

Here we establish the notations, geometries, and coordinate systems to be used. Consider a single point scatterer, or target, located at the origin with the scattering geometry as in Figure 1.1. The incident direction is $\hat{\mathbf{k}}_i$, the scattered direction is $\hat{\mathbf{k}}_s$, with

$$\hat{\mathbf{k}}_i = \sin\theta_i \cos\phi_i \, \hat{\mathbf{x}} + \sin\theta_i \sin\phi_i \, \hat{\mathbf{y}} - \cos\theta_i \, \hat{\mathbf{z}} \tag{1.1}$$

$$\hat{\mathbf{k}}_s = \sin\theta_s \cos\phi_s \, \hat{\mathbf{x}} + \sin\theta_s \sin\phi_s \, \hat{\mathbf{y}} + \cos\theta_s \, \hat{\mathbf{z}} \tag{1.2}$$

$$\hat{\mathbf{h}}_i = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_i}{\left|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_i\right|}, \qquad \hat{\mathbf{v}}_i = \hat{\mathbf{h}}_i \times \hat{\mathbf{k}}_i, \qquad \hat{\mathbf{h}}_s = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_s}{\left|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_s\right|}, \qquad \hat{\mathbf{v}}_s = \hat{\mathbf{h}}_s \times \hat{\mathbf{k}}_s . \tag{1.3}$$

This definition of coordinates is known as the "FSA (forward-scattering alignment) convention", or so-called wave coordinates [28].

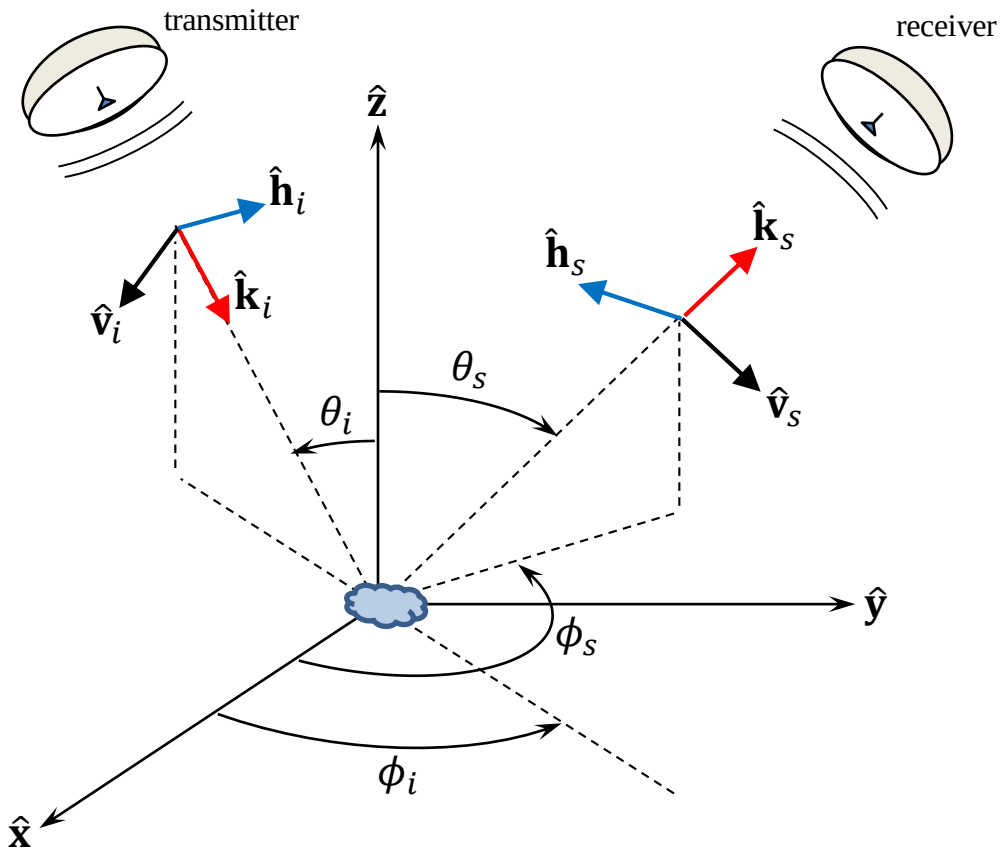

Figure 1.1. Geometry for the FSA (forward-scattering alignment) convention.

Let

$$\mathbf{E}_i = \left(E_h^i \hat{\mathbf{h}}_i + E_v^i \hat{\mathbf{v}}_i\right) \exp\left[ik\hat{\mathbf{k}}_i \cdot \mathbf{r} - \omega t\right] \tag{1.4}$$

be an incident plane wave propagating in the $\hat{\mathbf{k}}_i$ direction, where $k = 2\pi/\lambda$ is the wavenumber, and $\omega = 2\pi f$ is the angular frequency. $\hat{\mathbf{h}}_i$ and $\hat{\mathbf{v}}_i$ are mutually orthogonal unit vectors also orthogonal to $\hat{\mathbf{k}}_i$, as defined in (1.3) and Figure 1.1. Likewise, let

$$\mathbf{E}_s = \left(E_h^s \hat{\mathbf{h}}_s + E_v^s \hat{\mathbf{v}}_s\right) \exp\left[ik\hat{\mathbf{k}}_s \cdot \mathbf{r} - \omega t\right] \tag{1.5}$$

be the scattered field (in the far-field) in the $\hat{\mathbf{k}}_s$ direction, with $\hat{\mathbf{h}}_s$ and $\hat{\mathbf{v}}_s$ mutually orthogonal unit vectors also orthogonal to the scattered direction $\hat{\mathbf{k}}_s$. The terms "h-polarized" and "v-

polarized" will be used frequently with regards to the incident or scattered fields. The relation between the scattered and incident fields can be conveniently written in matrix form

$$\begin{bmatrix} E_h^s \\ E_v^s \end{bmatrix} = \frac{1}{kr} \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \begin{bmatrix} E_h^i \\ E_v^i \end{bmatrix} \tag{1.6}$$

where

$$[\mathbf{S}] = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \tag{1.7}$$

is the (dimensionless) scattering matrix for the single point scatterer (some other authors in the literature prefer to put the factor of $1/k$ within the matrix rather than outside, such that the scattering matrix has dimensions of length). The four entries of this 2x2 matrix are the scattering amplitudes for the respective polarizations. Note that the scattering matrix is dependent on the choice of polarization basis vectors $\hat{\mathbf{h}}_i, \hat{\mathbf{v}}_i, \hat{\mathbf{h}}_s, \hat{\mathbf{v}}_s$. These equations are written for linearly-polarized incident waves, but arbitrary elliptical polarizations can be readily accommodated with suitable complex phases.

Figure 1.1, and equations (1.1) to (1.3) represent a common choice of coordinate basis, the FSA convention. Another choice of coordinate basis is the "BSA (backscatter alignment) convention", or so-called antenna coordinates [28]:

$$\hat{\mathbf{k}}_i = \sin\theta_i \cos\phi_i \, \hat{\mathbf{x}} + \sin\theta_i \sin\phi_i \, \hat{\mathbf{y}} - \cos\theta_i \, \hat{\mathbf{z}} \tag{1.8}$$

$$\hat{\mathbf{k}}_s = -\sin\theta_s \cos\phi_s \, \hat{\mathbf{x}} - \sin\theta_s \sin\phi_s \, \hat{\mathbf{y}} - \cos\theta_s \, \hat{\mathbf{z}} \tag{1.9}$$

$$\hat{\mathbf{h}}_i = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_i}{\left|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_i\right|}, \qquad \hat{\mathbf{v}}_i = \hat{\mathbf{h}}_i \times \hat{\mathbf{k}}_i, \qquad \hat{\mathbf{h}}_s = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_s}{\left|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_s\right|}, \qquad \hat{\mathbf{v}}_s = \hat{\mathbf{h}}_s \times \hat{\mathbf{k}}_s. \tag{1.10}$$

The BSA geometry is shown in Figure 1.2.

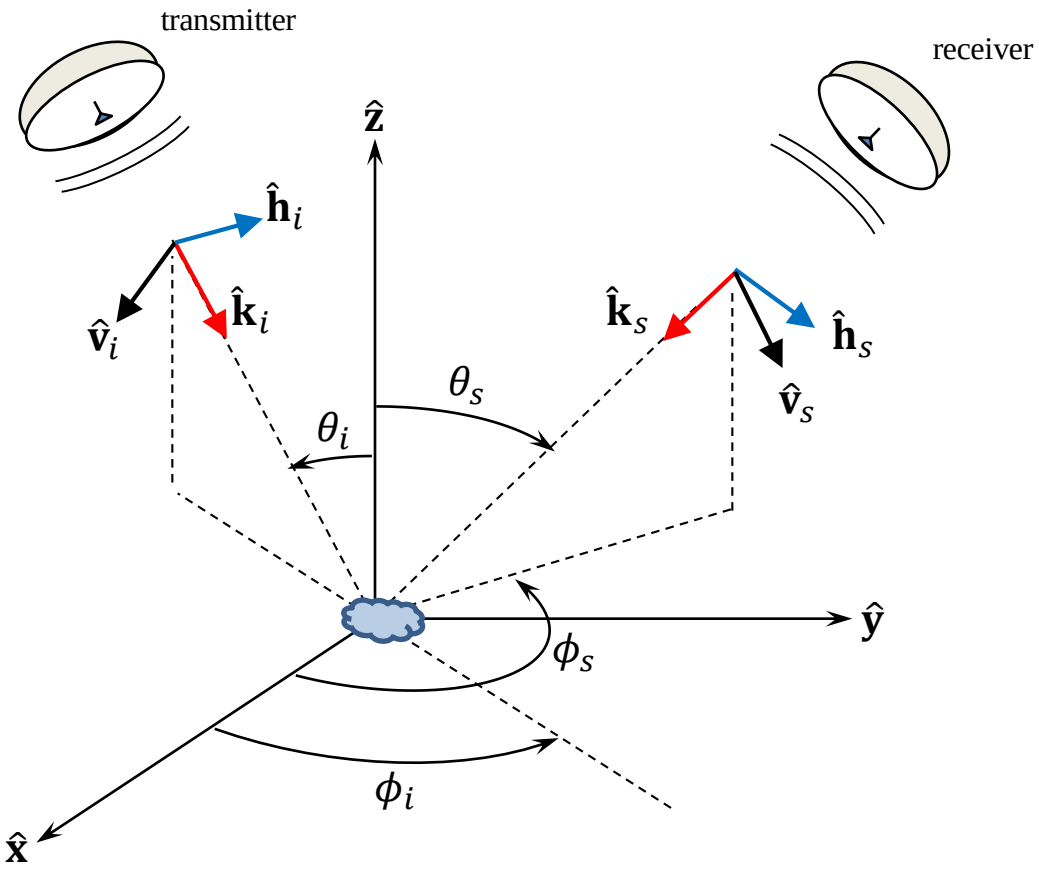

Figure 1.2. Geometry for the BSA (backscatter alignment) convention.

Throughout this thesis, both conventions are used. The BSA convention is more convenient when in the backscatter (monostatic) case, because then $\hat{\mathbf{k}}_i = \hat{\mathbf{k}}_s$, while FSA is more convenient in the bistatic case when the incident and scattered directions are arbitrary. The scattering matrices in either convention can be converted to the other convention simply:

$$[\mathbf{S}(\theta_i, \phi_i, \theta_s, \phi_s)]_{BSA} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} [\mathbf{S}(\theta_i, \phi_i, \theta_s, \phi_s)]_{FSA} . \qquad (1.11)$$

Whenever the scattering matrix is used, the convention used (BSA or FSA) will not be labelled as a subscript, but will instead be clarified in the text. Note that the quantities $|S_{hh}|^2$, $|S_{hv}|^2$, $|S_{vh}|^2$, $|S_{vv}|^2$ are the same in either convention.

1.3 Radar cross-sections and distributed targets

Suppose the incident radar is h-polarized and illuminates the point target with a power density of $\mathcal{S}_h^i$, with units of W/m^2 (power per unit area, where the area is normal to the propagation direction). The power dP_v^s scattered into the far-field into an infinitesimal solid angle $d\Omega$ in the direction $\hat{\mathbf{k}}_s$, in the v-polarization, is then

$$dP_v^s = \frac{|S_{vh}|^2}{k^2}\mathcal{S}_h^i d\Omega\,. \tag{1.12}$$

Analogous equations hold for other combinations of incident and scattered polarizations. The radar scattering cross-section, σ_{vh}, also called the radar cross-section or RCS, is by convention [29, 30] 4π of the scattered power per unit solid angle per unit illumination intensity

$$\sigma_{vh} = 4\pi\frac{|S_{vh}|^2}{k^2} \tag{1.13}$$

and has dimensions of area. Again, analogous equations hold for other combinations of incident and scattered polarizations.

So far all discussion has been concerning a single point target. In cases of a distributed target or collection of scatterers extending over an area, such as the ground, the normalized radar cross-section, σ_{vh}^0, (NRCS, also referred to as differential scattering coefficient) is defined as the radar cross-section per unit horizontal ground area [28]

$$\sigma_{vh}^0 = \frac{\langle\sigma_{vh}\rangle}{A} \tag{1.14}$$

where A is a unit horizontal ground area, and $\langle\sigma_{vh}\rangle$ is the average value of σ_{vh} over that area. σ_{vh}^0 is dimensionless and as usual, analogous equations hold for other combinations of incident and scattered polarizations. For a thin horizontal layer of randomly distributed identical small scatterers, with an average scatterer number density of n (i.e. n has units of m^{-3}) and layer thickness of Δz,

$$\sigma_{vh}^0 = \frac{4\pi n\Delta z}{k^2}\langle|S_{vh}|^2\rangle\,. \tag{1.15}$$

If the scatterers are not identical, equation (1.15) still holds with the angular brackets being interpreted as also averaging over the distribution of scatterers.

As a shorthand notation, instead of $\sigma^0_{hh}, \sigma^0_{vh}, \sigma^0_{hv}, \sigma^0_{vv}$, often HH, VH, HV, VV are used instead for the normalized radar cross-sections:

$$\text{HH is shorthand for } \sigma^0_{hh} \tag{1.16a}$$

$$\text{HV is shorthand for } \sigma^0_{hv} \tag{1.16b}$$

$$\text{VH is shorthand for } \sigma^0_{vh} \tag{1.16c}$$

$$\text{VV is shorthand for } \sigma^0_{vv} \tag{1.16d}$$

Due to the large dynamic range of possible normalized radar cross-sections and abundance of factors that combine multiplicatively, decibels (dB) are often used instead of linear units. They can be readily converted via

$$(quantity\ in\ decibels)\text{dB} = 10\log_{10}(quantity\ in\ linear\ units)\,. \tag{1.17}$$

1.4 Scatterer covariance matrix

As equation (1.12) suggests, to relate incident and scattered powers or intensities, second-order terms e.g. $S_{hh}S^*_{hh}$, $S_{hv}S^*_{hv}$, etc. are required. By re-arranging the scattering amplitudes from a matrix into a vector, these second-order terms can be collected into a 4x4 scatterer covariance matrix

$$[\mathbf{C}] = \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix} \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix}^{\dagger} = \begin{bmatrix} S_{hh}S^*_{hh} & S_{hh}S^*_{hv} & S_{hh}S^*_{vh} & S_{hh}S^*_{vv} \\ S_{hv}S^*_{hh} & S_{hv}S^*_{hv} & S_{hv}S^*_{vh} & S_{hv}S^*_{vv} \\ S_{vh}S^*_{hh} & S_{vh}S^*_{hv} & S_{vh}S^*_{vh} & S_{vh}S^*_{vv} \\ S_{vv}S^*_{hh} & S_{vv}S^*_{hv} & S_{vv}S^*_{vh} & S_{vv}S^*_{vv} \end{bmatrix} \tag{1.18}$$

where the † symbol stands for conjugate transpose. For the case of backscatter (i.e. $\theta_s = \theta_i\,, \phi_s = \phi_i + \pi$), reciprocity dictates that $S_{hv} = -S_{vh}$ in FSA and $S_{hv} = S_{vh}$ in BSA [31]. Hence for backscatter in BSA, often the 3x3 scatterer covariance matrix is used instead:

$$[\mathbf{C}] = \begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix} \begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix}^{\dagger} = \begin{bmatrix} S_{hh}S_{hh}^{*} & \sqrt{2}\,S_{hh}S_{hv}^{*} & S_{hh}S_{vv}^{*} \\ \sqrt{2}\,S_{hv}S_{hh}^{*} & 2\,S_{hv}S_{hv}^{*} & \sqrt{2}\,S_{hv}S_{vv}^{*} \\ S_{vv}S_{hh}^{*} & \sqrt{2}\,S_{vv}S_{hv}^{*} & S_{vv}S_{vv}^{*} \end{bmatrix}. \quad (1.19)$$

The diagonal entries of the covariance matrix are proportional to the respective normalized radar cross-sections, but the covariance matrix contains additional information about the relative phase between different polarizations in the off-diagonal entries.

1.5 Optical theorem

The optical theorem (see e.g. [32], [30], [33]), also known as the forward scattering theorem, relates the extinction to the forward scattering amplitude rather generally for scattering phenomena. Newton [34] traces its more-than-a-century old history. To state it for our case, the extinction cross-section for a single scatterer is, for h-polarized and v-polarized incident radar beams respectively,

$$\sigma_{ext,h}(\theta_i, \phi_i) = \frac{4\pi}{k^2}\,\mathrm{Imag}[S_{hh}(\theta_i, \phi_i, \theta_s = \pi - \theta_i, \phi_s = \phi_i)] \quad (1.20a)$$

$$\sigma_{ext,v}(\theta_i, \phi_i) = \frac{4\pi}{k^2}\,\mathrm{Imag}[S_{vv}(\theta_i, \phi_i, \theta_s = \pi - \theta_i, \phi_s = \phi_i)] \quad (1.20b)$$

where the forward scattering amplitudes are in FSA.

For instance, the intensity or power of a h-polarized incident beam propagating in the $\hat{\mathbf{x}}$ direction through a cloud of identical scatterers with number density n (n has units of m^{-3}), decays as

$$I_h(x) = I_h(x = 0)\exp\left[-n\sigma_{ext,h}\left(\theta_i = \frac{\pi}{2}, \phi_i = 0\right)x\right] \quad (1.21)$$

and likewise for v-polarization. If the scatterers are not identical, we can assign an average cross-section over the distribution of scatterers such that we retain the expression

$$I_h(x) = I_h(x = 0)\exp\left[-n\,\langle\sigma_{ext,h}\left(\theta_i = \frac{\pi}{2}, \phi_i = 0\right)\rangle\,x\right] \quad (1.22)$$

where the angular brackets denote averaging over the distribution of scatterers.

The electric field amplitude is proportional to the square root of the intensity; for ease of future notation, we introduce the corresponding average-per-scatterer extinction cross-section for the field

$$\langle \kappa_h(\theta_i, \phi_i) \rangle = \frac{1}{2} \langle \sigma_{ext,h}(\theta_i, \phi_i) \rangle = \frac{2\pi}{k^2} \langle \mathrm{Imag}[S_{hh}(\theta_i, \phi_i, \theta_s = \pi - \theta_i, \phi_s = \phi_i)] \rangle \quad (1.23)$$

and likewise for v-polarization.

Chapter 2

MODELLING OF L-BAND RADAR BACKSCATTER FROM FORESTS

2.1 Overview of modelling approach

In this chapter, a model to simulate the L-band radar backscatter from forests will be presented. As mentioned in the introduction, this model is intended primarily for regional to global spatial scales. Globally, there is a wide variation in the density and types of vegetation cover over land. We shall use the land cover classification adopted by the International Geosphere-Biosphere Program (IGBP) [35]. (A map of global land cover in the IGBP classification scheme can be found in Figure 3.5.) Classes 1-5 correspond to forests classes. Numerous different species of trees and inhomogeneous distributions are expected within the radar footprint. Instead of attempting to describe them all in detail, broad simplifications are made here. The forest is modelled as a homogenous layer of randomly oriented dielectric cylinders corresponding to crown canopy branches (leaves and structures of other morphologies are neglected), and a lower homogeneous layer of preferentially vertically oriented dielectric cylinders corresponding to tree trunks. Four separate terms contributing to the radar backscatter, corresponding to different scattering mechanisms, are considered. They are: 1. Backscatter from the crown canopy layer only; 2. direct backscatter from the ground; 3. double-reflections off the ground and canopy layer; 4. double-reflections off the ground and tree trunks. This is depicted schematically in Figure 2.1. These terms are then added incoherently together to give the total backscatter. For each polarization, the radar cross section per unit ground area is represented as:

$$\sigma^0 = \sigma^0_{cn} + \sigma^0_{gnd,direct} + \sigma^0_{cn-gnd,db} + \sigma^0_{trk-gnd,db} \,. \qquad (2.1)$$

Direct backscatter from the trunk layer is negligible because of the preferential vertical orientation of the trunks.

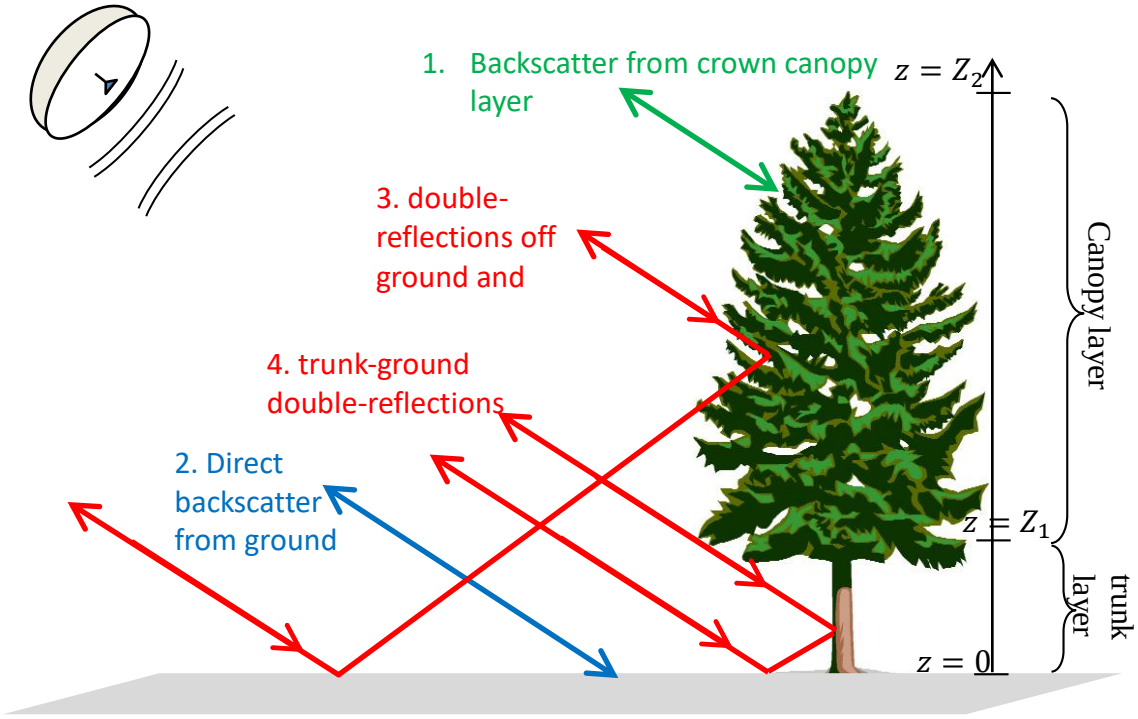

Figure 2.1. Model of L-band radar backscatter from natural vegetation, as the incoherent sum of four separate terms corresponding to different scattering mechanisms.

The modelling approach adopted here largely follows that taken by Durden et al. [14] and Burgin et al. [21], with some important novel developments. This approach required a detailed specification of the sizes and geometries of all the cylinders involved, which typically involved intensive measurements of several sample trees at specific field study sites. In those experiments, such a complete characterization was ideal for direct validation of the modelling approach against concurrent SAR measurements. However, because we intend to apply our model to many different parts of the globe, usually without detailed knowledge of the required physical and geometric parameters, we try to keep the model with as few free input parameters as possible. The importance of this parsimony was also emphasized by Tabatabaeenejad et al. [9] and Kim et al. [22], bearing in mind that in application, model inversion and parameter estimation are expected to be performed from limited measurement channels – just the three polarization channels HH, HV, VV in the case of the Aquarius and SMAP radars. A novel development in our approach is to go beyond the

species-specific allometry in Tabatabaeenejad et al. [9] and Kim et al. [22], and apply a general plant allometry model, which thus also mitigates the need for much intensive field data for training and/or over-tuning towards specific sites and samples; in both a literal and figurative sense, to not lose the forest for the trees. This will be further explained and elaborated upon in Section 2.7 and Section 2.8.

The retained free input parameters to our model are: the total canopy branch volume per unit ground area, $V_{b,tot}$, in units of m^3/m^2; the relative permittivity of the vegetation, ε_v; the relative permittivity of the ground, ε_g; and the RMS height of the roughness of the ground surface, h, in units of m. There are additional parameters pertaining to the orientation distribution of the cylinders that can be chosen, but these will not be freely varying.

Another important novel departure we make from Durden et al. [14] and Burgin et al. [21] is that, for the backscatter term from the canopy layer only, while we still employ a single-scattering approximation for computational simplicity, we make a correction for multiple scattering effects within the canopy; this correction is particularly significant for forests.

Details for computing each of the four modelled terms are elaborated in the sections below.

2.2 Backscatter from crown canopy layer only

This section describes the computation of the first term contributing to the overall radar backscatter, $\sigma^0_{cn,pol}$ for $pol = hh, vv, hv$, or in shorthand, $\mathrm{HH}_{cn}, \mathrm{HV}_{cn}, \mathrm{VV}_{cn}$. The canopy layer is modelled as a homogeneous layer of randomly oriented dielectric cylinders with some specified size distribution and orientation distribution. The cylinder scatterers are considered as "point particles" and the effect of their size appears only through their scattering matrix. The relatively small volume of cylinders (compared to air around it) allows the approach known in the literature as the distorted Born approximation [36, 37]. In this approximation, the scatterers are embedded in an equivalent medium with a complex relative permittivity $\varepsilon_{\mathrm{eff}}$ that gives agreement with the approximate mean field. Using Foldy's approximation, and taking air to have a refractive index of 1 and all magnetic relative permeabilities to be 1, the effective refractive index of the equivalent medium is related to the forward scattering amplitude of the scatterers [36, 37, 38], and is polarization dependent:

$$\sqrt{\varepsilon_{\text{eff},h-pol}} = 1 + \frac{2\pi}{k^3} n_{cn} \langle S_{hh}(\theta_i, \phi_i, \theta_s = \pi - \theta_i, \phi_s = \phi_i) \rangle \qquad (2.2a)$$

$$\sqrt{\varepsilon_{\text{eff},v-pol}} = 1 + \frac{2\pi}{k^3} n_{cn} \langle S_{vv}(\theta_i, \phi_i, \theta_s = \pi - \theta_i, \phi_s = \phi_i) \rangle \qquad (2.2b)$$

where n_{cn} has units of inverse volume and is the number density of scatterers (i.e. cylinders) in the canopy layer, and $k = 2\pi/\lambda$ is the wavenumber in air. Here the forward scattering amplitudes are in the FSA convention (see Section 0), and angular brackets denote averaging over the cylinder distribution. The cylinders "see" the incident radar wave modified correspondingly by an extinction and phase delay, and make a single scattering back to the radar receiver. (Multiple-scatterings are accounted for by a correction later.) Cylinders deeper in the layer make less contribution to the total backscatter due to extinction from higher parts. The electric field decays exponentially with depth with an extinction coefficient (for h- and v-polarizations respectively) related to the imaginary part of the forward scattering amplitude for the cylinders

$$n_{cn} \langle \kappa_{h,cn}(\theta_i, \phi_i) \rangle = n_{cn} \frac{2\pi}{k^2} \langle \text{Imag}[S_{hh}(\theta_i, \phi_i, \theta_s = \pi - \theta_i, \phi_s = \phi_i)] \rangle \qquad (2.3a)$$

$$n_{cn} \langle \kappa_{v,cn}(\theta_i, \phi_i) \rangle = n_{cn} \frac{2\pi}{k^2} \langle \text{Imag}[S_{vv}(\theta_i, \phi_i, \theta_s = \pi - \theta_i, \phi_s = \phi_i)] \rangle \qquad (2.3b)$$

consistent with the optical theorem (Section 2.4); the extinction coefficient for the intensity is twice that for the field. The phase delay can be found from the real part of the forward scattering amplitude, and needs to be considered if it is polarization-dependent and phase differences between polarizations are required. Otherwise if only normalized radar cross-sections are required, this phase can be ignored.

Scattering matrices for both the forward and backward scattering from the cylinders are required. For a single cylinder with relative permittivity ε_v, radius r, length L, and orientated with cylinder axis direction (θ_c, ϕ_c), the 2x2 dimensionless bistatic scattering matrix $[\mathbf{S}(\theta_i, \phi_i, \theta_s, \phi_s, \theta_c, \phi_c, r, L)]$ for incidence from direction (θ_i, ϕ_i) and scattering into direction (θ_s, ϕ_s) is provided in Appendix A (equation (A.2)), which follows closely the expressions given by van Zyl and Kim [11] and Bohren and Huffman [30]. The incidence

and scattering angles (θ_i, ϕ_i) and (θ_s, ϕ_s) may follow either the FSA or BSA conventions (see Section 1.2); recall that the 2x2 scattering matrix in the BSA and FSA conventions only differ by a change of sign in S_{hh} and S_{hv} (see equation (1.11)). The relative permittivity within each cylinder is assumed to be homogeneous, and all cylinders in the vegetation layer are assumed to have the same relative permittivity, so we do not display the dependence of $[\mathbf{S}]$ on ε_v.

We can now discuss in further detail the averaging operations over cylinder distribution that have been referred to with angular brackets. Let $p(\theta_c, \phi_c, r, L)$ be the density function of the cylinder size and orientation distribution, satisfying the normalization

$$\int_{r,L} \int_{\theta_c,\phi_c} p(\theta_c, \phi_c, r, L) \sin\theta_c \, d\theta_c d\phi_c dr dL = 1 \, . \tag{2.4}$$

Averaging over the cylinder distribution thus means, for instance,

$$\langle [\mathbf{S}(\theta_i, \phi_i, \theta_s, \phi_s)] \rangle = \int_{r,L} \int_{\theta_c,\phi_c} [\mathbf{S}(\theta_i, \phi_i, \theta_s, \phi_s, \theta_c, \phi_c, r, L)] p(\theta_c, \phi_c, r, L) \sin\theta_c \, d\theta_c d\phi_c dr dL \, . \tag{2.5}$$

For simplicity, the size and orientation of the cylinders are modelled as being independent, i.e. $p(\theta_c, \phi_c, r, L)$ can be factorized accordingly. Integrals over the cylinder orientations (θ_c, ϕ_c) are performed first, and then over the cylinder sizes (r, L). The specific choice of cylinder size distribution will be elaborated upon in Section 2.7. The orientation distribution for the upper crown canopy layer is chosen to be uniform in ϕ_c, and from the family parameterized by a reference angle θ_0 and an exponent m [39]

$$p(\theta_c, \phi_c) = \frac{|\cos^2(\theta_c - \theta_0)|^m}{2\pi \int_{\theta_c=0}^{\theta_c=\pi} \sin\theta_c \, |\cos^2(\theta_c - \theta_0)|^m d\theta_c} \, . \tag{2.6}$$

For values of m not too large, the distribution is quite broad and smooth over θ_c and ϕ_c. The 4π sphere is discretized into a uniform angular grid over θ_c and ϕ_c for numerical computation, and the integrals computed by approximating them as simple Riemann sums. For extinction calculations, a 5° angular discretization suffices. For backscatter and bistatic

scattering calculations, an angular discretization of $0.35\lambda/\pi L$ was used. This angular discretization is chosen in view of the width of the sinc function in equation (A. 1) in Appendix A. If $0.35\lambda/\pi L$ is greater than 5°, the angular discretization is set at 5° instead. When λ/L is small, $[\mathbf{S}]$ is small everywhere apart from the vicinity of the forward scattering cone (see Appendix A). Thus when λ/L is small, backscatter is dominated by cylinders oriented perpendicular to the incident direction, and computation can be accelerated by neglecting cylinders that are not within an angle of $2\lambda/L$ from a perpendicular orientation. In view of the cylinder orientation distribution being uniform in azimuth angle ϕ_c, we can arbitrarily pick $\phi_i = 0$. For backscatter, $\phi_s = \phi_i + \pi$ and $\theta_s = \theta_i$. The direct single-scattering backscatter from the whole canopy layer is found by summing the contributions of all the single-cylinder radar backscattering cross-section values over the distribution of cylinder sizes and orientations, and considering the two-way extinction as a function of depth. The normalized backscatter radar cross-sections for single-scattering from the canopy layer are

$$\mathrm{HH}_{cn,ss} = \frac{4\pi n_{cn}}{k^2} \int_{z=Z_1}^{z=Z_2} \exp\left(-\frac{4n_{cn}\langle\kappa_{h,cn}(\theta_i)\rangle(Z_2 - z)}{\cos\theta_i}\right)\langle|S_{hh}(\theta_i)|^2\rangle dz \qquad (2.7a)$$

$$\mathrm{VV}_{cn,ss} = \frac{4\pi n_{cn}}{k^2} \int_{z=Z_1}^{z=Z_2} \exp\left(-\frac{4n_{cn}\langle\kappa_{v,cn}(\theta_i)\rangle(Z_2 - z)}{\cos\theta_i}\right)\langle|S_{vv}(\theta_i)|^2\rangle dz \qquad (2.7b)$$

$$\mathrm{HV}_{cn,ss} = \frac{4\pi n_{cn}}{k^2} \int_{z=Z_1}^{z=Z_2} \exp\left(-\frac{2n_{cn}\langle\kappa_{h,cn}(\theta_i) + \kappa_{v,cn}(\theta_i)\rangle(Z_2 - z)}{\cos\theta_i}\right)\langle|S_{hv}(\theta_i)|^2\rangle dz \qquad (2.7c)$$

where $Z_2 - Z_1$ is the total height of the upper vegetation layer. The dependencies on ϕ_i, θ_s, and ϕ_s are no longer shown because $\phi_i = 0$, $\phi_s = \phi_i + \pi = \pi$, and $\theta_s = \theta_i$. Recall that $\langle\kappa_{h,cn}(\theta_i)\rangle$ and $\langle\kappa_{v,cn}(\theta_i)\rangle$ have units of area and are per-cylinder extinction cross-sections of the canopy layer for the h-polarized and v-polarized electric fields, respectively. Also note that n_{cn} and $(Z_2 - Z_1)$ always occur together in the combination $n_{cn}(Z_2 - Z_1)$ for the expressions for the normalized radar backscattering cross-section values, i.e. in this

calculation, the relevant parameter is instead the average number of cylinders per unit ground area, not n_{cn} and $(Z_2 - Z_1)$ separately.

Thus far, the validity of the single-scattering approximation has been implicitly relied upon. $\mathrm{HH}_{cn,ss}, \mathrm{VV}_{cn,ss}$ and $\mathrm{HV}_{cn,ss}$ shall be referred to as the "single-scattering solutions" for the backscatter from the canopy layer. When the vegetation gets thick, multiple scattering paths may become important. We attempt to take this into account by applying a polarization-dependent multiple scattering correction factor, denoted $\mathcal{F}$, such that our full solution for the backscatter from the canopy layer that includes multiple scattering effects is

$$\mathrm{HH}_{cn} = \mathcal{F}_{\mathrm{HH}}(\tau_{cn}(\theta_i), \theta_i, \varepsilon_v)\mathrm{HH}_{cn,ss} \tag{2.8a}$$

$$\mathrm{VV}_{cn} = \mathcal{F}_{\mathrm{VV}}(\tau_{cn}(\theta_i), \theta_i, \varepsilon_v)\mathrm{VV}_{cn,ss} \tag{2.8b}$$

$$\mathrm{HV}_{cn} = \mathcal{F}_{\mathrm{HV}}(\tau_{cn}(\theta_i), \theta_i, \varepsilon_v)\mathrm{HV}_{cn,ss}\,. \tag{2.8c}$$

The multiple scattering correction factor $\mathcal{F}$ depends on the incident angle θ_i, the cylinder relative permittivity ε_v, and the simplified average optical thickness of the canopy layer, $\tau_{cn}(\theta_i)$, where

$$\tau_{cn}(\theta_i) = \frac{n_{cn}\big(\langle \kappa_{h,cn}(\theta_i)\rangle + \langle \kappa_{v,cn}(\theta_i)\rangle\big)(Z_2 - Z_1)}{\cos\theta_i}\,. \tag{2.9}$$

For compactness, τ_{cn} instead of $\tau_{cn}(\theta_i)$ may be written subsequently, with implied dependence on the incidence angle θ_i.

To estimate $\mathcal{F}_{\mathrm{HH}}$ and $\mathcal{F}_{\mathrm{VV}}$, the method of radiative transfer [40] was used to calculate the backscatter without making the single-scattering assumption. The radiative transfer calculation was performed for several values of τ_{cn}, θ_i, and ε_v, with a uniformly random cylinder orientation distribution for (θ_c, ϕ_c), and a specific distribution over cylinder radii r and lengths L; details of this radiative transfer calculation are provided in Appendix B. A further correction is made to account for coherent backscatter enhancement not modelled by the radiative transfer equations. Interpolation is used to find $\mathcal{F}$ for intermediate values of $\tau_{vol}, \theta_i, \varepsilon_v$. $\mathcal{F}_{\mathrm{HV}}$ is estimated using Monte-Carlo estimates of the ratio between cross-

polarization and co-polarization returns for double scattering from the same cylinder distribution. More details are provided in Appendix B. Note that we only compute these multiple-scattering correction factors once for a fixed cylinder distribution, and apply these correction factors even for other cylinder distributions. The reason for this is that the multiple scattering computations are far more demanding than the single-scattering calculation. The multiple scattering correction factor was checked at a few values of $\tau_{cn}, \theta_i, \varepsilon_v$ to be similar to if a cosine-squared cylinder orientation distribution was used instead of a uniformly random orientation distribution, partially justifying this simplification.

The multiple scattering correction factor $\mathcal{F}(\tau_{cn}, \theta_i, \varepsilon_v)$ is plotted below in Figure 2.2 to Figure 2.4 as a function of the optical thickness $\tau_{cn}(\theta_i)$ for several computed values of θ_i and ε_v. (These values of θ_i correspond to the beam incidence angles of the Aquarius radar, to which the model will be applied in Chapter 3.) Note that this multiple-scattering being considered is within the branches of the canopy layer only. Multiple scattering pathways that simultaneously involve scattering off the ground and also off the branches or trunk more than once, have been neglected. This assumption is reasonable because when the optical thickness is large, such pathways incur a large extinction, whereas when the optical thickness is small, multiple scattering is weak. A less valid assumption is the neglect of multiple scattering pathways that simultaneously involve scattering off the canopy and the trunk have also been neglected; this assumption is made purely for simplicity. Finally, as a reminder, double-reflections involving the branches and the ground are being considered as a separate terms in equation (2.1).

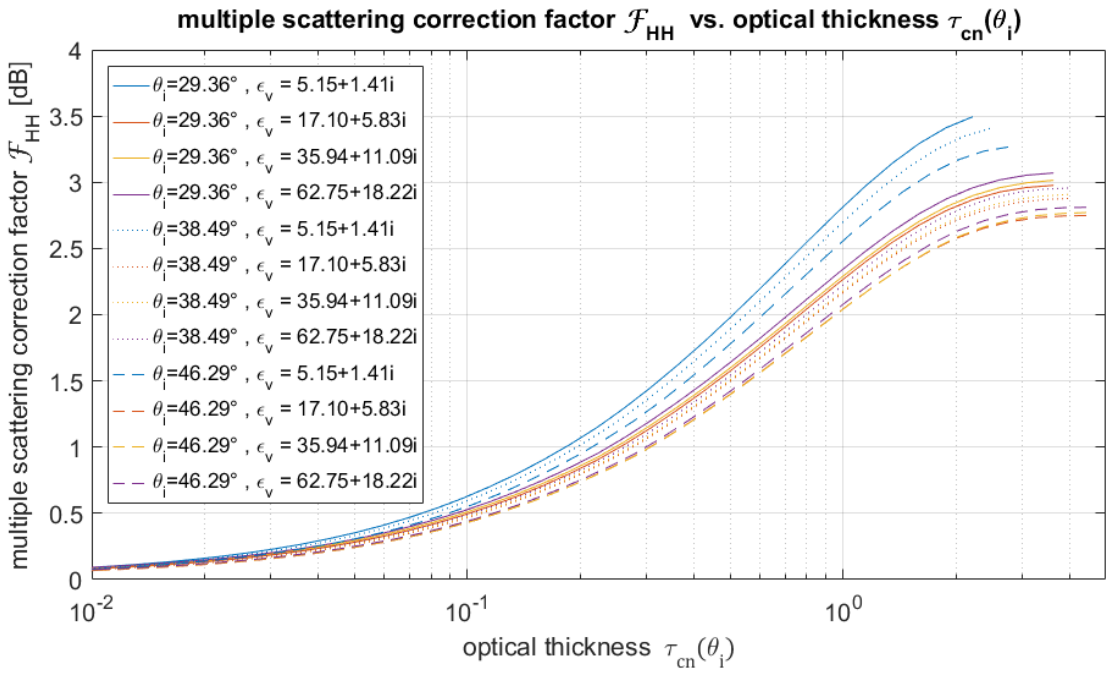

Figure 2.2. Estimated multiple-scattering correction factor $\mathcal{F}_{\mathrm{HH}}(\tau_{cn}(\theta_i), \theta_i, \varepsilon_v)$ as a function of optical thickness $\tau_{cn}(\theta_i)$, for several values of θ_i and ε_v .

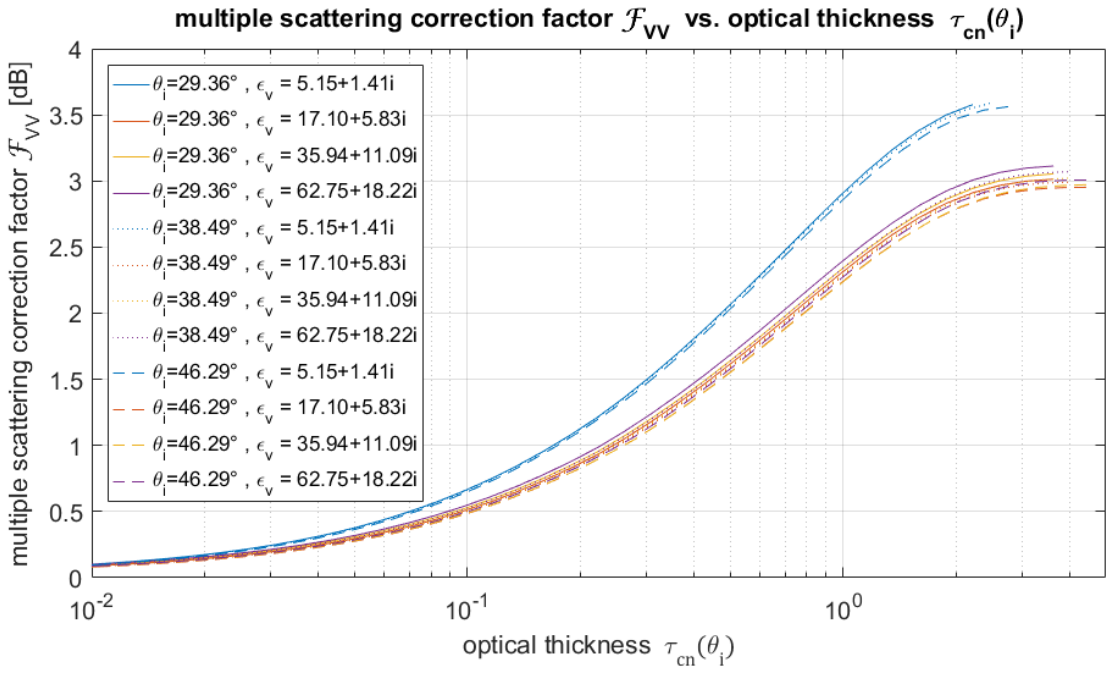

Figure 2.3. Estimated multiple-scattering correction factor $\mathcal{F}_{\mathrm{VV}}(\tau_{cn}(\theta_i), \theta_i, \varepsilon_v)$ as a function of optical thickness $\tau_{cn}(\theta_i)$, for several values of θ_i and ε_v .

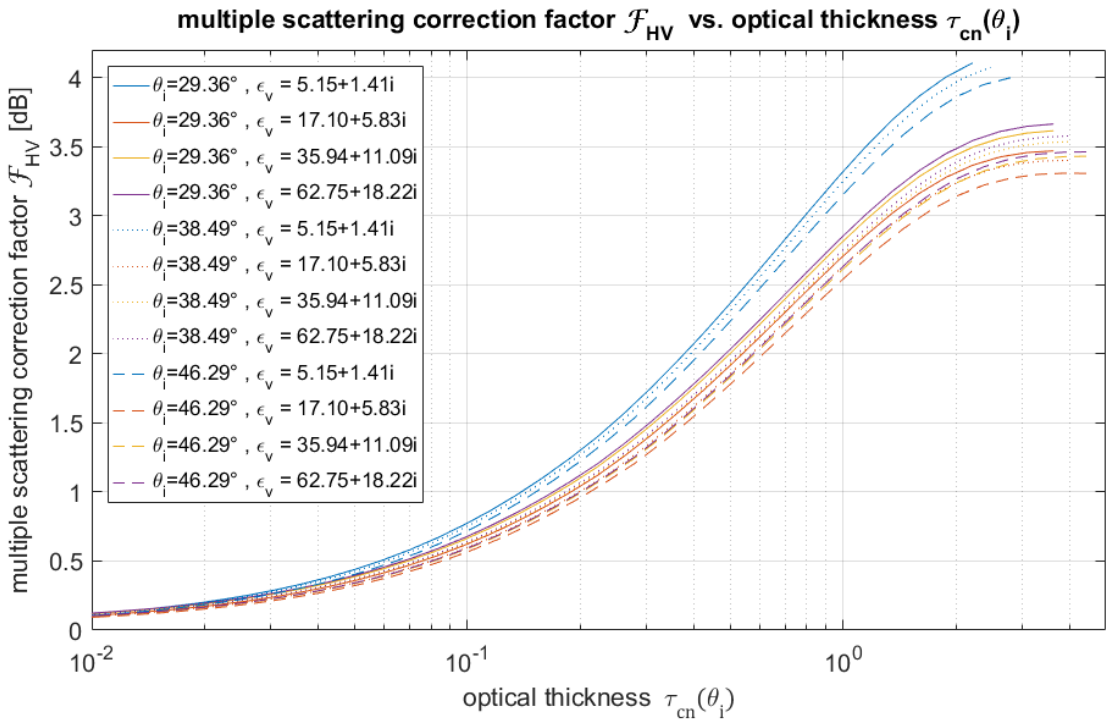

Figure 2.4. Estimated multiple-scattering correction factor $\mathcal{F}_{\mathrm{HV}}(\tau_{cn}(\theta_i), \theta_i, \varepsilon_v)$ as a function of optical thickness $\tau_{cn}(\theta_i)$, for several values of θ_i and ε_v .

2.3 Direct backscatter from the ground

This section describes the computation of the second term contributing to the overall radar backscatter, $\langle \sigma^0_{pol} \rangle_{gnd,direct}$ for $pol = hh, vv, hv$. For direct backscatter from the ground, we use the simplified IEM model by Fung and Chen [41]. The ground is modelled as a flat, horizontal, slightly rough surface with dielectric constant ε_g. The roughness is assumed to be a Gaussian process with RMS height h and exponential correlation with correlation length ζ. θ_i is the radar incidence angle and $k = 2\pi/\lambda$ is the wavenumber as usual. The normalized backscatter radar cross-section is then given in terms of these parameters by [41]:

$$\mathrm{HH}_{gnd,direct} = \frac{k^2}{2}\exp[-2k^2h^2\cos^2\theta_i]\sum_{n=1}^{\infty}|I^n_{hh}|^2\frac{W^{(n)}(2k\sin\theta_i, 0)}{n!} \qquad (2.10a)$$

$$\mathrm{VV}_{gnd,direct} = \frac{k^2}{2}\exp[-2k^2h^2\cos^2\theta_i]\sum_{n=1}^{\infty}|I^n_{vv}|^2\frac{W^{(n)}(2k\sin\theta_i, 0)}{n!} \qquad (2.10b)$$

where

$$I_{hh}^{n} = (2^{n}\exp[-k^{2}h^{2}\cos^{2}\theta_{i}]\, f_{hh} + \mathcal{G}_{hh})(kh\cos\theta_{i})^{n} \qquad (2.11a)$$

$$I_{vv}^{n} = (2^{n}\exp[-k^{2}h^{2}\cos^{2}\theta_{i}]\, f_{vv} + \mathcal{G}_{vv})(kh\cos\theta_{i})^{n} \qquad (2.11b)$$

For an isotropic correlation function, the dependence is only on the magnitude of the distance between two points. The surface spectrum (by the Wiener-Khinchin theorem) of an exponential correlation raised to the n-th power is, in polar coordinates,

$$W^{(n)}(2k\sin\theta_{i}, 0) = \frac{\zeta^{2}}{n^{2}}\left(1 + \left(\frac{k\zeta}{n}\right)^{2}\right)^{-\frac{3}{2}} \qquad (2.12)$$

where ζ is the correlation length. $\mathcal{G}_{hh}$, $\mathcal{G}_{vv}$, f_{hh}, f_{vv} are given through the following computations:

$$R_{v} = \frac{\varepsilon_{g}\cos\theta_{i} - \sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}}{\varepsilon_{g}\cos\theta_{i} + \sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}}, \qquad R_{h} = \frac{\cos\theta_{i} - \sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}}{\cos\theta_{i} + \sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}}$$

$$T_{v} = 1 + R_{v}, \qquad T_{h} = 1 + R_{h}$$

$$T_{vm} = 1 - R_{v}, \qquad T_{hm} = 1 - R_{h}$$

$$\mathcal{G}_{vv} = \left(\sin\theta_{i}\tan\theta_{i} - \frac{1}{\varepsilon_{g}}\sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}\right)T_{v}^{2} - 2\sin^{2}\theta_{i}\left(\sec\theta_{i} + \frac{1}{\sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}}\right)T_{v}T_{vm}$$
$$+ \left(\sin\theta_{i}\tan\theta_{i} + \varepsilon_{g}(1 + \sin^{2}\theta_{i})/\sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}\right)T_{vm}^{2}$$

$$\mathcal{G}_{hh} = -\left(\sin\theta_{i}\tan\theta_{i} - \sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}\right)T_{h}^{2} + 2\sin^{2}\theta_{i}\left(\sec\theta_{i} + \frac{1}{\sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}}\right)T_{h}T_{hm}$$
$$- \left(\sin\theta_{i}\tan\theta_{i} + (1 + \sin^{2}\theta_{i})/\sqrt{\varepsilon_{g} - \sin^{2}\theta_{i}}\right)T_{hm}^{2}$$

$$R_{v0} = \frac{\sqrt{\varepsilon_{g}} - 1}{\sqrt{\varepsilon_{g}} + 1}, \qquad R_{h0} = -R_{v0}$$

$$F_t = \frac{8R_{v0}^2 \sin^2\theta_i \left(\cos\theta_i + \sqrt{\varepsilon_g - \sin^2\theta_i}\right)}{\cos\theta_i \sqrt{\varepsilon_g - \sin^2\theta_i}}$$

$$\frac{S_t}{S_{t0}} = \frac{\sum_{n=1}^{\infty} \frac{(kh\cos\theta_i)^{2n}}{n!} W^{(n)}(2k\sin\theta_i, 0) \left| F_t + \frac{8R_{v0}}{\cos\theta_i} \right|^2}{\sum_{n=1}^{\infty} \frac{(kh\cos\theta_i)^{2n}}{n!} W^{(n)}(2k\sin\theta_i, 0) \left| F_t + \frac{2^{n+2}R_{v0}}{\exp\left[(kh\cos\theta_i)^2\right]\cos\theta_i} \right|^2}$$

$$R_{tv} = \frac{S_t}{S_{t0}} R_v + \left(1 - \frac{S_t}{S_{t0}}\right) R_{v0}\,, \qquad R_{th} = \frac{S_t}{S_{t0}} R_h + \left(1 - \frac{S_t}{S_{t0}}\right) R_{h0}$$

$$f_{vv} = \frac{2R_{tv}}{\cos\theta_i}\,, \qquad f_{hh} = -\frac{2R_{th}}{\cos\theta_i}\,.$$

In our numerical computation, we use only the first twenty terms of the infinite series, as a trade-off between computational accuracy and speed.

Fung and Chen [41] also provide expressions for the cross-polarized backscatter. This cross-polarized backscatter term from the ground is neglected here because it is small compared to the cross-polarized backscatter from the vegetation layer, i.e. let

$$\mathrm{HV}_{gnd,direct} = 0. \tag{2.13}$$

Strictly speaking, this approximation is only valid when the ground surface is reasonably flat, but because forested areas have strong cross-polarized backscatter from the vegetation, this approximation is good except in very extreme cases e.g. mountainous areas. When applying our model in subsequent chapters, we shall exclude areas with high terrain slope.

To simplify and reduce the total number of parameters in the overall model, we make the assumption that the correlation length and the RMS height are related [22, 42] via $\zeta = 10h$. This reduces the number of parameters for the ground to two, namely, the roughness RMS height h and the relative permittivity ε_g. The ground dielectric relative permittivity, ε_g, is closely related to the soil moisture; fluctuations in ε_g are almost synonymous with fluctuations in soil moisture, because the relative permittivity of liquid water at L-band is much higher than the relative permittivity of most rocks. A model by Mironov et al. [43] for

the relationship between soil moisture and ε_g is provided in Appendix C. We shall not need to explicitly use this relationship until Chapter 5, where it will be discussed further; until then, it suffices for us to work in terms of ε_g.

2.4 Double-reflections off the ground and crown canopy layer

This section describes the computation of the third term contributing to the overall radar backscatter, $\langle \sigma^0_{pol} \rangle_{vn-gnd,db}$ for $= hh, vv, hv$. Figure 2.5 is a schematic of the double-reflection off the ground and a cylinder in the upper crown canopy layer. The cylinder is at a height z above the ground, $Z_1 \leq z \leq Z_2$. Z_1 is the total height of the lower trunk layer, and $Z_2 - Z_1$ is the total height of the upper crown canopy layer. Note that there are two exactly opposite pathways that add coherently. Let path 3a be the path that scatters off the cylinder first, and then the ground, before returning to the radar antenna. Let path 3b be the opposite path that scatters off the ground first, and then the cylinder, before returning to the radar antenna. To find the double-bounce backscatter, the bistatic scattering off the cylinder, bistatic scattering off the ground, extinction through the upper crown canopy layer, and extinction through the lower trunk layer are required.

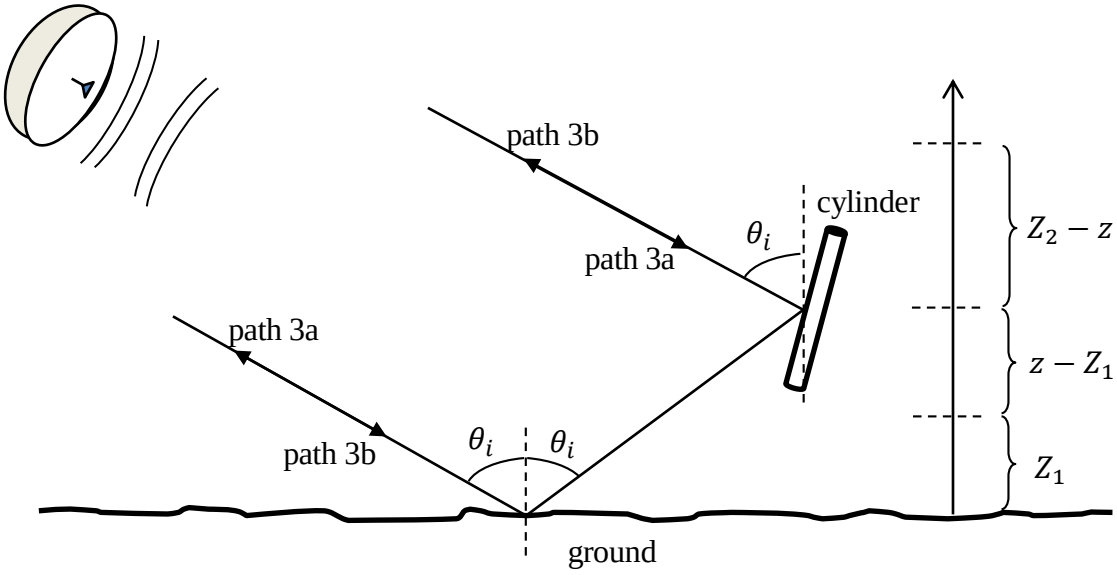

Figure 2.5. Double reflections off the ground and a cylinder in the upper vegetation volume layer are a coherent sum of two opposite pathways, labelled 3a and 3b.

Let

$$\left[\mathbf{S}_{cyl,3a}\right] = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{cyl,3a} \tag{2.14}$$

be the 2x2 dimensionless bistatic scattering matrix off the cylinder for incidence from direction (θ_i, ϕ_i) and scattering into direction $(\theta_s = \pi - \theta_i, \phi_s = \phi_i + \pi)$, and $[\mathbf{G}]$ be the 2x2 dimensionless bistatic scattering matrix off the ground for incidence from direction (θ_i, ϕ_i) and scattering into direction $(\theta_s = \theta_i, \phi_s = \phi_i)$, both written in their local BSA coordinates. As before, $\left[\mathbf{S}_{cyl,3a}\right]$ is provided by the expressions in Appendix A. For $[\mathbf{G}]$, the Fresnel reflection coefficients are used, with a correction for surface roughness found from a physical optics approximation (also often called the Kirchhoff approximation) [44, 14]

$$[\mathbf{G}] = \begin{bmatrix} G_{hh} & 0 \\ 0 & G_{vv} \end{bmatrix} = \exp(-2k^2h^2\cos^2\theta_i) \begin{bmatrix} R_h & 0 \\ 0 & R_v \end{bmatrix} \tag{2.15}$$

where R_h and R_v are the Fresnel reflection coefficients

$$R_h = \frac{\cos\theta_i - \sqrt{\varepsilon_g - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\varepsilon_g - \sin^2\theta_i}} \tag{2.16a}$$

$$R_v = \frac{\varepsilon_g\cos\theta_i - \sqrt{\varepsilon_g - \sin^2\theta_i}}{\varepsilon_g\cos\theta_i + \sqrt{\varepsilon_g - \sin^2\theta_i}} \tag{2.16b}$$

and $\exp(-2k^2h^2\cos^2\theta_i)$ is the roughness correction factor.

The extinction cross-sections (normalized to one cylinder) through the upper crown canopy layer were $\langle\kappa_{h,cn}\rangle$ and $\langle\kappa_{v,cn}\rangle$ for h-polarized and v-polarized electric field respectively. Likewise, let $\langle\kappa_{h,trk}\rangle$ and $\langle\kappa_{v,trk}\rangle$ be the extinction cross-sections (normalized to one cylinder) through the lower trunk layer for h-polarized and v-polarized electric field, with units of area, and n_{trk} be the average number density of cylinders in the trunk layer, with units of inverse volume.

The total one-way extinction of the electric field through the lower trunk layer is then, for the two polarizations,

$$\alpha_{h,trk}(Z_1) = \exp\left[-\frac{n_{trk}\langle\kappa_{h,trk}\rangle Z_1}{\cos\theta_i}\right] \tag{2.17a}$$

$$\alpha_{v,trk}(Z_1) = \exp\left[-\frac{n_{trk}\langle\kappa_{v,trk}\rangle Z_1}{\cos\theta_i}\right] \tag{2.17b}$$

For convenience in multiplication with 2x2 scattering matrices, let

$$[\boldsymbol{\alpha}_{trk}(Z_1)] = \begin{bmatrix} \alpha_{h,trk}(Z_1) & 0 \\ 0 & \alpha_{v,trk}(Z_1) \end{bmatrix}. \tag{2.18}$$

Likewise, for the upper crown canopy layer, the one-way extinction of the electric field through the part of the upper crown canopy layer above the cylinder, written as a 2x2 matrix, is

$$[\boldsymbol{\alpha}_{cn}(Z_2 - z)] = \begin{bmatrix} \exp\left[-\frac{n_{cn}\langle\kappa_{h,cn}\rangle(Z_2 - z)}{\cos\theta_i}\right] & 0 \\ 0 & \exp\left[-\frac{n_{cn}\langle\kappa_{v,cn}\rangle(Z_2 - z)}{\cos\theta_i}\right] \end{bmatrix}. \tag{2.19}$$

Similarly, the one-way extinction of the electric field through the part of the upper crown canopy layer below the cylinder is $[\boldsymbol{\alpha}_{cn}(z - Z_1)]$, and the total one-way extinction of the electric field through the upper crown canopy layer is $[\boldsymbol{\alpha}_{cn}(Z_2 - Z_1)]$.

We can now write the overall 2x2 scattering matrix for path 3a as

$$\begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{3a} = [\boldsymbol{\alpha}_{cn}(Z_2 - Z_1)][\boldsymbol{\alpha}_{trk}(Z_1)][\mathbf{G}]\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\times [\boldsymbol{\alpha}_{trk}(Z_1)][\boldsymbol{\alpha}_{cn}(z - Z_1)]\left[\mathbf{S}_{cyl,3a}\right][\boldsymbol{\alpha}_{cn}(Z_2 - z)] \tag{2.20}$$

where the factor of $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ accounts for the direction of the axes in the BSA convention.

Similarly, the overall 2x2 scattering matrix for path 3b is

$$\begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{3b} = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{3a}^{T}$$

$$= [\boldsymbol{\alpha}_{cn}(Z_2 - z)]\left[\mathbf{S}_{cyl,3a}\right]^{T}[\boldsymbol{\alpha}_{cn}(z - Z_1)][\boldsymbol{\alpha}_{trk}(Z_1)]\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}[\mathbf{G}]$$

$$\times [\boldsymbol{\alpha}_{trk}(Z_1)][\boldsymbol{\alpha}_{cn}(Z_2 - Z_1)]\,. \qquad (2.21)$$

Paths 3a and 3b combine coherently to give the symmetric 2x2 scattering matrix

$$\begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{cyl,db} = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{3a} + \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{3b}$$

$$= [\boldsymbol{\alpha}_{cn}(Z_2 - z)]\left(\left[\mathbf{S}_{cyl,3a}\right]^{T}\begin{bmatrix} Q_{hh} & 0 \\ 0 & Q_{vv} \end{bmatrix} + \begin{bmatrix} Q_{hh} & 0 \\ 0 & Q_{vv} \end{bmatrix}^{T}\left[\mathbf{S}_{cyl,3a}\right]\right)[\boldsymbol{\alpha}_{cn}(Z_2 - z)] \qquad (2.22)$$

where

$$\begin{bmatrix} Q_{hh} & 0 \\ 0 & Q_{vv} \end{bmatrix} = [\boldsymbol{\alpha}_{cn}(z - Z_1)][\boldsymbol{\alpha}_{trk}(Z_1)]\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}[\mathbf{G}][\boldsymbol{\alpha}_{trk}(Z_1)][\boldsymbol{\alpha}_{cn}(z - Z_1)]\,. \qquad (2.23)$$

Rearranging,

$$\begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix}_{cyl,db} = [\mathfrak{A}][\mathbb{2}][\mathbf{Q}]\begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix}_{cyl,3a} \qquad (2.24)$$

where

$$[\mathfrak{A}]$$

$$= \begin{bmatrix} \exp\left[-2n_{cn}\langle\kappa_{h,cn}\rangle(Z_2 - z)/\cos\theta_i\right] & 0 & 0 \\ 0 & \exp\left[-n_{cn}\langle\kappa_{h,cn} + \kappa_{v,cn}\rangle(Z_2 - z)/\cos\theta_i\right] & 0 \\ 0 & 0 & \exp\left[-2n_{cn}\langle\kappa_{h,cn}\rangle(Z_2 - z)/\cos\theta_i\right] \end{bmatrix}$$

$$[\mathbb{2}] = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$[\mathbf{Q}] = \begin{bmatrix} Q_{hh} & 0 & 0 & 0 \\ 0 & Q_{hh} & Q_{vv} & 0 \\ 0 & 0 & 0 & Q_{vv} \end{bmatrix}.$$

Next the 3x3 scatterer covariance matrix for the single cylinder can be constructed:

$$\begin{bmatrix} S_{hh}S_{hh}^* & \sqrt{2}\,S_{hh}S_{hv}^* & S_{hh}S_{vv}^* \\ \sqrt{2}\,S_{hv}S_{hh}^* & 2\,S_{hv}S_{hv}^* & \sqrt{2}\,S_{hv}S_{vv}^* \\ S_{vv}S_{hh}^* & \sqrt{2}\,S_{vv}S_{hv}^* & S_{vv}S_{vv}^* \end{bmatrix}_{cyl,db}$$

$$= [\mathfrak{U}][\mathbf{2}][\mathbf{Q}] \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix}_{cyl,3a} \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix}_{cyl,3a}^{\dagger} [\mathbf{Q}]^{\dagger}[\mathbf{2}][\mathfrak{U}]^{\dagger}\,. \qquad (2.25)$$

By averaging over the cylinder distribution and integrating over the layer, the normalized radar cross-sections can then be found using the diagonal entries of the averaged 3x3 covariance matrix

$$\begin{bmatrix} \langle\sigma_{hh}^0\rangle & * & * \\ * & 2\langle\sigma_{hv}^0\rangle & * \\ * & * & \langle\sigma_{vv}^0\rangle \end{bmatrix}_{cn-gnd,db} = \frac{4\pi n_{cn}}{k^2} \times$$

$$\int_{z=Z_1}^{z=Z_2} [\mathfrak{U}][\mathbf{2}][\mathbf{Q}] \left\langle \begin{bmatrix} S_{hh}S_{hh}^* & S_{hh}S_{hv}^* & S_{hh}S_{vh}^* & S_{hh}S_{vv}^* \\ S_{hv}S_{hh}^* & S_{hv}S_{hv}^* & S_{hv}S_{vh}^* & S_{hv}S_{vv}^* \\ S_{vh}S_{hh}^* & S_{vh}S_{hv}^* & S_{vh}S_{vh}^* & S_{vh}S_{vv}^* \\ S_{vv}S_{hh}^* & S_{vv}S_{hv}^* & S_{vv}S_{vh}^* & S_{vv}S_{vv}^* \end{bmatrix}_{cyl,3a} \right\rangle [\mathbf{Q}]^{\dagger}[\mathbf{2}][\mathfrak{U}]^{\dagger} dz \qquad (2.26)$$

where the off-diagonal entries on the left-hand side have simply not been displayed, and the angular brackets on the right-hand side denote averaging cylinders over $p(\theta_c, \phi_c, r, L)$ as described earlier in the Section 2.2 on backscatter from the upper crown canopy layer only.

2.5 Trunk-ground double reflections

This section describes the computation of the fourth term contributing to the overall radar backscatter, $\langle\sigma_{pol}^0\rangle_{trk-gnd,db}$ for $= hh, vv, hv$. Like the upper crown canopy layer, tree trunks are modelled as cylinders at a height z above the ground, $0 \le z \le Z_1$, but are handled separately because trunks tend to have a vertical orientation. Because of the predominant vertical orientation, some simplifications can be made. The first simplification was the neglect of direct backscatter from the tree trunks, leaving trunks to contribute only via this trunk-ground double reflection term. Another simplification is the neglect of $S_{hv,trk,3a}$ and $S_{vh,trk,3a}$ for predominantly vertically oriented cylinders, where

$$[\mathbf{S}_{trk,3a}] = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{trk,3a} \tag{2.27}$$

is the 2x2 dimensionless bistatic scattering matrix off a trunk for incidence from direction (θ_i, ϕ_i) and scattering into direction $(\theta_s = \pi - \theta_i, \phi_s = \phi_i + \pi)$, completely analogous to equation (2.14).

Consequently,

$$\mathrm{HV}_{trk-gnd,db} = 0 \tag{2.28a}$$

$$\mathrm{HH}_{trk-gnd,db} = \frac{4\pi n_{trk} Z_1}{k^2} 4 \exp\left[-4 \frac{n_{cn}\langle\kappa_{h,cn}\rangle(Z_2 - Z_1) + n_{trk}\langle\kappa_{h,trk}\rangle Z_1}{\cos\theta_i}\right] \times |G_{hh}|^2 \langle |S_{hh}|^2_{trk,3a}\rangle \tag{2.28b}$$

$$\mathrm{VV}_{trk-gnd,db} = \frac{4\pi n_{trk} Z_1}{k^2} 4 \exp\left[-4 \frac{n_{cn}\langle\kappa_{v,cn}\rangle(Z_2 - Z_1) + n_{trk}\langle\kappa_{v,trk}\rangle Z_1}{\cos\theta_i}\right] \times |G_{vv}|^2 \langle |S_{vv}|^2_{trk,3a}\rangle \,. \tag{2.28c}$$

Typically, $\langle |S_{hh}|^2_{trk,3a}\rangle > \langle |S_{vv}|^2_{trk,3a}\rangle$ and $\langle\kappa_{v,trk}\rangle > \langle\kappa_{h,trk}\rangle$ because of the vertical orientation of the trunks, and $|G_{hh}|^2 > |G_{vv}|^2$ because of the Fresnel reflection coefficients, so $\mathrm{HH}_{trk-gnd,db}$ dominates $\mathrm{VV}_{trk-gnd,db}$.

The trunk cylinders are modelled with an orientation distribution that is uniform in ϕ_c and Gaussian distributed about the vertical direction [14], with a small RMS tilt σ_c

$$p(\theta_c, \phi_c) = \frac{\exp\left[-\frac{1}{2}\left(\frac{\theta_c}{\sigma_c}\right)^2\right]}{2\pi \int_{\theta_c=0}^{\theta_c=\pi} \sin\theta_c \exp\left[-\frac{1}{2}\left(\frac{\theta_c}{\sigma_c}\right)^2\right] d\theta_c} \,. \tag{2.29}$$

For computation, the integrals required for averaging over the trunk distribution are approximated numerically as sums. The integral over (θ_c, ϕ_c) is performed first. In view of the narrow distribution about the vertical, an angular grid is set up only up to $\theta_c = 4\sigma_c$ from the vertical. For extinction calculations, the angular spacing was set to $\sigma_c/5$ in θ_c, and 10°

in ϕ_c. For bistatic scattering calculations, an angular spacing of $0.35\lambda/\pi L$ was used in both θ_c and ϕ_c. This angular spacing is chosen in view of the width of the sinc function in equation (A.1) in Appendix A. If $0.35\lambda/\pi L$ is greater than $\sigma_c/5$, the angular spacing is set at $\sigma_c/5$ instead. Because the bistatic scattering is dominated by cylinders oriented approximately perpendicular to the direction $\mathbf{k}_i + \mathbf{k}_s$ (where $\mathbf{k}_i, \mathbf{k}_s$ are in the BSA convention), for accelerated computation, only cylinders that are within an angle of $2\lambda/L$ from an orientation perpendicular to $\mathbf{k}_i + \mathbf{k}_s$ are included. The specific choice of distribution over trunk cylinder radii and lengths (r, L) will be elaborated upon in Section 2.8.

2.6 Cylinder relative permittivity $\boldsymbol{\varepsilon_v}$

For the cylinder relative permittivity, the model of Ulaby and El-Rayes [45] is used. Following the model, the relative pemittivity of vegetation material is related to the volumetric moisture content of vegetation, M_v, and the microwave frequency f in GHz, via

$$\varepsilon_v = v_{fw}\varepsilon_f + v_b\varepsilon_b + \varepsilon_r. \tag{2.30}$$

In the first term, v_{fw} is the volume fraction of free water and ε_f is its relative permittivity:

$$v_{fw} = M_v(0.82M_v + 0.166) \tag{2.31}$$

$$\varepsilon_f = 4.9 + \frac{75}{1 - \frac{if}{18}} + 18\frac{\gamma_{sal}}{f}i \tag{2.32}$$

$$\gamma_{sal} = 0.16S - 0.0013S^2 \tag{2.33}$$

where γ_{sal} is the ionic conductivity, and S the water salinity in parts per thousand on a weight basis.

In the second term, v_b is the volume fraction of the bulk vegetation-bound water mixture and ε_b is its dielectric constant:

$$v_b = \frac{31.4M_v^2}{1 + 59.5M_v^2} \tag{2.34}$$

$$\varepsilon_b = 2.9 + \frac{55}{1 - i\sqrt{\frac{if}{0.18}}}\,. \qquad (2.35)$$

The third term ε_r is a nondispersive residual component

$$\varepsilon_r = 1.7 + 3.2M_v + 6.5M_v^2. \qquad (2.36)$$

In all these expressions, $i = \sqrt{-1}$ is the imaginary number and the opposite sign convention was taken from Ulaby and El-Rayes to keep consistent with the expressions in Appendix A (i.e. electric field oscillates as $\exp(-2\pi i f t)$). Unless otherwise stated, we typically choose M_v to be 0.5 and set S to 8.5. This gives a value of $\varepsilon_v = 35.94 + 11.09\mathrm{i}$, which is in reasonable agreement with measurements taken by Chauhan and Lang [15] at a walnut orchard and Durden et al. [46] at a coniferous forest near Mount Shasta. This value of ε_v is taken as the default value for subsequent modelling, unless variations in ε_v are explicitly considered in context.

2.7 Distribution of cylinder radius and length in the crown canopy layer

This section describes and explains the choice of cylinder size distribution (over cylinder radius r and cylinder length L) for the upper crown canopy layer. A cylinder size distribution constrained to have as few free parameters as possible is desired. As mentioned in the overview, the reason is that in application to Aquarius and SMAP data, there are only 3 output radar observables to compare against or perform model inversion with: the normalized radar backscattering cross-sections HH, VV, and HV. In principle, the forward model can also compute the full covariance matrix containing the relative phases between the polarizations, by computing integrals containing $S_{hh}S_{vv}^*, S_{vv}S_{hv}^*$, etc. in equations analogous to (2.7a-c) for the cylinders.

Before presenting the specific choice of cylinder size distribution, a preliminary analysis of the dependence of the forward model on the radius, r, is performed. This will shed some useful insights with regards to how the cylinder size distribution affects the radar backscatter. This preliminary analysis is performed for an upper crown canopy layer of uniformly randomly oriented cylinders, i.e. in the cosine-squared distribution for (θ_c, ϕ_c), the exponent

parameter was set to $m = 0$, no multiple scattering correction, no lower trunk layer, and neglecting the ground and any interactions with it, i.e. only the single-scattering solutions $\mathrm{HH}_{cn,ss}$, $\mathrm{HV}_{cn,ss}, \mathrm{VV}_{cn,ss}$ from equations (2.7a-c). All the cylinders in the layer have identical sizes. The length was pegged to radius via $L = (r/1\mathrm{cm})^{2/3}$ m (the reason for this choice will be explained in Section 2.7), and the cylinder number density n_{cn} was chosen in such a way as to keep the total volume of cylinders per unit ground area $n_{cn}(Z_2 - Z_1)\pi r^2 L$ fixed at $10^{-3}\mathrm{m}^3/\mathrm{m}^2$. Note that for uniformly randomly oriented cylinders, by symmetry we expect $\langle \kappa_{h,cn} \rangle = \langle \kappa_{v,cn} \rangle$, $\mathrm{HH}_{cn,ss} = \mathrm{VV}_{cn,ss}$. Figure 2.6 to Figure 2.8 show the results of this preliminary analysis. The one-way extinction (for power or intensity), $\exp(-\tau_{cn})$, through the layer, and radar backscatter $\mathrm{HH}_{cn,ss}$, $\mathrm{HV}_{cn,ss}$, are plotted against cylinder radius r. The radar wavelength was set to 24cm, and the incidence angle 40° from the vertical.

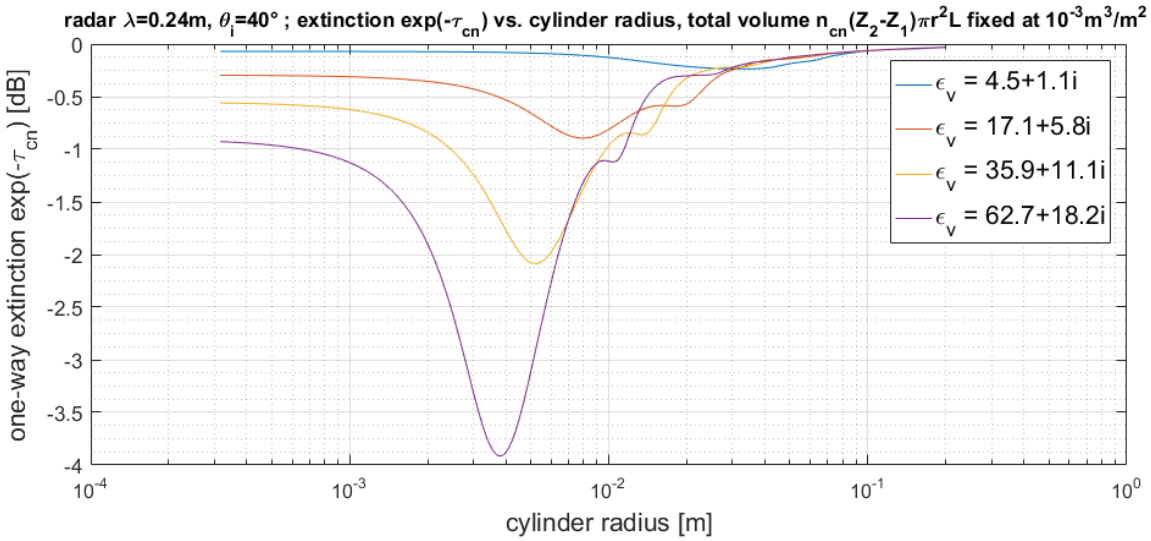

Figure 2.6. L-band (λ=24cm) one-way extinction at incidence angle of 40° for a layer of uniformly randomly oriented identical cylinders, as a function of cylinder radius r for several values of dielectric relative permittivity (blue curve: $\varepsilon_v = 4.5 + 1.1\mathrm{i}$; red curve: $\varepsilon_v = 17.1 + 5.8\mathrm{i}$; yellow curve $\varepsilon_v = 35.9 + 11.1\mathrm{i}$; $\varepsilon_v = 62.7 + 18.2\mathrm{i}$). Total volume of cylinders per unit ground area is fixed at $10^{-3}\mathrm{m}^3/\mathrm{m}^2$.

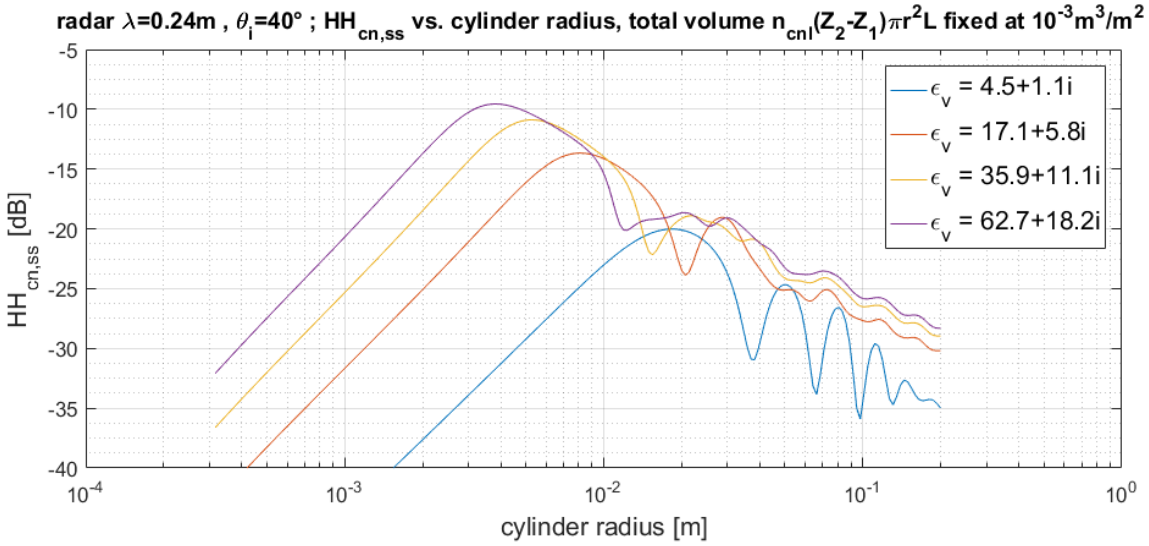

Figure 2.7. Normalized radar backscattering cross-section $\mathrm{HH}_{cn,ss}$ from equation (2.7a) at incidence angle of 40° for a layer of uniformly randomly oriented identical cylinders, as a function of cylinder radius r for several values of dielectric relative permittivity (blue curve: $\varepsilon_v = 4.5 + 1.1\mathrm{i}$; red curve: $\varepsilon_v = 17.1 + 5.8\mathrm{i}$; yellow curve $\varepsilon_v = 35.9 + 11.1\mathrm{i}$; $\varepsilon_v = 62.7 + 18.2\mathrm{i}$). Total volume of cylinders per unit ground area is fixed at $10^{-3}\mathrm{m}^3/\mathrm{m}^2$, and cylinder $L = (r/1\mathrm{cm})^{2/3}$ m.

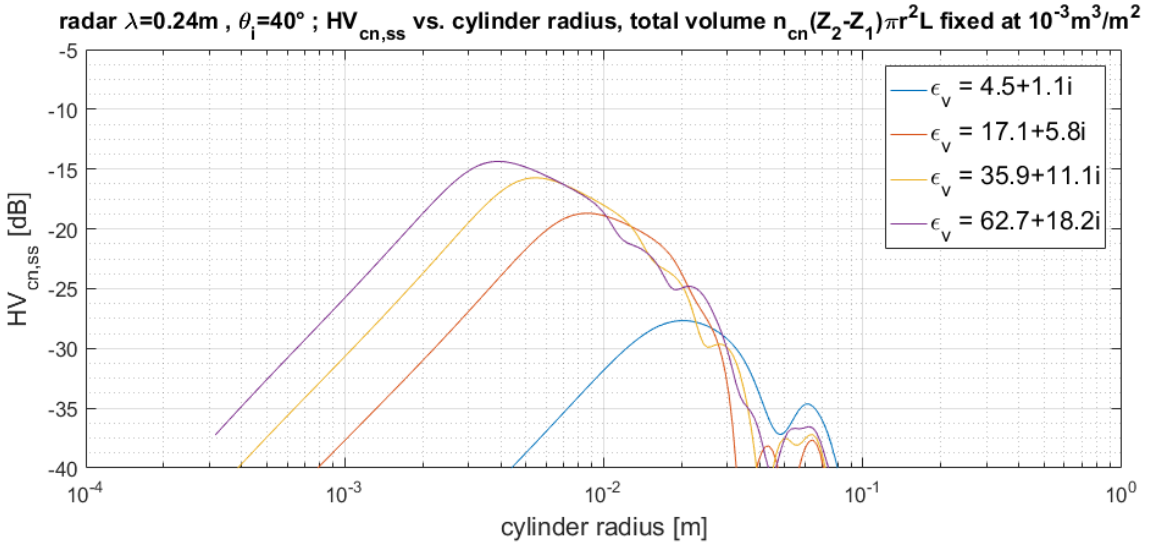

Figure 2.8. Normalized radar backscattering cross-section $\mathrm{HV}_{cn,ss}$ from equation (2.7c) at incidence angle of 40° for a layer of uniformly randomly oriented identical cylinders, as a function of cylinder radius r for several values of dielectric relative permittivity (blue curve: $\varepsilon_v = 4.5 + 1.1\mathrm{i}$; red curve: $\varepsilon_v = 17.1 + 5.8\mathrm{i}$; yellow curve $\varepsilon_v = 35.9 + 11.1\mathrm{i}$; $\varepsilon_v = 62.7 + 18.2\mathrm{i}$). Total volume of cylinders per unit ground area is fixed at $10^{-3}\mathrm{m}^3/\mathrm{m}^2$, and cylinder $L = (r/1\mathrm{cm})^{2/3}$ m.

Even though here the cylinder length L was fixed in relation to r and the results for different choices for L are not displayed, the extinction is actually independent of L, if the total volume is fixed. The reason is that cylinders with longer length have a proportionately longer extinction contribution, but the total number of cylinders is also proportionately fewer. The backscatter contribution from the canopy is also only weakly dependent on cylinder length, if the total volume is fixed. Given a fixed total volume $n_{cn}(Z_2 - Z_1)\pi r^2 L$, if L is increased, each cylinder has a larger backscatter scaling as the square of L, but the total number of cylinders n_{cn} is inversely proportional to L, and further there are also fewer cylinders close to perpendicular to the incident direction that contribute significantly to the backscatter (within an angle proportional to λ/L).

The most obvious trend in Figure 2.6 to Figure 2.8 is that the interaction of the radar, when considered per unit volume of cylinder, is strongest at the resonance size $\left(\approx 2\pi r\sqrt{\varepsilon_v}/\lambda\right)$. The other important point is that the vegetation dielectric constant affects the backscatter both through changing the resonance size of cylinders as well as the strength of the interaction. For $\varepsilon_v \approx 36 + 11\mathrm{i}$, the resonance size is about 0.5-1 cm radius, or 1-2 cm diameter. Very small branches (<1 mm radius) approach the Rayleigh regime and have negligible backscatter. Branches much larger than the resonance size also have less interaction with the radar per unit volume, but their total volume may be significant. This issue extends to tree trunks (whose total volume dominates the branches, but whose contribution to the radar backscatter is primarily only in the horizontal polarization through the trunk-ground double-bounce mechanism $\mathrm{HH}_{\mathrm{trk-gnd,db}}$) and is evidently a primary source of the well-known difficulty of estimating above-ground biomass for dense vegetation if only L-band normalized radar backscatter cross-sections are available. As such, for biomass retrieval, either radar normalized backscatter cross-sections at P-band due to the longer wavelength is preferred [47], or if a shorter wavelength is to be used, additional information from interferometry to measure tree heights [48] and/or phase information to isolate the double-bounce mechanism using the phase difference between horizontal and vertical polarization [12, 13] should be used.

From a modelling perspective, for vegetated areas, in particular forests and dense natural vegetation, the single most important parameter for the canopy volume is the total volume (per unit ground area) of resonance-sized branches. Ideally, full knowledge of the distribution of that volume as a function of cylinder radius is needed. Fractal tree models [49] and "computer-grown" trees based on architectural plant models have been used [50] for radar backscatter simulations, but these do not explicitly specify the distribution as a function of cylinder radius. Reviewing some of the radar modelling literature [9, 46, 51, 16, 15, 52], attempts have been made at measuring the distribution of branches, but a clear, simple functional form with widespread applicability for global-scale modelling has not been proposed. For that, we turn to a general model proposed by West et al. [53] in the ecology literature. While their theory is much more general, parts of which remain controversial, we only seek the distribution of branches as a function of radius, not the validity of their entire theory. Their model [53] views a plant as a "branching hierarchical network running from the trunk (level k=0) to the petioles (level k=K>0)", with the number of branches of a given size inversely proportional to the square of the branch radius

$$N_k \propto r_k^{-2} \tag{2.37}$$

reminiscent of an observation by Leonardo da Vinci [54]: cross-sectional area is preserved whenever a tree branches. This area-preserving branching condition was la also associated with the pipe model by Shinozaki et al. [55] [56]. The branch radii at each level are related by

$$\frac{r_{k+1}}{r_k} = \text{constant.} \tag{2.38}$$

Extending the discrete hierarchical levels to a continuous distribution, the discrete levels can be viewed as occupying evenly spaced bins of width $\Delta \ln r$ on a log-scale of branch radius:

$$N_k \propto A\, n_{cn} p(r)\, \Delta \ln r \propto r^{-2} \tag{2.39}$$

where A is the ground area, and $n_{vol} p(r)$ is the number of branches per unit radius per unit ground area. The distribution for the number of branches per unit radius per unit ground area is thus inverse cubed:

$$n_{cn}p(r) \propto r^{-3}. \quad (2.40)$$

What is the range of branch radii for which this distribution is valid? There must certainly be some minimum (since the smallest branches terminate as leaves) and maximum (since there is a maximum size to branches) branch radii beyond which these relationships fail. From Shinozaki et al. [55, 56], this minimum value occurs around 1-2mm, while the maximum value depends on the tree species, since different tree species have different maximum branch sizes. Fortunately, the approximate range of validity includes the range of resonance cylinder radii (at $\lambda = 0.24$ m and $\varepsilon_v \approx 36 + 11\text{i}$), from 1 mm to 3 cm. Our approach for forests is thus to model the branches only up to an arbitrary cutoff maximum radius of 3 cm and down to a minimum radius of 1 mm. As mentioned earlier, the Rayleigh regime smaller than 1 mm is not expected to contribute significantly to the radar backscatter. However, for the maximum cutoff, further simulations show that the forward model does have some sensitivity to the choice of cutoff in a way that also depends on other input parameters. Hence, failure to model branches larger than 3 cm may have some detrimental impact on the accuracy of the forward model, since the true maximum branch radius may be larger than that, especially in the tropical jungles. The benefit of this simplification is the avoidance of having to handle the maximum branch size as an additional species-dependent unknown variable, which would be tricky to implement on a global scale. Additionally, 3 cm does correspond approximately to the maximum branch radius from various field measurements in boreal forests [57, 46, 52].

With the radius distribution specified, the cylinder lengths remain. Mechanical considerations and botanical data of trees are consistent with a relationship between length and radius of the form given by McMahon and Kronauer [58] :

$$L_{tot}(r) \propto r^{\frac{2}{3}} \quad (2.41)$$

Where $L_{tot}(r)$ is the average total "path length" from the point a branch has radius r to a virtual twig with radius 0. For the sake of analysis, constants $L_{tot,b}$ and r_b are temporarily introduced such that

$$L_{tot}(r) = L_{tot,b}\left(\frac{r}{r_b}\right)^{\frac{2}{3}}. \quad (2.42)$$

Then the length of a branch segment between radius cr and $c^{-1}r$ is $\left(c^{2/3} - c^{-2/3}\right)L_{tot,b}(r/r_b)^{2/3}$. The choice of c may be related to the branching ratio. For instance, if the typical ratio of the radius between a parent and daughter branch is 2, c may be taken to be $\sqrt{2}$. $L_{tot,b}$ is potentially measurable. However our aim is to reduce the number of parameters as much as possible and it is not immediately clear what an appropriate choice of $L_{tot,b}$ and c are, nor if they even are constants across different species of trees. The $r^{2/3}$ relationship suffices for us to construct the volume distribution without requiring knowledge of $L_{tot,b}$ and c.

To recapitulate, the cylinder size distribution required for the forest canopy layer is as follows. The reference cylinder radius r_b is arbitrarily fixed to be $r_b = 1$ cm. The cylinder length L_b (corresponding to reference radius r_b) is fixed to be $L_b = 1$ m. The lower bound of the distribution is fixed at $r_{min,b} = 1$ mm. The upper bound of the distribution is chosen to be $r_{max,b} = 3$ cm.

The number of canopy layer cylinders per unit radius per unit ground area is of the form

$$n_{cn}p(r) = N_b\left(\frac{r}{r_b}\right)^{-3} \quad (2.43)$$

where the scaling parameter N_b, with units of $\mathrm{m}^{-1}\mathrm{m}^{-2}$, can be thought of as describing the "density" of resonance-sized branches. Associated with each branch of radius r is a length

$$L(r) = L_b\left(\frac{r}{r_b}\right)^{\frac{2}{3}} \quad (2.44)$$

where L_b is the branch length associated with radius $r = r_b$. Note that $L(r)$ may not actually be the true branch length, but rather the cylinder length chosen for our electromagnetic modelling, and thus N_b and L_b may not individually be quantities with true physical meaning,

nor directly measureable. What is physically meaningful, and measurable (though perhaps with tedious effort), is the volume distribution

$$V_b(r) = n_{cn} p(r) \pi r^2 L(r) = N_b L_b \pi r_b^2 \left(\frac{r}{r_b}\right)^{-\frac{1}{3}}. \quad (2.45)$$

$V_b(r)dr$ is to be interpreted as the total volume, per unit ground area, of branches with radius between r and $r + dr$. Also measurable would be the total volume $V_{b,tot}$ per unit ground area, in branches with radius between $r_{min,b}$ and $r_{max,b}$, expressed as

$$V_{b,tot} = \int_{r_{min,b}}^{r_{max,b}} V_b(r)dr = N_b L_b \pi r_b^2 \int_{r_{min,b}}^{r_{max,b}} \left(\frac{r}{r_b}\right)^{-\frac{1}{3}} dr\,. \quad (2.46)$$

$V_{b,tot}$ has dimensions of $\mathrm{m}^3/\mathrm{m}^2$. All other parameters are fixed and if $r_{max,b}$ is also fixed, then $V_{b,tot}$ and N_b are proportional, and we then use $V_{b,tot}$ as the single, primary parameter in our model controlling "amount of vegetation".

Note that this formulation can alternatively be derived from the assumptions of area-preserving branching and the $L_{tot}(r) \propto r^{2/3}$ relationship in the following way: Let the total branch cross-sectional area (per unit ground area) in the area-preserving assumption be A. Then by considering a volume element,

$$V_b(r)dr = A dL_{tot} \quad (2.47)$$

we retrieve the $r^{-1/3}$ dependence of $V_b(r)$:

$$V_b(r) = A\frac{dL_{tot}}{dr} = A L_{tot,b} {r_b}^{-1} \frac{2}{3}\left(\frac{r}{r_b}\right)^{-\frac{1}{3}} \quad (2.48)$$

and by comparison with earlier expression for $V_b(r)$, we can identify

$$N_b = \frac{2L_{tot,b}}{3L_b \pi r_b^3} A \quad (2.49)$$

which reveals that N_b may also be interpreted as being proportional to the stand basal area, or more accurately, the cross-sectional area at base of live crown, per unit ground area. This proportionality and interpretation, however, assumes $L_{tot,b}$ to be a known universal constant, which may not be true.

These general relationships allow us to drastically reduce the number of free parameters in the forward model, and also frees us from depending on local plot-specific empirical distributions of branch cylinder sizes. For numerical computation of the integrals (2.5) and (2.26), the cylinder radius distribution is represented using 40 log-uniform radius bins between $r_{min,b}$ and $r_{max,b}$, and sums performed accordingly.

2.8 Distribution of cylinders in the trunk layer

The final distribution needed is the trunk cylinder size distribution. As before, a preliminary analysis of the dependence of the forward model on the trunk cylinder radius r is performed. This preliminary analysis is performed for a trunk layer of cylinders with their orientations following a Gaussian distribution about the vertical, as described in Section 2.5, with RMS tilt $\sigma_c = 1°$ or $\sigma_c = 5°$. There is no canopy layer for this preliminary trunk layer analysis (set $Z_2 = Z_1$). All the cylinders in the layer have identical sizes. The length, L, is chosen such that the length-to-radius ratio L/r is some fixed value. The cylinder number density n_{trk} is chosen in such a way as to keep the total volume of cylinders per unit ground area $n_{trk} Z_1 \pi r^2 L$ fixed at $10^{-3}\mathrm{m}^3/\mathrm{m}^2$. Figure 2.9 to Figure 2.11 show the results of this preliminary analysis. The one-way extinction (for power or intensity) through the layer for h- and v- polarizations, $\exp\left[-2n_{trk}\langle\kappa_{h,trk}\rangle Z_1/\cos\theta_i\right]$ and $\exp\left[-2n_{trk}\langle\kappa_{v,trk}\rangle Z_1/\cos\theta_i\right]$, and normalized radar cross-sections for the trunk-ground double-bounce $\mathrm{HH}_{trk-gnd,db}$, are plotted against cylinder radius. Also plotted is the ratio $\mathrm{HH}_{trk-gnd,db}/\mathrm{VV}_{trk-gnd,db}$. However for the trunk-ground double-bounce normalized radar cross-sections plotted here, no extinction was applied (pretend $\langle\kappa_{h,trk}\rangle = \langle\kappa_{v,trk}\rangle = 0$), and the ground was set to be perfectly reflecting (set $G_{hh} = G_{vv} = 1$). The radar wavelength was set to 24 cm, and the incidence angle to 38.49°, which is the incidence angle for one of the Aquarius radar beams, and also close to the incidence angle for SMAP.

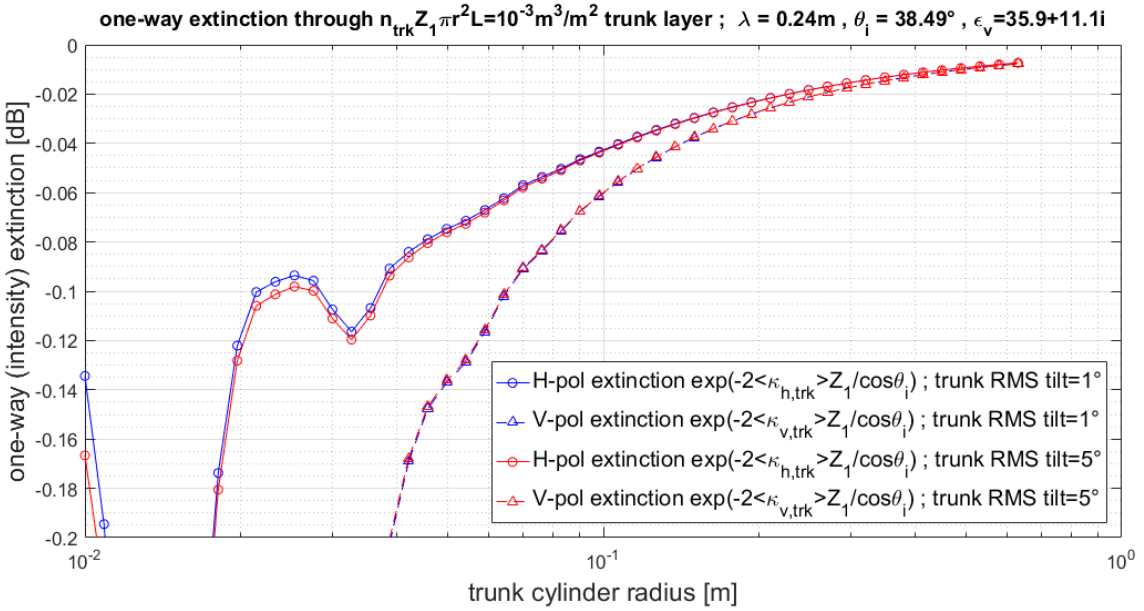

Figure 2.9. One-way extinction at incidence angle of 38.49° for the trunk layer comprising identical cylinders, as a function of trunk cylinder radius r, for trunk RMS tilt angles of 1° (blue curves) and 5° (red curves), and for h-polarization (open circles) and v-polarization (triangles). The total volume of cylinders per unit ground area $n_{trk}Z_1\pi r^2 L$ is kept fixed at $10^{-3}\mathrm{m}^3/\mathrm{m}^2$.

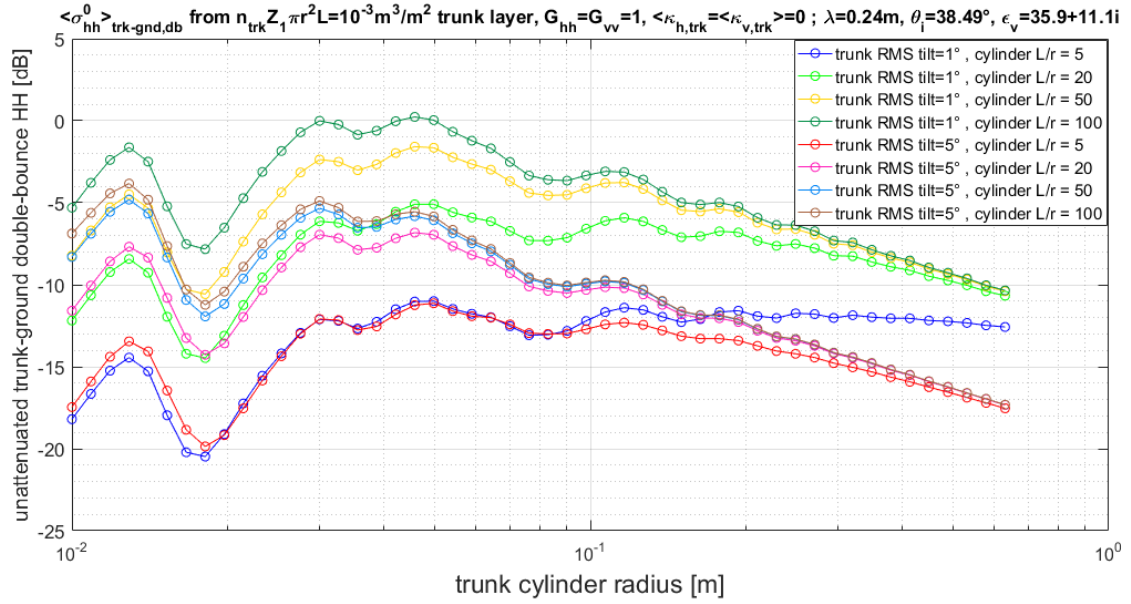

Figure 2.10. Normalized radar backscattering cross-section for the trunk-ground double-bounce $\mathrm{HH}_{trk-gnd,db}$ at incidence angle of 38.49° as a function of trunk cylinder radius, for various trunk RMS tilt angles and cylinder length-to-radius ratios (see legend). No extinction was applied (pretend $\langle\kappa_{h,trk}\rangle = \langle\kappa_{v,trk}\rangle = 0$), and the ground was perfectly reflecting (set $G_{hh} = G_{vv} = 1$). The total volume of cylinders per unit ground area $n_{trk}Z_1\pi r^2 L$ is kept fixed at $10^{-3}\mathrm{m}^3/\mathrm{m}^2$.

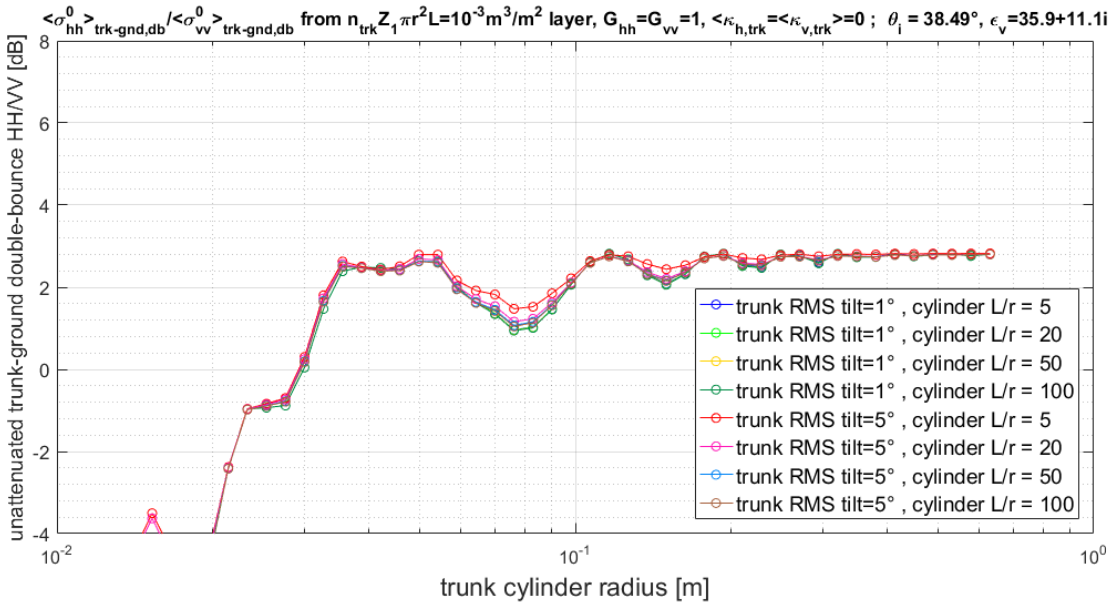

Figure 2.11. Same as Figure 2.10, but for the trunk-ground double-bounce $\mathrm{HH}_{trk-gnd,db}/\mathrm{VV}_{trk-gnd,db}$ ratio.

Several important observations can be made about Figure 2.9 to Figure 2.11. Firstly, as expected from the vertical orientations, the extinction is greater at v-polarization ($\langle\kappa_{v,trk}\rangle > \langle\kappa_{h,trk}\rangle$), and is not strongly dependent on the trunk RMS tilt σ_c. (Also, as mentioned in Section 3.7, if the total cylinder volume is fixed, the extinction is not dependent on the cylinder length.) Secondly, the ratio $\mathrm{HH}_{trk-gnd,db}/\mathrm{VV}_{trk-gnd,db}$ tends towards a fixed value at large cylinder radii and is also not strongly dependent on the trunk RMS tilt σ_c. Finally, most of the interesting dependencies are in the graph for $\mathrm{HH}_{trk-gnd,db}$. As can be seen, at large cylinder radii, the trunk RMS tilt is the more important parameter, while at small cylinder radii, the cylinder length L is the more important parameter. The underlying reason is that, as mentioned before, only cylinders sufficiently close (within an angle proportional to λ/L) to perpendicular to the direction $\mathbf{k}_i + \mathbf{k}_s$ contribute significantly to the backscatter). Given a fixed total volume $n_{trk}Z_1\pi r^2 L$, if the cylinder length L is increased, each cylinder has a larger bistatic scattering contribution scaling as the square L, but the total number of cylinders n_{trk} is inversely proportional to L. If the cylinder length and trunk RMS tilt are small enough such that the $\sigma_c \ll \lambda/L$, all the trunk cylinders contribute significantly (each

with approximately the same contribution). So in this regime, the total $\mathrm{HH}_{trk-gnd,db}$ should increase with increasing cylinder length L but have little dependence on the trunk RMS tilt σ_c. On the other hand, if L is large enough and σ_c is also large enough such that $\sigma_c \gg \lambda/L$, not all the cylinders have a significant contribution; the fraction of cylinders having a significant contribution is inversely proportional to both σ_c and L. The dependence on L thus cancels out, similar to the situation for the upper canopy layer, and the total $\mathrm{HH}_{trk-gnd,db}$ is inversely proportional to the trunk RMS tilt σ_c in this regime.

As before for the canopy layer, the volume distribution as a function of radius is key. Even though there are fewer larger trees than smaller trees, more of the total volume in tree trunks is in the larger trees. Following West et al. [59] and Enquist et al. [60], we use an inverse square distribution for the trunk cylinder radii

$$n_{trk}p_{trk}(r) \propto r^{-2} \tag{2.50}$$

which gives a volume distribution

$$V_{trk}(r) = n_{trk}p_{trk}(r)\pi r^2 L(r). \tag{2.51}$$

For the cylinder lengths, we again apply the relation (2.44)

$$L(r) = L_b(r/r_b)^{\frac{2}{3}} \tag{2.52}$$

to the trunks as well. The range of validity of the inverse square distribution for the trunk cylinder radii, in particular the maximum valid trunk radius, is likely to be species dependent. In the upper canopy layer, $r_{max,b} = 3$ cm was chosen for the branches, so for consistency $r_{min,trk} = 3$ cm is chosen for the lower trunk layer. For the maximum trunk radius $r_{max,trk}$, it is chosen such that

$$\left(\frac{3cm}{r_{max,trk}}\right)^{\frac{2}{3}} = \frac{V_{b,tot}}{V_{b,tot} + V_{trk,tot}} \tag{2.53}$$

where $V_{trk,tot} = \int_{r_{min,trk}}^{r_{max,trk}} V_{trk}(r)dr$ is the total volume in trunks. The rationale for the choice in (2.53) is to be consistent with having branches with $r \leq 3$ cm take up some chosen

volume fraction of the total volume, and also to be consistent with the approximate proportionality $V_{b,tot} \propto r_{max,b}^{2/3} - r_{min,b}^{2/3} \approx r_{max,b}^{2/3}$ (equation 2.46 for $r_{max,b} \gg r_{min,b}$) if it had been extended beyond its supposed range of validity from the branches to the whole tree. Choosing the volume fraction in branches with $r \leq 3\text{cm}$ to be $V_{b,tot}/(V_{b,tot} + V_{trk,tot}) = 0.2$ and solving yields a value of $r_{max,trk} = 33.5$ cm that we use as the maximum trunk radius in our model. Since the trunk extinctions and bistatic cross-sections are slowly varying when cylinder radius is large, of importance here is not so much whether $r_{max,trk} = 33.5$ cm is an accurate guess, but that we have fixed $V_{b,tot}$ to be 20% of the total volume. This 20% value is an estimate that seems consistent with that by Beaudoin et al. [51]. The other key parameter is the trunk RMS tilt σ_c. For a reasonable guess, from Chauhan et al. [16], we expect that $\sigma_c < 10°$. Here we use the $\sigma_c = 5°$ guess made by Durden et al. [14], also consistent with Beaudoin et al. [51], though all these studies were only on coniferous forests.

Under these assumptions, we compute and tabulate in Table 2.1 to Table 2.5 below some quantities relevant to computing equations (2.28b-c) for the trunk-ground double reflections.

$\exp\left[-2n_{trk}Z_1\langle\kappa_{h,trk}\rangle/\cos\theta_i\right]$				
	$\varepsilon_v = 5.15 + 1.41i$	$\varepsilon_v = 17.1 + 5.8i$	$\varepsilon_v = 35.9 + 11.1i$	$\varepsilon_v = 62.8 + 18.2i$
$\theta_i = 29.36°$	-0.042282dB	-0.031799dB	-0.029454dB	-0.026318dB
$\theta_i = 38.49°$	-0.053868dB	-0.041922dB	-0.038671dB	-0.035311dB
$\theta_i = 46.29°$	-0.066636dB	-0.053518dB	-0.049232dB	-0.045651dB

Table 2.1. One-way h-polarization extinction $\exp\left[-2n_{trk}Z_1\langle\kappa_{h,trk}\rangle/\cos\theta_i\right]$ for a $V_{trk,tot} = 10^{-3}\text{m}^3\text{m}^{-2}$ trunk layer, for several values of incidence angle θ_i and trunk cylinder relative permittivity ε_v .

$\exp\left[-2n_{trk}Z_1\langle\kappa_{v,trk}\rangle/\cos\theta_i\right]$				
	$\varepsilon_v = 5.15 + 1.41i$	$\varepsilon_v = 17.1 + 5.8i$	$\varepsilon_v = 35.9 + 11.1i$	$\varepsilon_v = 62.8 + 18.2i$
$\theta_i = 29.36°$	-0.054415dB	-0.044106dB	-0.046044dB	-0.044469dB
$\theta_i = 38.49°$	-0.070223dB	-0.058606dB	-0.060656dB	-0.059221dB
$\theta_i = 46.29°$	-0.087169dB	-0.074483dB	-0.076645dB	-0.075310dB

Table 2.2. Same as Table 2.1, but for v-polarization.

$16\pi n_{trk} Z_1 \langle \lvert S_{hh} \rvert^2_{trk,3a} \rangle / k^2$				
	$\varepsilon_v = 5.15 + 1.41i$	$\varepsilon_v = 17.1 + 5.8i$	$\varepsilon_v = 35.9 + 11.1i$	$\varepsilon_v = 62.8 + 18.2i$
$\theta_i = 29.36°$	-10.9376dB	-10.7324dB	-10.1299dB	-9.9533dB
$\theta_i = 38.49°$	-12.5545dB	-10.8279dB	-10.0889dB	-9.7269dB
$\theta_i = 46.29°$	-13.5699dB	-10.9303dB	-10.1317dB	-9.6721dB

Table 2.3. Hypothetical normalized radar cross-section at h-polarization $16\pi n_{trk} Z_1 \langle \lvert S_{hh} \rvert^2_{trk,3a} \rangle / k^2$ for a $V_{trk,tot} = 10^{-3}\mathrm{m}^3\mathrm{m}^{-2}$ trunk layer with perfectly reflecting ground and without extinction considerations, for several values of incidence angle θ_i and trunk cylinder relative permittivity ε_v.

$16\pi n_{trk} Z_1 \langle \lvert S_{vv} \rvert^2_{trk,3a} \rangle / k^2$				
	$\varepsilon_v = 5.15 + 1.41i$	$\varepsilon_v = 17.1 + 5.8i$	$\varepsilon_v = 35.9 + 11.1i$	$\varepsilon_v = 62.8 + 18.2i$
$\theta_i = 29.36°$	-23.0373dB	-17.1299dB	-13.7693dB	-12.2022dB
$\theta_i = 38.49°$	-20.3795dB	-14.5201dB	-12.2302dB	-11.0765dB
$\theta_i = 46.29°$	-18.1205dB	-13.3495dB	-11.5116dB	-10.5306dB

Table 2.4. Same as Table 2.3, but for v-polarization $16\pi n_{trk} Z_1 \langle \lvert S_{vv} \rvert^2_{trk,3a} \rangle / k^2$.

$\langle \lvert S_{hh} \rvert^2_{trk,3a} \rangle / \langle \lvert S_{vv} \rvert^2_{trk,3a} \rangle$				
	$\varepsilon_v = 5.15 + 1.41i$	$\varepsilon_v = 17.1 + 5.8i$	$\varepsilon_v = 35.9 + 11.1i$	$\varepsilon_v = 62.8 + 18.2i$
$\theta_i = 29.36°$	12.0996dB	6.3975dB	3.6395dB	2.2489dB
$\theta_i = 38.49°$	7.8249dB	3.6923dB	2.1413dB	1.3496dB
$\theta_i = 46.29°$	4.5506dB	2.4192dB	1.3800dB	0.8585dB

Table 2.5. Same as Table 2.3, but for the ratio $\langle \lvert S_{hh} \rvert^2_{trk,3a} \rangle / \langle \lvert S_{vv} \rvert^2_{trk,3a} \rangle$.

To summarize this section, using many assumptions, we have constrained the trunk cylinder distribution to have no free parameters. The orientation distribution is uniform in ϕ_c and Gaussian distributed about the vertical with RMS tilt $\sigma_c = 5°$, the cylinder lengths are pegged to the radii by a 2/3 power law, the radius distribution is inverse square between 3 cm and 33.5 cm, and the total volume in trunk layer cylinders is fixed to be 4 times the total volume in the upper vegetation volume layer. Like before, numerical computation and integrals are carried out using a grid of log-uniform radius bins.

Chapter 3

APPLICATION OF L-BAND RADAR BACKSCATTER MODEL TO AQUARIUS DATA OVER GLOBAL FORESTS

3.1 Chapter overview

In this chapter, our L-band radar backscatter model shall be compared with the data from the spaceborne L-band scatterometer of Aquarius. The Aquarius mission is unique for having had both significantly long temporal coverage (slightly more than 3 years), frequent (weekly) repeat global coverage, and a L-band scatterometer with high accuracy and stability (<0.2dB) at all polarization channels [61, 62, 63]. Frequent repeats are important for studying changes in ground moisture, which may rise rapidly with precipitation in a matter of hours, and then dry down on a time scale of days.

In subsequent sections (Sections 3.3-3.7) of this chapter, the comparison between our L-band radar backscatter model and the data from the Aquarius scatterometer will be performed class-by-class for the five forest classes, with reference to the global land cover classes of the International Geosphere-Biosphere Programme (IGBP) [35], bearing in mind the coarse ~100km spatial resolution and 7-day temporal resolution. Physically reasonable representative values for the input parameters for the forward model shall be chosen for each class by fitting to the data and/or consulting values from the literature. Following Kim et al. [22], the real part of the ground relative permittivity ε_g is typically not expected to fall outside the range 2.7-40, and the ground surface roughness RMS height h is typically expected within the 0.5-5cm range. Fitted branch volume values $V_{b,tot}$ shall be compared to independent aboveground biomass estimates. It shall be shown that our relatively parsimonious forward model is in overall reasonable quantitative agreement with the data at these global spatial scales, particularly for needleleaf forests, but slightly worse for broadleaf forests; the neglect of modelling leaves in the dense jungles are likely to have caused some inaccuracies. Sensitivity of L-band multi-polarization radar to vegetation structure [64, 65] was also verified at our global spatial scales: a preferentially horizontal orientation

distribution for the canopy cylinders was more suited for needleleaf forests, a uniformly random orientation distribution was more suited for broadleaf forests, and mixed forests somewhere in between. Sensitivity to sub-canopy flooding and differences between frozen/unfrozen states was expected [66, 67] and our model provides partial quantitative agreement by modelling them in terms of changes in ground surface roughness, ground dielectric relative permittivity, and vegetation dielectric relative permittivity; the frozen/unfrozen comparison will be in Section 3.8. Microwave vegetation optical depth values will be reported in Section 3.9.

3.2 Input datasets

The Aquarius/SAC-D mission was launched in 2011 with the primary goal of measuring sea surface salinity from space. Here we use the portion of data that was taken over land. The instrument carried 3 radiometers operating at 1.41GHz, with beams at incidence angles of about 29°, 38°, and 46°. It also carried a scatterometer at 1.26GHz that shares the feed horns with the radiometers. The scatterometer measured, for each incidence angle, normalized radar backscattering cross-sections HH, HV, VH, and VV, with footprints ~100km. Further instrument details can be found in Le Vine et al. [61], Fore et al. [62], and Yueh et al. [68]; in particular, Fig. 3 from the paper by Le Vine et al shows a schematic of the radiometer and scatterometer footprints of the 3 beams, reproduced here in Figure 3.1.

The spacecraft orbit is sun-synchronous at about 657km altitude with a 7-day repeat period. With the exception of cross-over points and high-latitude regions, a point on the ground is typically visited only once every 7-days by one of the three beams (thus only one of the local incidence angles 29°, 38°, or 46°) on either the ascending (local time ~6PM) or descending pass (local time ~6AM) of the spacecraft. The full data record ran from August 25, 2011 through June 7, 2015. We used the Level 2 scatterometer data from the Aquarius dataset version 4.0 release; the data was obtained from the NASA EOSDIS Physical Oceanography Distributed Active Archive Center (PO.DAAC) at the Jet Propulsion Laboratory, Pasadena, CA [69, 70]. The global Aquarius scatterometer dataset was then regridded into a 36km EASE-Grid 2.0 [71], i.e. the globe was gridded into 406x964 equal-area pixels (henceforth

referred to as EASE2 grid pixels). This regridding reassigns data based on which EASE2 grid pixel the centre of each radar footprint was nearest to. Note that the footprint size is significantly larger than the pixel size.

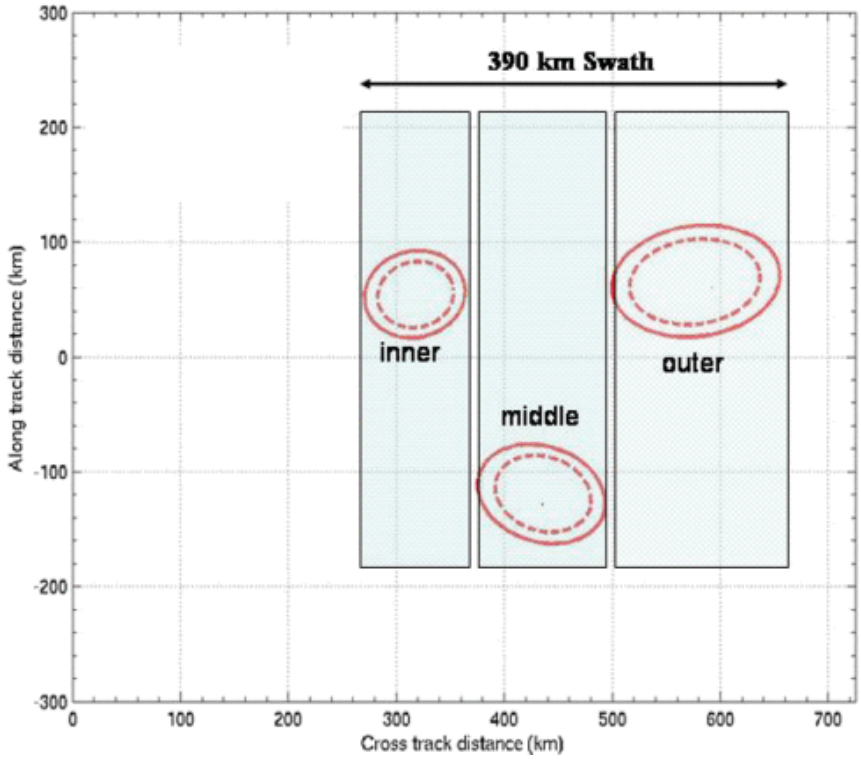

Figure 3.1. Aquarius 3-dB footprints and swath of the three beams for radiometer (solid lines) and scatterometer (dashed lines), as shown in Fig. 3. of Le Vine et al. [61] © IEEE 2007

Some data filtering was performed before comparison with the model. Provided together with the Level 2 scatterometer data in the Aquarius dataset were data quality flags for non-nominal data conditions. There were many possible factors and conditions that triggered flagging. One primary contributor was radio-frequency interference (RFI) corruption. Figure 3.2 displays the fraction of overall "good quality" time series data (25Aug2011 to 07Jun2015) over land from the Aquarius Level 2 scatterometer dataset. (The overall quality flags must be met for all polarizations.) Note from Figure 3.2 the significant impact of RFI from heavily populated and developed parts of the world. Note also the exclusion of pixels that have a non-negligible amount of surface water bodies within the scatterometer footprint.

Some swath and regridding artefacts are visible in Figure 3.2, but these artefacts should have little impact on the comparison with the model and the inferences drawn.

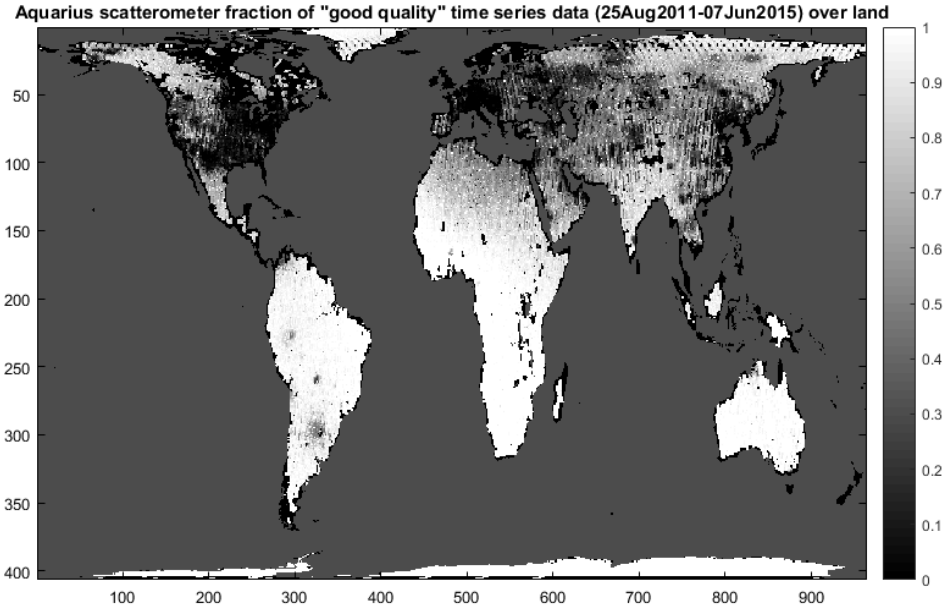

Figure 3.2. Global map (on EASE2 grid) of fraction of "good quality" time series data (25-Aug-2011 to 07-Jun-2015) over land from the Aquarius Level 2 scatterometer dataset. The grayscale runs from black (fraction=0) to white (fraction=1). Some swath and regridding artefacts are present.

Additional filtering was also performed to select only data from areas where the terrain slope was not too steep. This is because our L-band radar backscatter model as described in Chapter 3 implicitly assumed that the ground under the forest was flat. For this filtering, we obtained global digital elevation data from the Harmonized World Soil Database v1.2 [72]. For lands below 60° latitude, we obtained from the database distributions of absolute slopes, binned into eight slope classes (0-0.5%, 0.5-2%, 2-5%, 5-10%, 10-15%, 15-30%, 30-45%, >45%), on a global 5arc-minute latitude-longitude grid. The original data for these distributions were 3arcsecond NASA Shuttle Radar Topographic Mission (SRTM) data. We regridded the 5arc-minute grid into the 36km EASE2 grid and computed the median slope class for each pixel. SRTM did not cover areas beyond 60° latitude. The primary landmass south of 60° latitude is Antarctica, which is excluded from this study. For land north of 60° latitude, we obtained from the database digital elevation data on a 30arc-second latitude-longitude grid (the

original source for the database was 30 arcsecond elevation data from USGS GTOPO30). From this 30arc-second data we computed median absolute slopes within each 36km-EASE2grid pixel.

One can expect median slopes calculated from initially 30arc-second data to be less steep than if calculated from initially 3arc-second data. This is indeed true, as displayed in Figure 3.3, which shows a comparison between them for pixels in between 40°N and 60°N latitude.

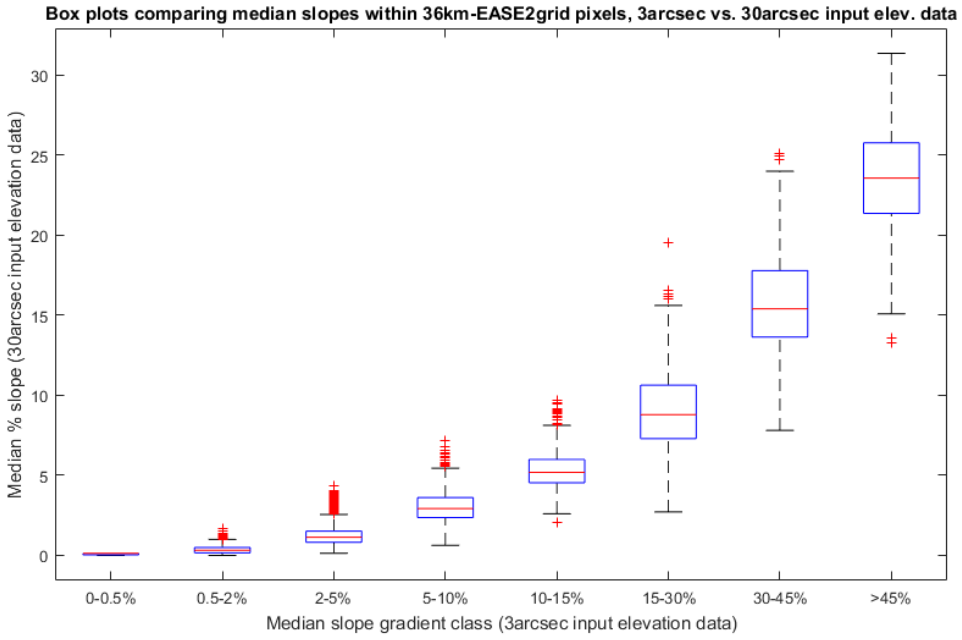

Figure 3.3. Box plots comparing median slopes within 36-km EASE2 grid pixels, calculated from initially 30arc-second data, vs. calculated from initially 3arc-second data. Compared pixels are between 40°N and 60°N latitude.

Our choice of cutoff for terrain slope is thus as follows. For 36-km EASE2 grid pixels below 60°N latitude (i.e. initial data was 3arcsecond SRTM data), we flag pixels with >15% median slope as "high terrain slope". For 36-km EASE2 grid pixels above 60°N latitude (i.e. initial data was 30 arcsecond USGS GTOPO30 data), we flag pixels with >7% median slope as "high terrain slope". These pixels are displayed in black in Figure 3.4. Data from these pixels will be excluded from comparison with the radar backscatter forward model.

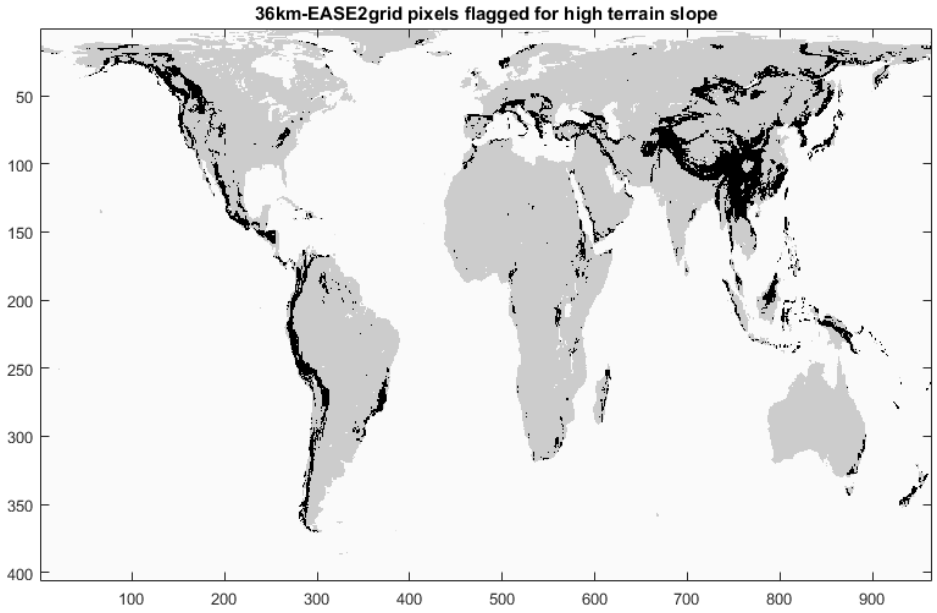

Figure 3.4. 36km EASE2grid pixels flagged for high terrain slope are displayed in black. See main text for details.

Radar backscatter is very different between frozen and unfrozen conditions, as is well known [73, 74]. The comparison and analysis of radar backscatter between frozen and unfrozen conditions will be in Section 4.7. For the analysis in Sections 4.2-4.6, a further data filtering was performed, retaining only data for unfrozen conditions; if otherwise unfiltered, the large change between frozen and unfrozen conditions would have dominated and swamped other effects. Specifically, this filter retained only data with corresponding temperature >5°C during a time period from two weeks before to two weeks after, and only EASE grid pixels with at least 20 weeks of such data. The temperature value used is the temperature of the 0-10cm subsurface layer, as provided by the ancillary NCEP GFS GDAS model product provided along with the Aquarius dataset. (Abbreviations: NCEP: National Centers for Environmental Prediction; GFS: Global Forecast System; GDAS: Global Data Assimilation System.)

Each EASE2 grid pixel on the globe was labelled with a land cover class, with reference to the global land cover classes of the International Geosphere-Biosphere Programme (IGBP). To obtain class labels for each pixel, the MODIS Land Cover Type Product (MCD12Q1) was used. The MODIS Land Cover Type Product (MCD12Q1) in the IGBP Land Cover Type Classification was obtained from the Global Land Cover Facility at a 5-minute (in both latitude and longitude) resolution for the year 2012 [75, 76]. Briefly, this land cover map is the product of a classification algorithm that uses the Moderate-Resolution Imaging Spectroradiometer (MODIS) visible/near-infrared data as primary input. This was then regridded into the 36-km EASE2 grid by taking the mode of the finer 5-minute resolution pixels within each EASE2 grid pixel. The regridded land cover map is shown in Figure 3.5.

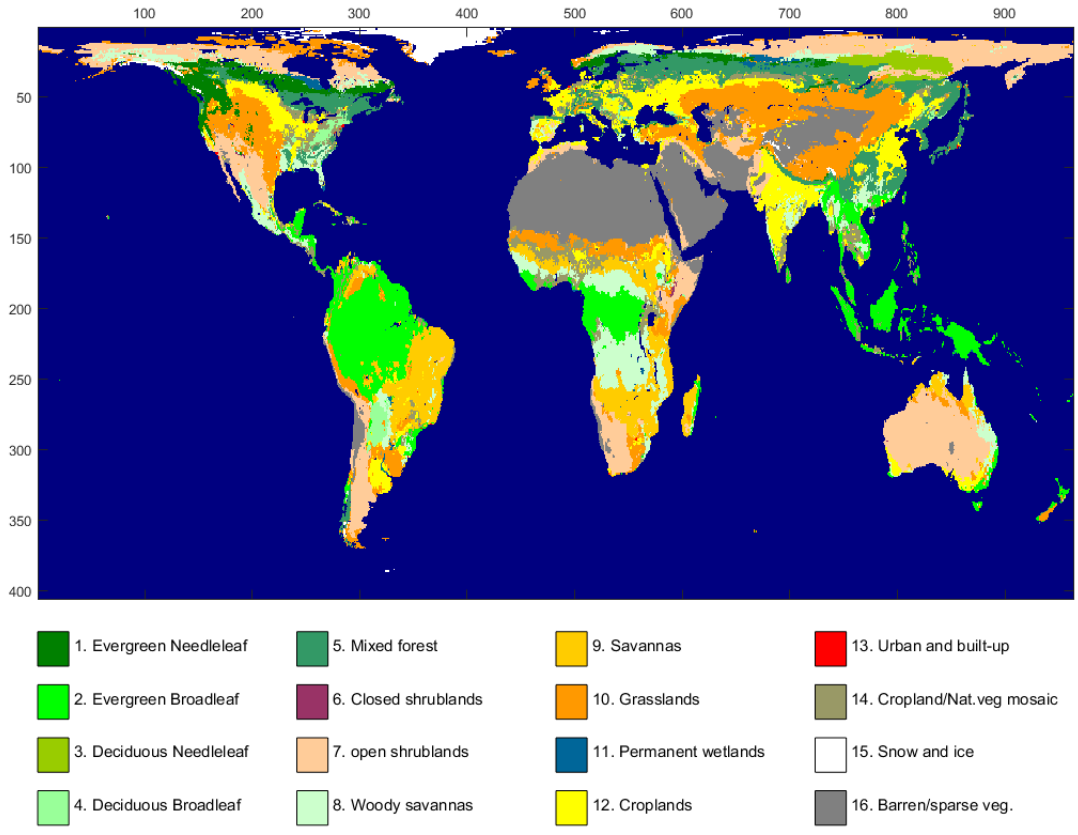

Figure 3.5. MODIS Land Cover map, IGBP (International Geosphere-Biosphere Programme) classification, 2012 data.

3.3 Evergreen needleleaf forests (IGBP class 1)

Using the MODIS land cover map, the geographic distribution of the evergreen needleleaf forests (IGBP class 1) is shown in Figure 3.5 and Figure 3.6 for emphasis. From the filtered data for unfrozen conditions, a random 20% selection (to avoid clutter on subsequent scatterplots) of the EASE2 grid pixels is displayed in Figure 3.6.

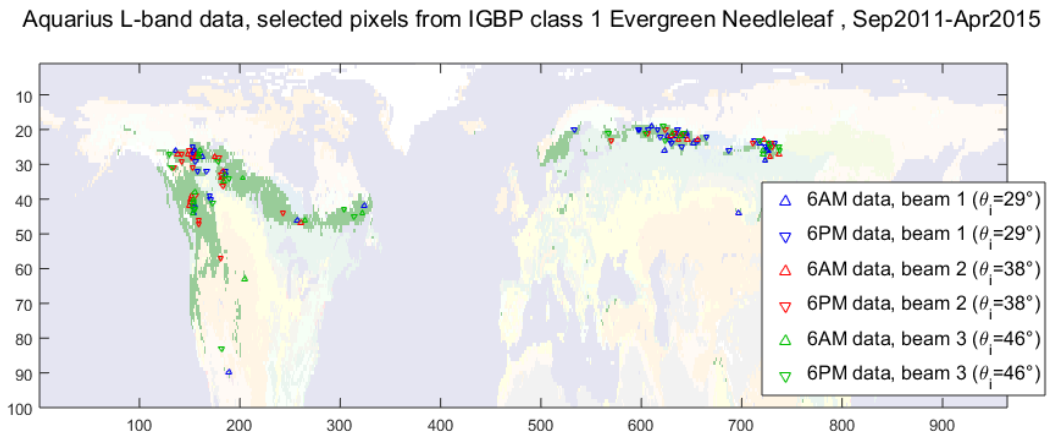

Figure 3.6. Geographic distribution of evergreen needleleaf forests (IGBP class 1) and randomly selected 20% of filtered Aquarius data EASE2 grid pixels, marked by open triangle symbols.

The scatterometer data for these pixels is displayed in Figure 3.7, which depicts two scatterplots: one of VV/HH[dB] vs HH[dB], and another of HV/HH[dB] vs HH[dB]. In Figure 4.2-2, the three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°), and data for descending-pass (local time 6AM) and ascending-pass (local time 6PM) are represented by upward-pointing and downward-pointing triangles, respectively. The symbol descriptions are in the figure caption. Each open triangle represents the mean value (of the plotted quantities) of data taken over unfrozen periods from Sep2011-Apr2015 from its corresponding pixel shown in Figure 3.6. As noted earlier, only a small minority of pixels may have data for more than one incidence angle or for both ascending and descending passes, and even then the actual footprints may not be the same. Thus comparisons between incidence angles or 6AM/6PM may only be made assuming similarity between pixels. Furthermore, the azimuth angles for the same beam are different on ascending and descending pass, and not by a simple 180° because of the footprint positions

depicted in Figure 3.1 as well as the inclination of the sun-synchronous orbit. In view of these confounding factors, we avoid jumping to any conclusions based only on comparisons between 6AM and 6PM Aquarius data unless the differences are very obvious.

The filled triangles in Figure 3.7 are the median (marginal medians, i.e. the component-wise median) value of the open triangles, but with median taken over all data pixels of this class instead of only the 20% displayed. Filled circles show the results from the forest forward model of Chapter 2. The parameters $m = 1$ and $\theta_0 = \pi/2$ (cosine-squared orientation distribution about the horizontal) in equation (2.6) were used, and the vegetation cylinder relative permittivity chosen as $\varepsilon_v = 29.9 + 9.5i$. These orientation parameters, which will also be applied to IGBP class 3 (deciduous needleleaf forests), were chosen because many coniferous trees have branches with a preferential horizontal orientation; our distribution is similar to various results from direct measurements made by Saleh et al. for Maritime pine trees [77], Chauhan et al. for hemlock trees [16], and Jiang et al. for Dahurian larches [57]. Other model input parameters are: total volume in branches $V_{b,tot} = 3.1 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, ground relative permittivity $\varepsilon_g = 8.8$, ground roughness RMS height $h = 2.6\mathrm{cm}$. These latter three parameters ($V_{b,tot}$, ε_g , h) were chosen by fitting to the pooled (ascending and descending) data. The fitting procedure minimized the sum of component-wise absolute distances between the model and the data, i.e.

$$\underset{V_{b,tot}\,,\varepsilon_g,h}{\text{minimize}} \quad \sum_{data\ pix} \sum_{i=1}^{3} \left| x_i^{(data\ pix)} - x_i^{(model)}\left(V_{b,tot}\,,\varepsilon_g,h\right) \right| \qquad (3.1)$$

$$x_1 = \mathrm{HH[dB]}\,, \qquad x_2 = \mathrm{VV/HH[dB]}\,, \qquad x_3 = \mathrm{HV/HH[dB]}\,.$$

To clarify, $x_i^{(data\ pix)}$ is the mean value of radar feature i ($i = 1,2,3$) for the EASE2 pixel indexed by "$data\ pix$" , over unfrozen periods from Sep2011-Apr2015 for data.

Strictly speaking, ε_g is complex and should have an imaginary component, which we have neglected. From Wang and Schmugge [78] and Hallikainen et al. [79], we know that this imaginary part depends both on the soil moisture as well as the soil type; it is typically at least several times smaller than the real part. Because the soil affects the radar backscatter

only through backscatter and reflection, not involving transmission through the soil, the modelling error incurred by neglecting the imaginary part is thus small; Oh et al. [80] also recognized that neglecting the imaginary part of ε_g is a valid approximation for the Fresnel reflectivity term off the soil surface.

Several remarks can be made regarding Figure 3.7. Firstly, the radar backscatter at all polarizations generally decrease with increasing incidence angle from 29° to 46°, as expected from even the simplest volume scattering models [29]. Secondly, the forward model seems to provide a reasonable agreement to the data, albeit using fitted parameters. The fitted ground parameters fall within their expected ranges. Let us compare the fitted branch volume with aboveground biomass estimates from the literature. We assume a value of wood density (or specific gravity; this is the ratio of "oven-dry weight" to "green volume") of ρ=0.45g/cm^3 based on Nelson et al. [81] and Zanne et al. [82] for IGBP classes 1, 3, and 5, and retain our earlier assumption from chapter 3 of $V_{trk,tot} = 4V_{b,tot}$. Taking aboveground biomass $M_{tot} = \rho\left(V_{b,tot} + V_{trk,tot}\right)$, we find that $V_{b,tot} = 3.1$ litres/m^2 corresponds to 69.8tons/ha of aboveground biomass. Looking at the literature, we find aboveground biomass estimates of 30-100 tons/ha by Margolis et al. [83] for boreal forests in North America based on LiDAR techniques; estimates of 40-120 tons/ha by Shepashenko et al. [84] for Siberian forest phytomass (including roots) using Russian state forest inventories and regression analysis; estimates of 30-150 tons/ha depending on stand age by Beaudoin et al. [51] for maritime pine at the Landes forest in France. Thirdly, there might be an overall trend of L-band radar backscatter being slightly higher from ascending-pass (local time 6PM) data compared to descending-pass (local time 6AM) data, but this difference, when considered in an average sense over time and multiple pixels, is small (compared to the spatial variation between pixels and longer period temporal variations to be discussed later). Friesen et al. found evening (ascending track; 9:30PM-12AM) radar backscatter values higher than morning (descending 9:30AM-12PM) at C-band VV by about 1dB over Russia/Canada, using ESCAT data [85]. We bear this in mind, but proceed to pool the 6AM and 6PM data for several of the subsequent analyses.

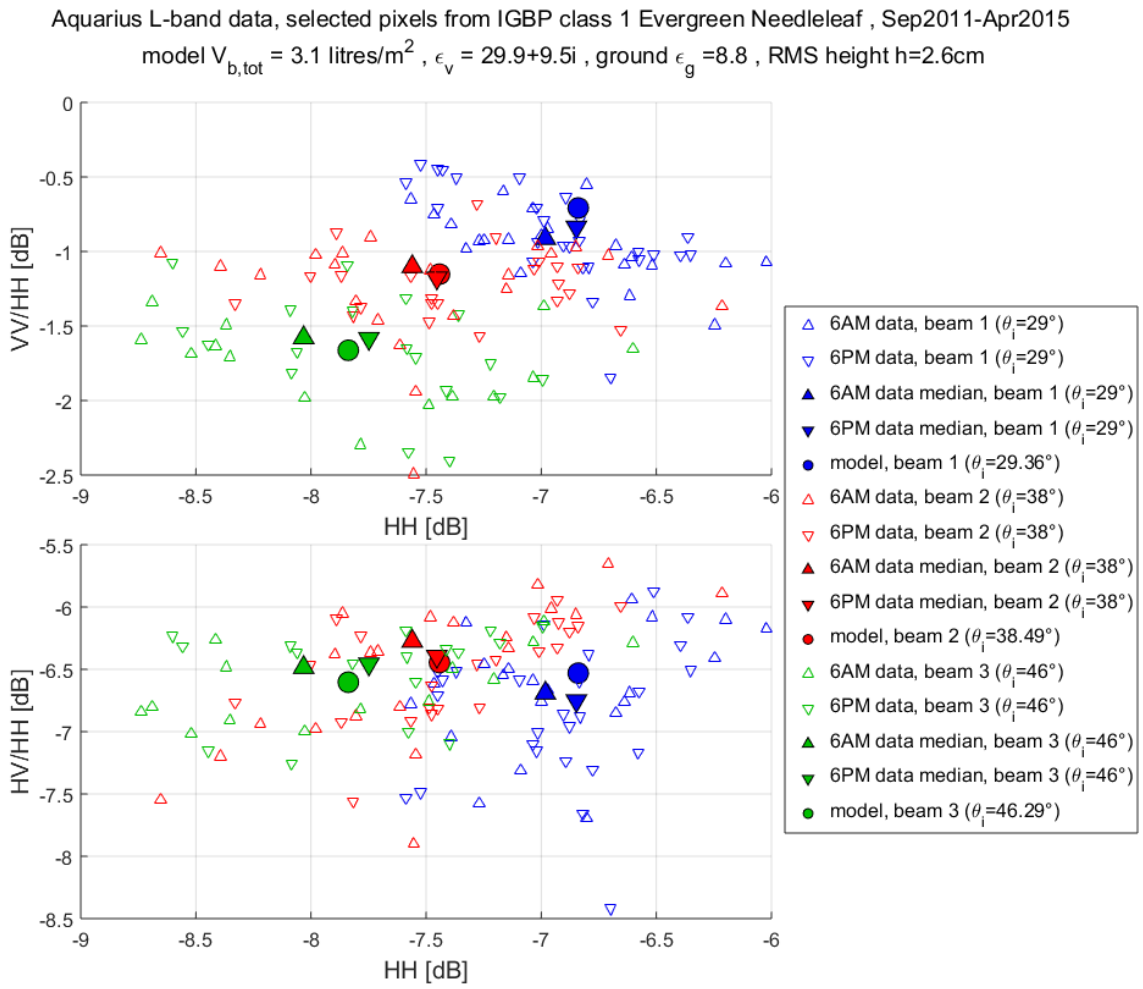

Figure 3.7. Aquarius L-band scatterometer data for EASE2 grid pixels classified as evergreen needleleaf forests (IGBP class 1). The three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°). Upward-pointing triangles are for descending-pass (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM). Each open triangle represents data from one of the pixels shown in Figure 3.6, taking the mean over unfrozen periods from Sep2011-Apr2015 for each pixel. Filled triangles are the median over open triangles, but with median taken over all data pixels of this class instead of only the 20% displayed. Filled circles show the forest forward model from Chapter 2, with input parameters $m = 1$ and $\theta_0 = \pi/2$ (preferential horizontal orientation) in equation (2.6), total volume in branches $V_{b,tot} = 3.1 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, ground relative permittivity $\varepsilon_g = 8.8$, ground roughness RMS height $h = 2.6\mathrm{cm}$.

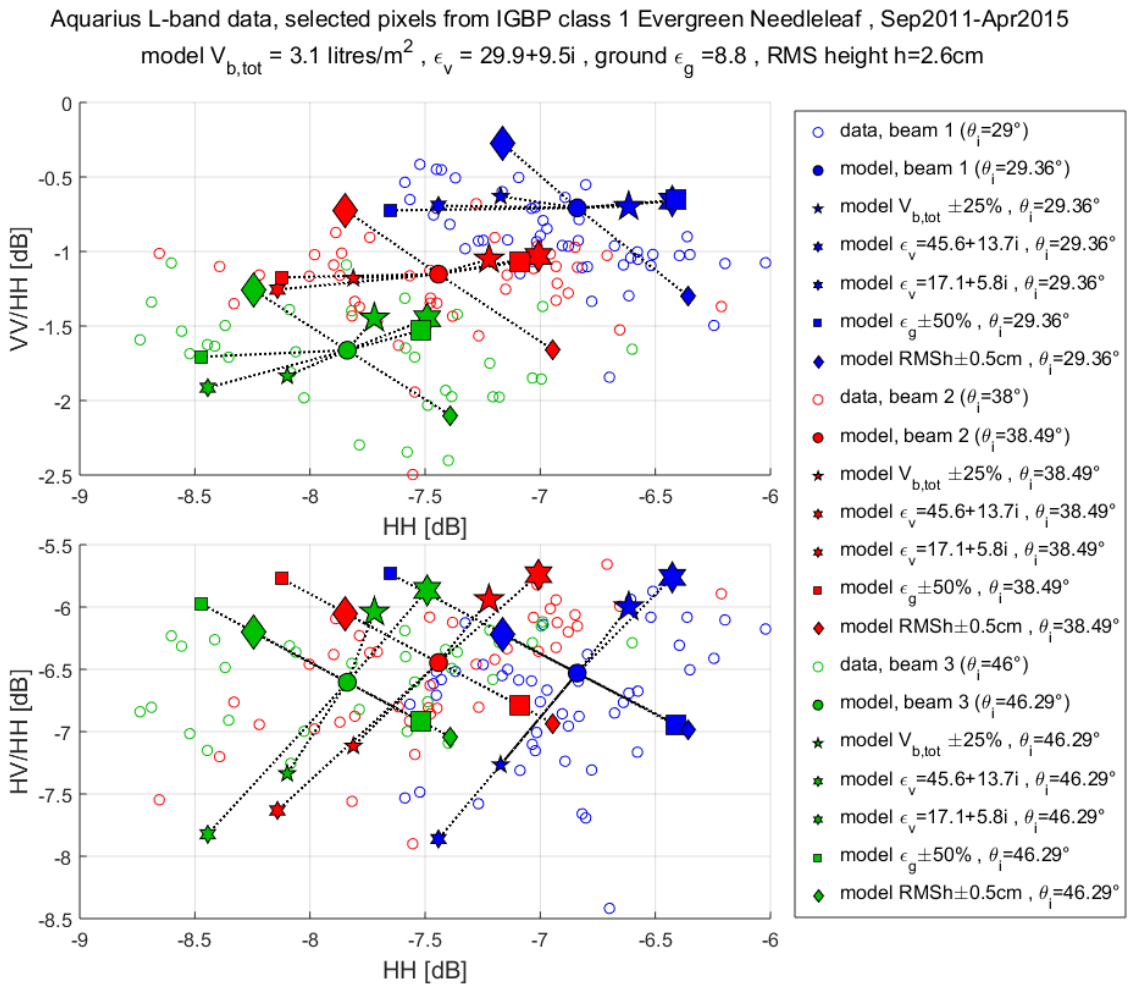

Figure 3.8. Same as Figure 3.7, but with no distinction made between descending- and ascending-pass data; all are marked with open circles. Filled circles show the forward model with input parameters as before in Figure 3.7 ($m = 1$ and $\theta_0 = \pi/2$ in equation (2.6), $V_{b,tot} = 3.1 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, $\varepsilon_v = 29.9 + 9.5i$, $\varepsilon_g = 8.8$, $h = 2.6\mathrm{cm}$). Other filled symbols explore the nearby model parameter space if one parameter is perturbed. Larger and smaller five-pointed stars represent a change in $V_{b,tot}$ by +25% and -25% respectively. Larger six-pointed stars represent an increase in ε_v to $\varepsilon_v = 45.6 + 13.7i$, while smaller six-pointed stars represent a decrease to $\varepsilon_v = 17.1 + 5.8i$. Larger and smaller squares represent a change ε_g by +50% and -50% respectively. Larger and smaller diamonds represent a change in h by +0.5cm and -0.5cm respectively.

The model is then used to analyze the spatial variation in the radar backscatter data across different pixels. Figure 3.8 displays the same data as Figure 3.7, now with the descending-pass (local time 6AM) and ascending-pass (local time 6PM) data pooled and all marked with open circles instead of triangles. The three different beam incidence angles are again

represented with blue, red, and green colors (29°, 38°, 46° respectively). Overlaid on the data are filled symbols showing the results from the forward model. The filled circles show the forward model with parameters as before. Other filled symbols explore the nearby model parameter space if one parameter is perturbed. Larger five-pointed stars represent an increase in $V_{b,tot}$ by 25%, while smaller five-pointed stars represent a decrease in in $V_{b,tot}$ by 25%. Larger six-pointed stars represent an increase in ε_v to $\varepsilon_v = 45.6 + 13.7i$, while smaller six-pointed stars represent a decrease to $\varepsilon_v = 17.1 + 5.8i$. Larger squares represent an increase in ε_g by 25%, while smaller squares represent a decrease in ε_g by 25%. Larger diamonds represent an increase in h by 0.5cm, while smaller diamonds represent a decrease in h by 0.5cm. From the scatterplots, it seems that effects of the two vegetation parameters are difficult to distinguish from each other. In view of the scatterplot of VV/HH[dB] vs HH[dB], it appears that the ground roughness RMS height h is one of the physical parameters involved in the spatial variations. However, when both scatterplots are considered together, no one single physical parameter appears sufficient to explain all or most of the spatial variation across different pixels.

It is interesting to note that when the scatterplot of HV/HH[dB] vs HH[dB] is considered, ground parameters (ε_g and h) vary the radar backscatter in an approximately orthogonal direction on the plot as compared to varying vegetation parameters ($V_{b,tot}$ and ε_v), so it is easier to distinguish ground effects vs. vegetation effects on this scatterplot, as compared to the scatterplot of VV/HH[dB] vs HH[dB], where there is greater degeneracy between the effects of perturbing different physical parameters. This can be physically understood as follows. In our forward model, the primary scattering mechanism contributing to HV[dB] is backscatter from the forest canopy layer. The ground parameters have little direct effect on HV[dB], so when HH[dB] changes due to ground parameters, HV/HH[dB] changes in an anti-correlated way. In this part of the parameter space, vegetation backscatter is not yet saturated, so vegetation parameters affect backscatter at all polarizations in a correlated manner, but change HV[dB] slightly more than HH[dB] (because of the backscatter from the canopy, as well “dilution” of HH[dB] by the other scattering mechanisms). Thus HV/HH[dB] and HH[dB] are also correlated in changes due to vegetation parameters. Note

however that this is not necessarily true everywhere in the parameter space, e.g. in the saturation regime.

Next, temporal variations in the radar backscatter are considered. We summarize some of the temporal information for each pixel in the following way. The sample covariance of the radar data (arranged as a vector of three features HH[dB], VV/HH[dB] , HV/HH[dB], with mean subtracted) over the weekly time-series is computed as a 3x3 matrix. (Recall that frozen weeks were excluded.) Let λ_1 , λ_2 , λ_3 be the 3 real nonnegative eigenvalues, in descending order, in the orthogonal eigendecomposition of this covariance matrix. It turns out from the data that typically this 3x3 covariance matrix may be acceptably approximated by its rank-1 approximation. This is shown in Figure 3.9, which plots histograms of pixel counts binned by the fraction $\lambda_1 / (\lambda_1 + \lambda_2 + \lambda_3)$. Note that all eligible pixels of IGBP class 1 (evergreen needleleaf), not just the 20% random subsample, were counted in these histograms. The fraction $\lambda_1 / (\lambda_1 + \lambda_2 + \lambda_3)$ represents the fraction of the total covariance in the 3x3 matrix that is retained by its rank-1 approximation, and as can be seen from Figure 3.9, this fraction is typically greater than 0.7 or 0.8 for most pixels at all three incidence angles. We also note from Figure 3.9 that this fraction shows no big differences between the ascending-pass (local 6PM) and descending-pass (local 6AM). Having verified the validity of the rank-1 approximation, the dominant eigenvalue λ_1 and its corresponding dominant eigenvector are then used for comparison with the forward model in Figure 3.10. Figure 3.10 is essentially the same as Figure 3.8, but also showing temporal variations in the data. As before, each open circle in Figure 3.10 corresponds to one of the pixels shown in Figure 3.6 and represents the mean taken over unfrozen periods from Sep2011-Apr2015; but now each open circle also comes with a line through it that visually summarizes the temporal variation. Specifically, the direction of this line displays the dominant eigenvector of the 3x3 covariance matrix for that pixel as discussed earlier, and the half-length of the displayed line is $\sqrt{\lambda_1}$.

From Figure 3.10, it appears from the eigenvectors that for most of the pixels, HH[dB] is changing over time, VV/HH[dB] is not changing much over time, and HV/HH[dB] is changing slightly over time, and in a manner that is anti-correlated with the change in HH[dB]. Comparing with the model, it can be seen that if we had to choose only one physical

factor to explain the temporal variations in the weekly data (over unfrozen periods), it would be the ground relative permittivity ε_g, i.e. changes in soil moisture. It is the only physical parameter out of the four considered here that, when perturbed, gives variation in the radar data-space that is roughly, but not exactly, in alignment with the observed directions of those dominant eigenvectors, in both scatterplots (VV/HH[dB] vs HH[dB], and HV/HH[dB] vs. HH[dB]). Furthermore, the lengths of the lines correspond to a reasonable amount of change in ε_g. Recall that the half-length of the displayed line is $\sqrt{\lambda_1}$; in our rank-1 approximation, the half-line can be thought of as displaying a standard deviation of variation. In the median case, the values of ground relative permittivity corresponding to ±1 standard deviation of variation in HH[dB] are approximately $\varepsilon_g = 5.8$ and $\varepsilon_g = 14.1$. Though details vary slightly depending on soil type, these are reasonable values for dry and moist soils respectively.

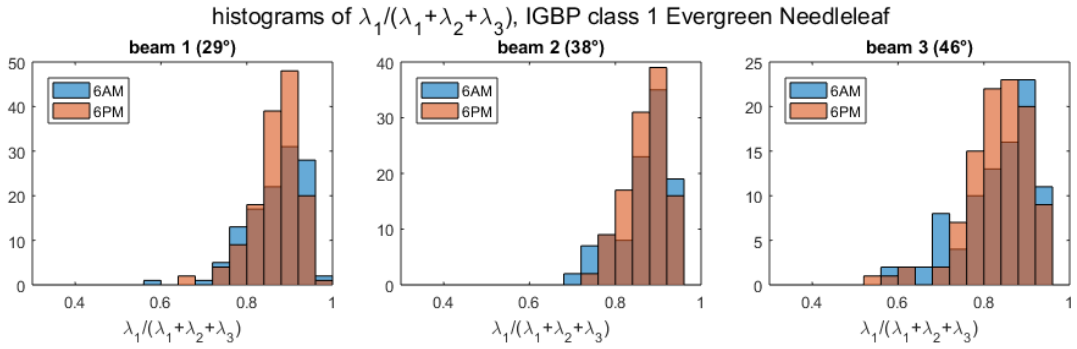

Figure 3.9. Histograms of pixel counts binned by the dominant eigenvalue as a fraction of the trace of the temporal covariance matrices, verifying that the dominant eigenvalue/eigenvector indeed capture most of the temporal variation, and that again there are no obvious differences between the ascending-pass (local 6PM) and descending-pass (local 6AM) data.

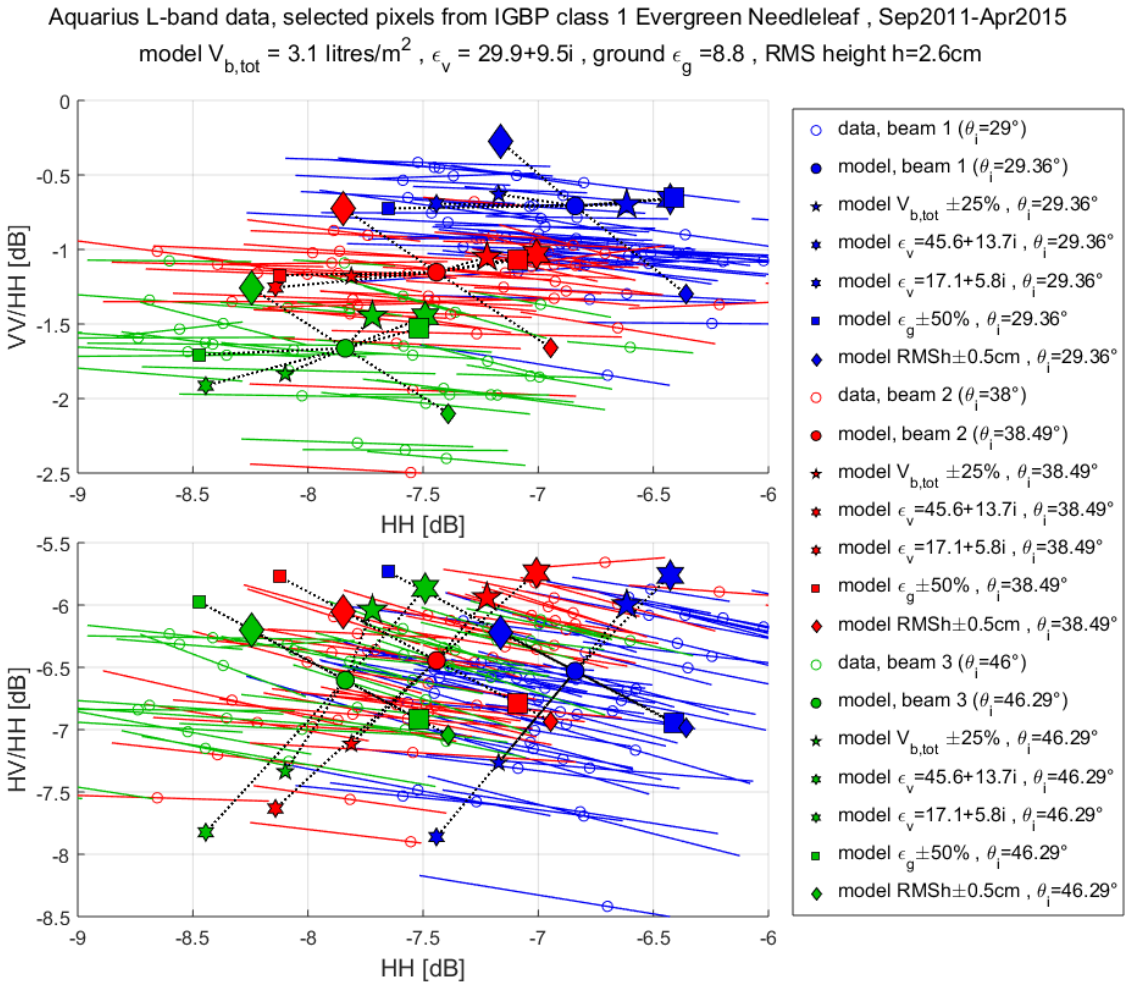

Figure 3.10. Same as Figure 3.8, but also showing temporal variations in the data. Each open circle, corresponding to one of the pixels shown in Figure 3.6 and representing the mean taken over unfrozen periods from Sep2011-Apr2015, now also comes with a line through it that visually summarizes the temporal variation. Specifically, the direction of this line is the dominant eigenvector of the covariance of the data (HH[dB], VV/HH[dB], HV/HH[dB]) for that pixel over the unfrozen period, and the half-length of the line is the square-root of the dominant eigenvalue.

We can take the temporal analysis a step further to see if there is any difference in the patterns of variation for shorter vs longer time scales. For longer time scales, we take the median of the data over two-month periods January-February, March-April, May-June, July-August, September-October, November-December. Data from different years are pooled and included together so long as they fall within the same two months. Variations between different two-month periods would reveal "low-frequency" fluctuations. For shorter time

scales, we implement a simple high-pass filter for the data time-series using a kernel [-0.25, 0.5, -0.25], i.e. if $u[t]$ is the time-series, the filtered time-series is

$$v[t] = -0.25u[t-1] + 0.5u[t] - 0.25u[t+1] \quad (3.2)$$

where the time-step of our data is 1 week, and $u[t]$ represent time-series for our radar features HH[dB], VV/HH[dB], HV/HH[dB] for each pixel. We then take the sample covariance of the filtered high-frequency time-series (arranged as a vector of three features HH[dB], VV/HH[dB], HV/HH[dB], with mean subtracted) computed as a 3x3 matrix (recall that frozen weeks were excluded), and perform eigendecomposition similar to when analyzing the full temporal covariance. The validity of a rank-1 approximation to this "high-frequency" covariance is then checked in Figure 3.11, which reveals, unsurprisingly, that the high-frequency covariance is "noisier" and the rank-1 approximation is poorer than for the full temporal covariance (Figure 3.9), in particular for 6PM data. Nevertheless, we shall still employ the rank-1 approximation in Figure 3.12.

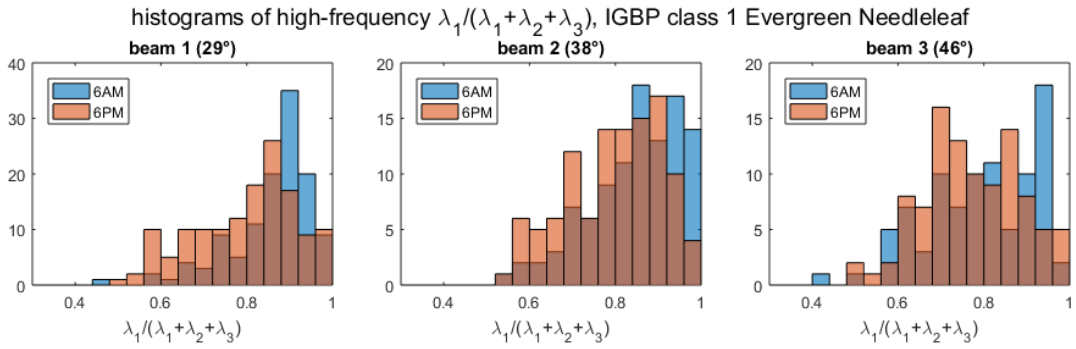

Figure 3.11. Histograms of pixel counts binned by the dominant eigenvalue as a fraction of the trace of the "high-frequency" temporal covariance matrices. The rank-1 approximation is poorer for the "high-frequency" covariance than for the total temporal covariance matrices.

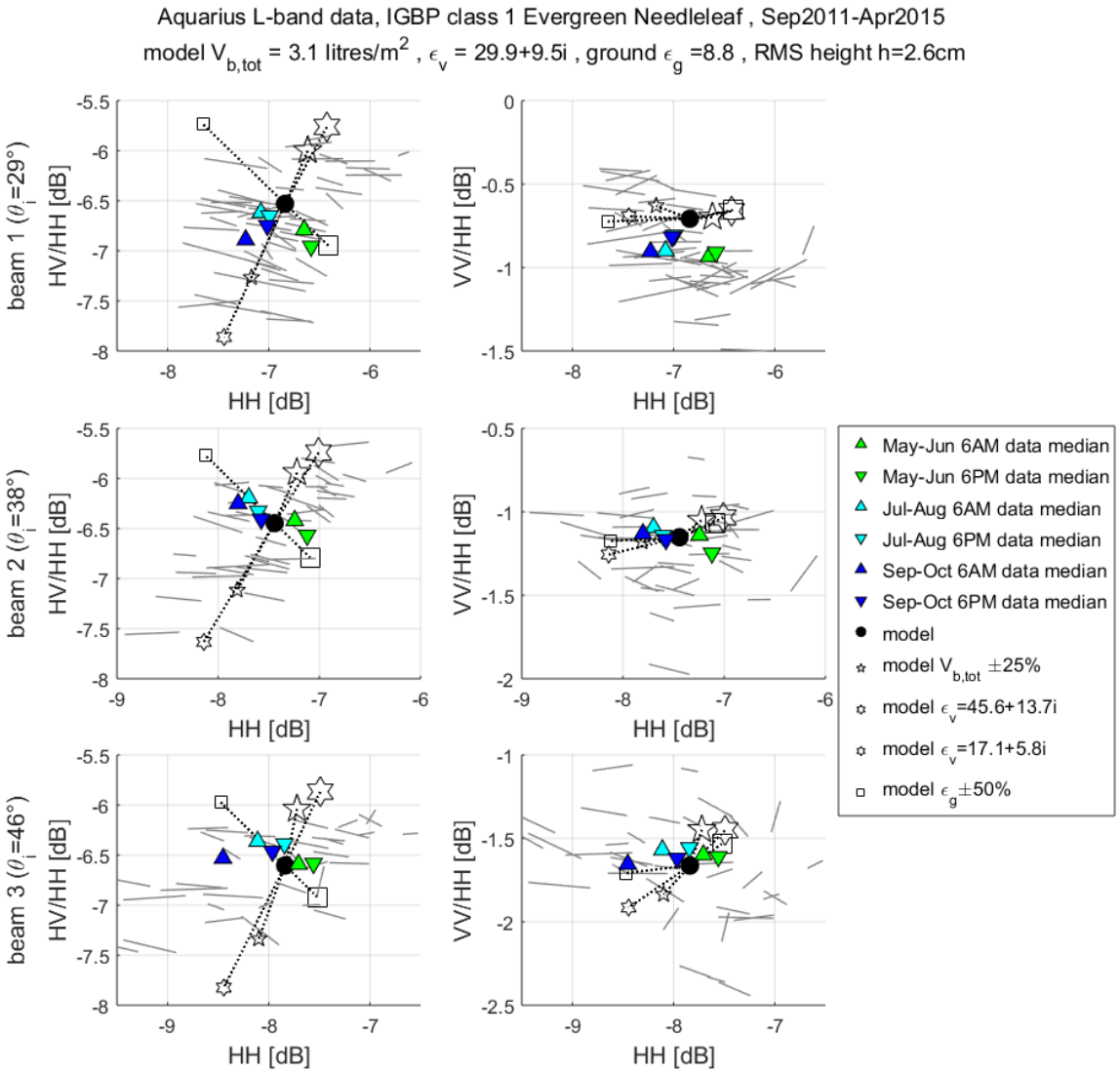

Figure 3.12. Further analysis of temporal variations in radar backscatter. The three beam incidence angles are now plotted in different rows (top: 29°, middle: 38°, bottom: 46°). The left column plots are HV/HH[dB] vs HH[dB], while the right column plots are VV/HH[dB] vs HH[dB]. Filled triangles are the median over all pixels in IGBP class 1, data taken from time periods indicated by color: green (May-Jun), cyan (Jul-Aug), blue (Sep-Oct). Thus these colored triangles display "low-frequency" temporal variations. Upward-pointing triangles are for descending-pass (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM). Solid gray lines visually summarize the "high-frequency" temporal variations of each selected pixel in Figure 3.6. Specifically, the direction of this line is the dominant eigenvector of the covariance of the data (HH[dB], VV/HH[dB], HV/HH[dB]) for that pixel over the unfrozen period and after a high-frequency filter, and the half-length of the line is the square-root of the dominant eigenvalue. Filled circles show the forward model with input parameters as before in Figure 3.7. Other open symbols explore the nearby model parameter space if one parameter is perturbed. Larger and smaller five-

pointed stars represent a change in $V_{b,tot}$ by +25% and -25% respectively. Larger six-pointed stars represent an increase in ε_v to $\varepsilon_v = 45.6 + 13.7i$, while smaller six-pointed stars represent a decrease to $\varepsilon_v = 17.1 + 5.8i$. Larger and smaller squares represent a change ε_g by +50% and -50% respectively. Larger and smaller diamonds represent a change in h by +0.5cm and -0.5cm respectively. Dotted lines merely serve to connect these open symbols with the filled circle.

Figure 3.12 displays these "low-frequency" and "high-frequency" temporal fluctuations. There are six subplots, organized into three rows and two columns. The three beam incidence angles are plotted in different rows (top: 29°, middle: 38°, bottom: 46°). The left column plots are HV/HH[dB] vs HH[dB], while the right column plots are VV/HH[dB] vs HH[dB]. Filled colored triangles display the two-month median data values; these are the "low-frequency" fluctuations. Only May-Jun, Jul-Aug, and Sep-Oct are displayed because most of the November-April period corresponds to frozen conditions. The solid gray lines are similar to the colored lines in Figure 3.10, except that they are for the high-frequency temporal covariance. Specifically, the direction of each gray line displays the dominant eigenvector of the 3x3 "high-frequency" covariance matrix for that pixel, and the half-length of the displayed line is $\sqrt{\lambda_1}$.

From Figure 3.12, we see that both "low-frequency" and "high-frequency" temporal fluctuations seem to be essentially similar in direction on the scatterplot as for the total temporal covariance (Figure 3.10). These suggest that our earlier inference, that temporal fluctuations in radar backscatter might primarily be due to changes in soil moisture, applies across temporal time-scales from 1-2 weeks to several months. Note however that our weekly data time-series is temporal subsampled because soil moisture changes often occur on time-scales shorter than a week or so; effects from aliasing may have impacted our inferences.

As an additional check on our interpretation of the spatial and temporal factors, consider Figure 3.13 and Figure 3.14, which plot the (total) temporal standard deviation of the HH-polarization radar backscatter, against fitted parameters (total volume in branches $V_{b,tot}$ for Figure 3.13 and ground roughness RMS height h for Figure 3.14). The temporal standard deviations were computed over unfrozen periods from Sep2011-Apr2015 per EASE2 grid pixel for all valid data pixels of IGBP class 1 (evergreen needleaf forests). The fitted values

were found for each of these pixels as follows. For Figure 3.13, a value of $V_{b,tot}$ was fitted to each EASE2 grid pixel by minimizing the sum of component-wise absolute distances between the data and model with all other model parameters as found earlier, i.e.

$$\underset{V_{b,tot}}{\text{minimize}} \sum_{i=1}^{3} \left| x_i^{(data\,pix)} - x_i^{(model)}\left(V_{b,tot}\,, \varepsilon_g = 8.8, h = 2.6\text{cm}\right) \right| \qquad (3.3)$$

$$x_1 = \text{HH[dB]}\,, \qquad x_2 = \text{VV/HH[dB]}\,, \qquad x_3 = \text{HV/HH[dB]}\,.$$

To clarify, $x_i^{(data\,pix)}$ is the mean value of radar feature i ($i = 1,2,3$) for that EASE2 pixel over unfrozen periods from Sep2011-Apr2015 for data. Likewise, for Figure 3.14, a value of ground roughness RMS h was fitted to each EASE2 grid pixel by minimizing the sum of component-wise absolute distances between the data and model with all other model parameters as found earlier, i.e.

$$\underset{h}{\text{minimize}} \sum_{i=1}^{3} \left| x_i^{(data\,pix)} - x_i^{(model)}\left(V_{b,tot} = 3.1 \times 10^{-3}\text{m}^3/\text{m}^2, \varepsilon_g = 8.8, h\right) \right|. \quad (3.4)$$

In our actual implementation, the minimization was performed by computing the minimization objective function over a grid of parameter values, and choosing from amongst them the value that gives the smallest objective value. This grid is visible in Figure 3.13 and Figure 3.14.

Figure 3.13 shows an inverse relationship between the fitted values of the total volume in branches $V_{b,tot}$ and the HH temporal standard deviation. This is consistent with fluctuation in ground relative permittivity ε_g (i.e. soil moisture) being an important cause of HH temporal variations. For forests, vegetation is significantly thick, larger $V_{b,tot}$ corresponds to greater canopy optical thickness $\tau_{cn}(\theta_i)$, and thus less sensitivity to the ground. (On the other hand, in the sparse vegetation regime where optical thickness is small, larger $V_{b,tot}$ strengthens the double-bounce scattering mechanisms and could enhance radar sensitivity to the ground.) The dependence on incidence angle is less intuitive because it is a subtle parameter-dependent interplay between the extinction, the direct backscatter from the

ground, the Fresnel reflection coefficients, and the angle-dependence of bistatic scattering coefficients off the cylinders. For these parameters here, the forward model predicts a slight decrease in HH temporal standard deviation with incidence angle increasing from 29° to 46°; there is but a tenuous suggestion of such a trend in the scatterplot.

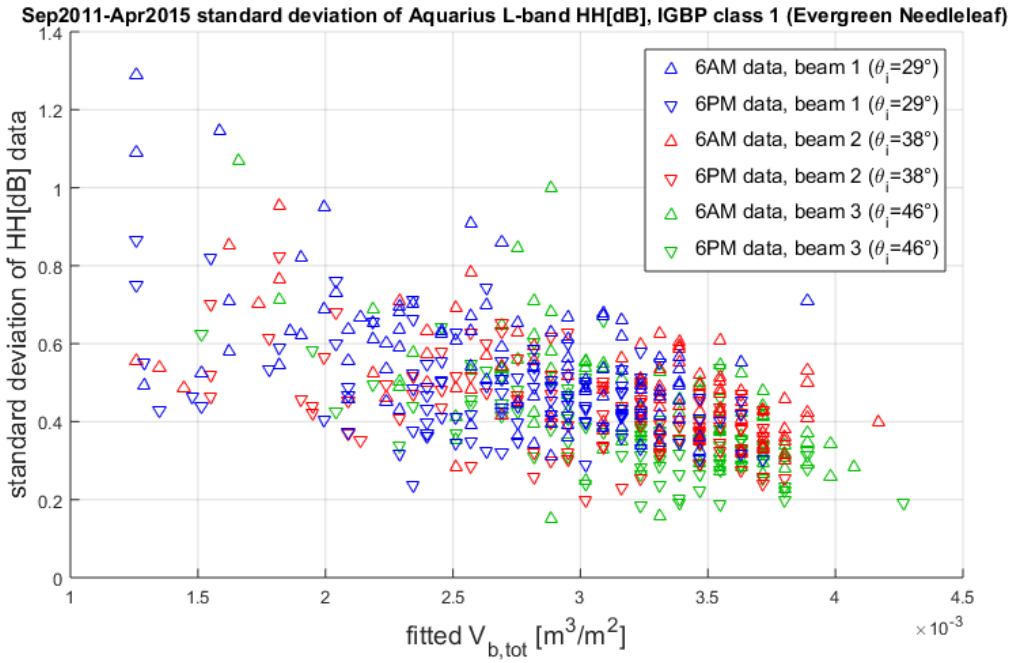

Figure 3.13. Temporal standard deviation over unfrozen periods from Sep2011-Apr2015 of Aquarius L-band HH[dB] data, vs. fitted values of $V_{b,tot}$. Each open triangle represents one EASE2 grid pixel of IGBP class 1 (evergreen needleleaf forests). While keeping other forward model parameters as before in Figure 3.8 ($m = 1$ and $\theta_0 = \pi/2$ in equation (2.6), vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, ground relative permittivity $\varepsilon_g = 8.8$, ground roughness RMS height $h = 2.6$cm), the total volume in branches $V_{b,tot}$ was now allowed to vary and fitted for each pixel; these are the values on the horizontal axis. Blue, red, and green triangles representing the three different beam incidence angles (29°, 38°, 46° respectively) as usual. Upward-pointing triangles are for descending-pass data (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM).

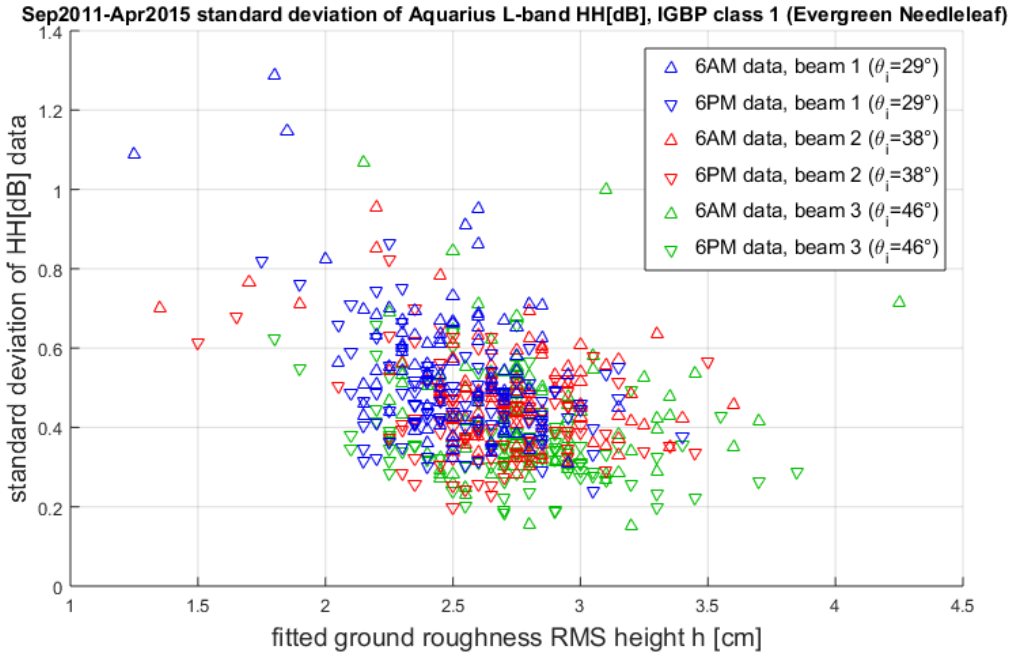

Figure 3.14. Temporal standard deviation over unfrozen periods from Sep2011-Apr2015 of Aquarius L-band HH[dB] data, vs. fitted values of ground surface roughness RMS height h. Each open triangle represents one EASE2 grid pixel of IGBP class 1 (evergreen needleleaf forests). While keeping other forward model parameters as before in Figure 3.8 ($m = 1$ and $\theta_0 = \pi/2$ in equation (2.6), vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, total volume in branches $V_{b,tot} = 3.1 \times 10^{-3} \mathrm{m}^3/\mathrm{m}^2$, ground relative permittivity $\varepsilon_g = 8.8$), the ground surface roughness RMS height h was now allowed to vary and fitted for each pixel; these are the values on the horizontal axis. Blue, red, and green triangles representing the three different beam incidence angles (29°, 38°, 46° respectively) as usual. Upward-pointing triangles are for descending-pass data (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM).

Figure 3.14 also suggests an inverse relationship between the fitted ground roughness parameter h and the HH temporal standard deviation, despite significant scatter. This is also consistent with soil moisture being a driver of HH temporal variations, and further that this effect is seen by the radar primarily through double-reflection mechanisms. Coherent reflection off a rougher ground surface is weaker (equation 2.15), so the impact by variations in ground relative permittivity ε_g on ground-trunk and ground-canopy double-reflection scattering mechanisms is also weaker for rougher ground surfaces. However, if the impact of ε_g fluctuations is primarily through direct backscatter from the ground instead (e.g. in the regime when the incidence angle is small, ground roughness h is high, and the vegetation is

thin), then rougher ground surfaces instead strengthen the direct backscatter from the ground, and we should see a positive relationship between h and HH temporal standard deviation.

Observe from both Figure 3.13 and Figure 3.14 that there is also a slight correlation of the fitted parameters with incidence angle. This correlation should not exist, and is probably an artefact of model deficiencies and overfitting (because for each pixel, one parameter is being fit to only three data numbers).

Let us also look in detail at the time series data from an example pixel. We shall use Aquarius backscatter data corresponding to EASE2 grid row 66, column 155; the IGBP land cover classification for this pixel is class 1 (evergreen needleleaf forests). Within this pixel is SNOTEL site #344 (Billie Creek Divide, 42.4°N, 122.27°W, elevation 1609m), where soil moisture and precipitation were measured. We obtained the data from the National Resources Conservation Service [86] of the US Department of Agriculture and the International Soil Moisture Network [87, 88]. Data from October 2011-Dec2014 are plotted in Figure 3.15. As can be seen from Figure 3.15, there is some correlation between HH radar backscatter and soil moisture – we see decreases in HH backscatter accompanying drying-downs in the soil moisture, and spikes in HH backscatter accompanying spikes in soil moisture brought by precipitation. There is generally an annual pattern of soil moisture starting high in the spring and decreasing into the summer, likely associated with snowmelt, observed not only at this ground station but at many other sites as well (not displayed); this is likely a widespread phenomenon in the boreal/hemiboreal/temperate forests in the Northern hemisphere as seen from the Aquarius data.

However, we can also see from Figure 3.15 that the correlations between radar backscatter and soil moisture is not always consistent, and there are some fluctuations in HH backscatter that do not accompany variations in soil moisture, and vice-versa. Further, according to our model, HV/HH should be negatively correlated with HH and soil moisture if all other factors are held constant; from the data, we certainly see that not all fluctuations in HV/HH are negatively correlated with HH. These remind us that though soil moisture is important, other factors affecting backscatter variations cannot be neglected.

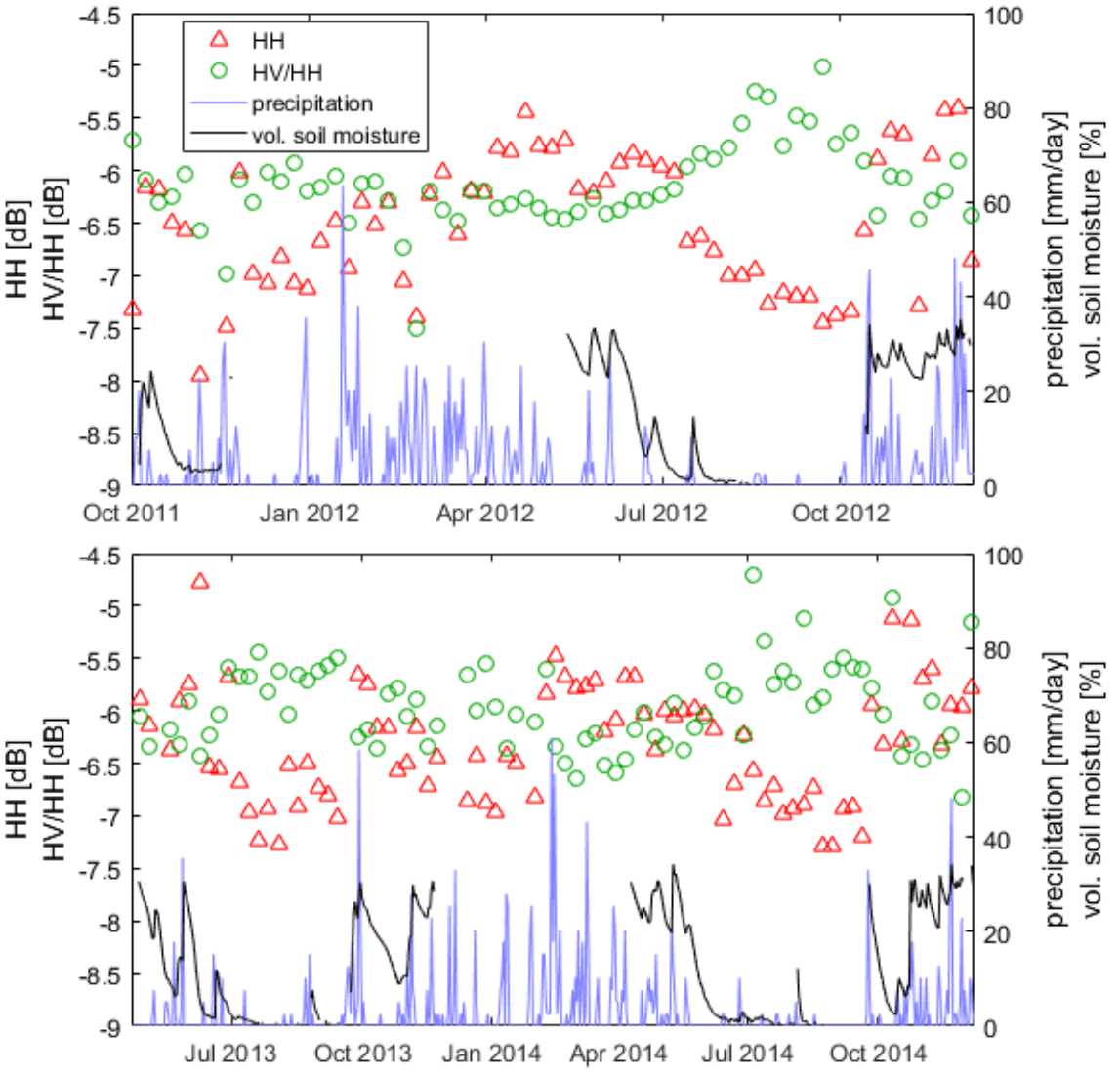

Figure 3.15. Aquarius backscatter data and ground measurements at SNOTEL site #344. HH backscatter is plotted with red triangles, HV/HH backscatter is plotted with green circles, both are displayed in dB units and share the axis scale on the left. Precipitation in mm/day is plotted with a solid blue line, volumetric soil moisture in percentage (cm^3/cm^3) is plotted with a solid black line, and both share the axis scale on the right.

Let us summarize the comparison in this section between model and data for IGBP class 1 (evergreen needleleaf forests) areas, taken at a coarse global spatial scale. Spatial variations in the L-band radar backscatter are likely due to a combination of variations in both vegetation and ground factors, in particular ground surface roughness. Changes in ground

dielectric relative permittivity, which is directly related to changes in soil moisture, are an important driver of temporal variations (as observed from the weekly samples during unfrozen periods) in the L-band radar backscatter, though again not the sole cause. Diurnal differences, though possibly present, are not conclusive from the Aquarius dataset alone. These findings and interpretations were based on analysis of the Aquarius multi-polarization scatterometer data, and using the forward model of Chapter 2.

3.4 Evergreen broadleaf forests (IGBP class 2)

Using the MODIS land cover map, the geographic distribution of the evergreen broadleaf forests (IGBP class 2) is shown in Figure 3.5 and Figure 3.16 for emphasis. Most of these pixels come from the Amazon basin and the Congo basin. A random 10% selection of the EASE2 grid pixels is displayed in Figure 3.16. Land area for evergreen broadleaf forests from Southeast Asia are slightly underrepresented compared to the Amazon or Congo areas for several reasons: greater amounts of RFI corruption, mountainous areas with high terrain slope, and coastal areas with the presence of water within the large radar footprints.

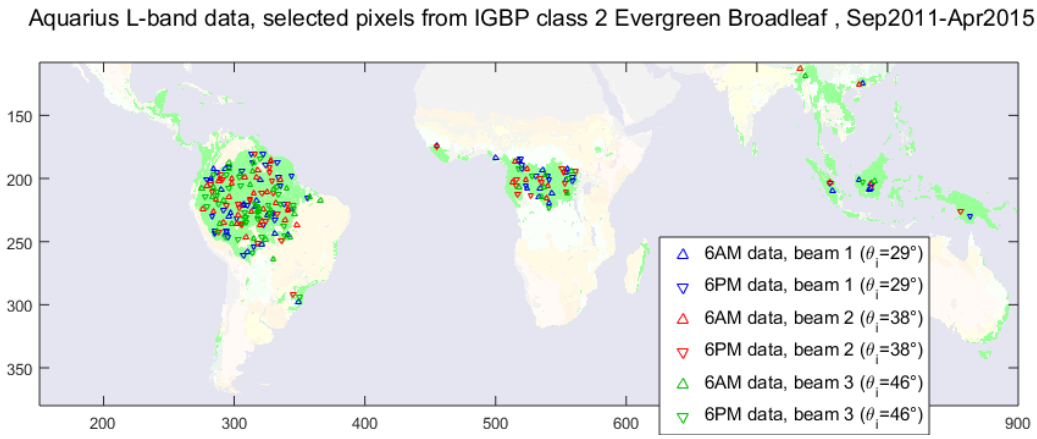

Figure 3.16. Geographic distribution of evergreen broadleaf forests (IGBP class 2) and randomly selected 10% of filtered Aquarius data EASE2 grid pixels, marked by open triangle symbols.

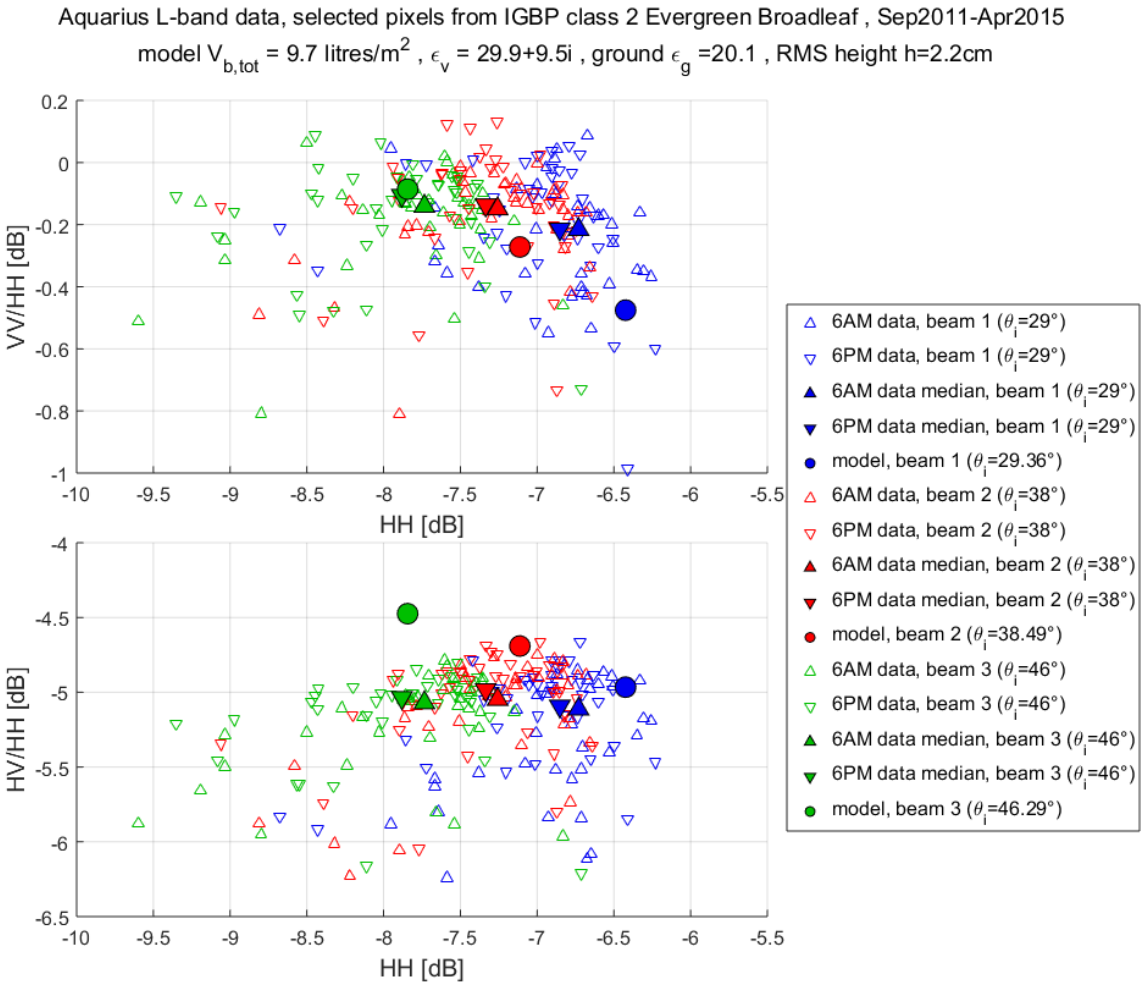

Figure 3.17. Aquarius L-band scatterometer data for EASE2 grid pixels classified as evergreen broadleaf forests (IGBP class 2). The three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°). Upward-pointing triangles are for descending-pass data (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM). Each open triangle represents data from one of the pixels shown in Figure 3.16, taking the mean over Sep2011-Apr2015 for each pixel. Filled triangles are the median over open triangles, but with median taken over all data pixels of this class instead of only the 10% displayed. Filled circles show the forest forward model from Chapter 2, with parameters $m = 0$ (uniformly random cylinder orientation distribution) in equation (2.6), total volume in branches $V_{b,tot} = 9.7 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, ground relative permittivity $\varepsilon_g = 20.1$, ground roughness RMS height $h = 2.2\mathrm{cm}$.

The scatterometer data for the selected pixels is displayed in Figure 3.17, which depicts two scatterplots: one of VV/HH[dB] vs HH[dB], and another of HV/HH[dB] vs HH[dB], similar to Figure 3.7 except for a change in IGBP class. The symbol descriptions are in the figure

caption. Each open triangle represents the mean value (of the plotted quantities) of data from Sep2011-Apr2015 from its corresponding pixel shown in Figure 3.16. Filled circles show the results from the forest forward model from Chapter 2. The parameter $m = 0$ (i.e. uniformly random orientation distribution of cylinders ; θ_0 does not matter in this case) in equation (2.6) was used, and the vegetation cylinder relative permittivity chosen as $\varepsilon_v = 29.9 + 9.5i$ as usual. Other model input parameters are: total volume in branches $V_{b,tot}$, ground relative permittivity ε_g, and ground roughness RMS height h. From Saatchi et al. [7], the above-ground biomass for evergreen broadleaf forests is about $M_{tot} = 300$ tons/ha. Based on work by Segura and Kanninen [89], Brown and Lugo [90], and Nogueira et al. [91], a value of ρ=0.62g/cm^3 was used as the value for the wood specific gravity (ratio of "oven-dry weight" to "green volume"). Combined with our earlier estimate of 0.2 for the fraction $V_{b,tot}/(V_{b,tot} + V_{trk,tot})$ and taking $M_{tot} = \rho(V_{b,tot} + V_{trk,tot})$, these give an estimated $V_{b,tot} = 9.7 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$. For the ground relative permittivity, we choose $\varepsilon_g = 20.1$, with reference to work by Wang et al. [92]. The only input parameter chosen using the Aquarius data was the ground RMS height $h = 2.2\mathrm{cm}$, chosen to give a reasonable fit to the pooled (ascending and descending) data. The fitting procedure minimized a function similar to (3.1) and (3.4) :

$$\underset{h}{\text{minimize}} \sum_{data\ pix} \sum_{i=1}^{3} \left| x_i^{(data\ pix)} - x_i^{(model)}\left(V_{b,tot} = 9.7 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2, \varepsilon_g = 20.1, h\right) \right|$$

$$x_1 = \mathrm{HH[dB]}\,, \qquad x_2 = \mathrm{VV/HH[dB]}\,, \qquad x_3 = \mathrm{HV/HH[dB]}\,.$$

From Figure 3.17, it can be seen that the forward model somewhat agrees with the data, but not as well as for IGBP class 1 (evergreen needleleaf forests). Part of this may be due to the fact that only one parameter was fitted here, while three parameters were fitted for evergreen needleleaf forests, but a more likely explanation is that the model is simply less appropriate in this case. Recall that in keeping the overall number of parameters small, only cylinders (corresponding to branches) in the canopy were modelled, and there was no consideration of leaves. This is likely to be less valid in evergreen broadleaf forests.

From Figure 3.17, there might be a tiny (0.1dB) difference overall between the ascending-pass (local time 6PM) data and descending-pass (local time 6AM) data. Various authors [93, 94, 95] have reported diurnal differences (higher at sunrise) in radar backscatter from the Amazon using shorter wavelength (Ku band) scatterometers; van Emmerik et al. [93] provided evidence that this is driven by vegetation water stress. As noted earlier in Section 3.3, due to other possible confounding factors such as differences in footprints and azimuth angles, it is thus difficult to be conclusive based on this 0.1dB median difference.

We then attempt to use the forward model to analyze the spatial and temporal variations in the radar backscatter data for IGBP class 2 (evergreen broadleaf forests). Displayed in Figure 3.18 are the same data points as Figure 3.17, now with the descending-pass (local time 6AM) and ascending-pass (local time 6PM) data pooled and all marked with open circles instead of triangles. Overlaid on the data are filled symbols showing the results from the forward model. The filled circles show the forward model with parameters as before. Other filled symbols explore the nearby model parameter space if one parameter is perturbed. The symbols are explained in the caption, and are similar to those used in Figure 3.10. Each line through each open circle represents the dominant eigenvector of the covariance of the data (HH[dB], VV/HH[dB], HV/HH[dB]) for that pixel, and the half-length of each line is the square-root of the dominant eigenvalue. These lines visually summarize the temporal variations (sampled weekly over Sep2011-Apr2015). Figure 3.19 plots histograms of pixel counts binned by the fraction $\lambda_1 / (\lambda_1 + \lambda_2 + \lambda_3)$; like in Figure 3.9, it serves to verify the dominance of the first eigenvalue/eigenvector and the validity of the rank-1 approximation. Unlike in Figure 3.9, however, here there is some hint of differences between the ascending-pass (local 6PM) and descending-pass (local 6AM) at larger incidence angles. For the ascending-pass (local 6PM) data, there is slightly more temporal variation not captured by the dominant eigenvalue/eigenvector.

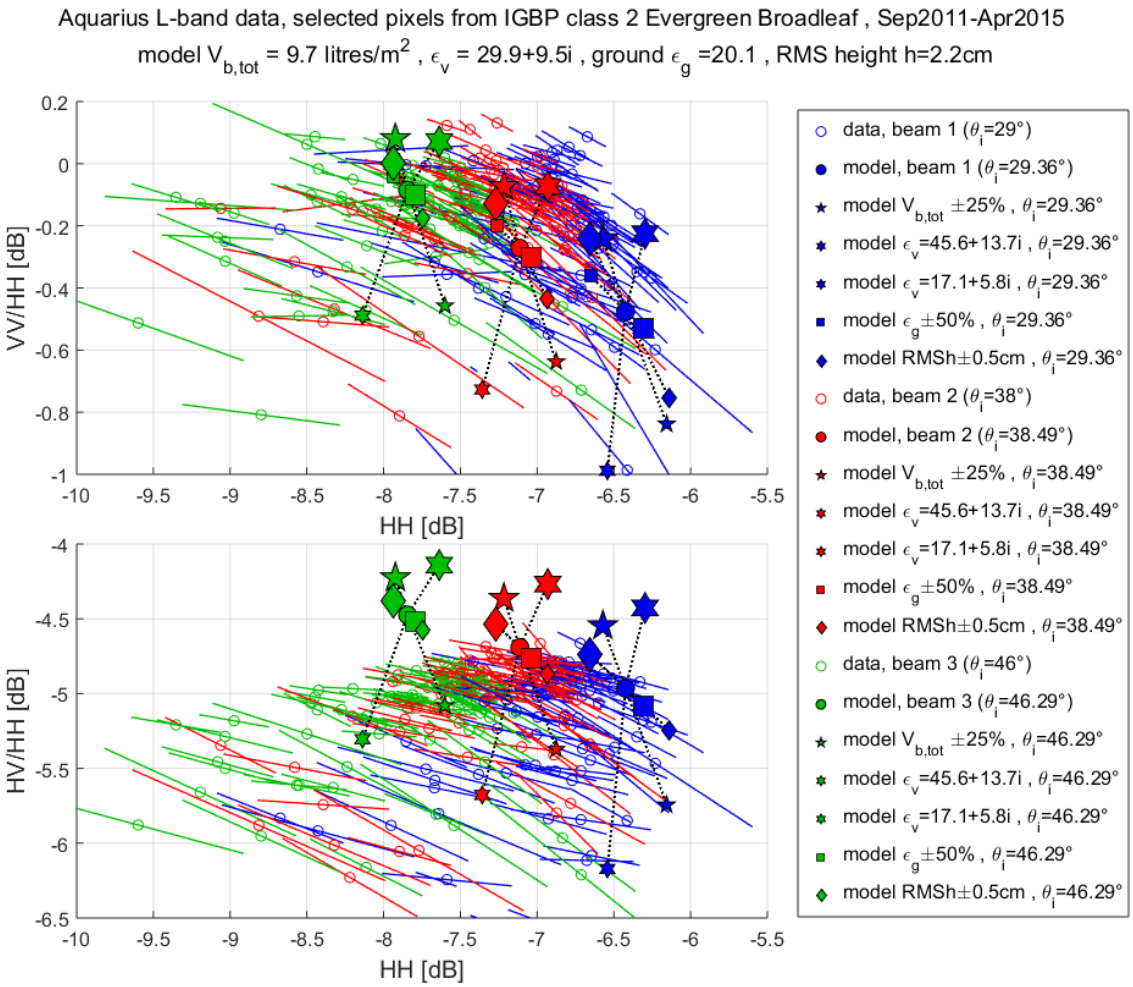

Figure 3.18. Same as Figure 3.17, but with no distinction made between descending- and ascending-pass data; all open triangles are replaced with open circles. Each open circle now also comes with a line through it that visually summarizes the temporal variation. The direction of this line is the dominant eigenvector of the covariance of the data (HH[dB], VV/HH[dB], HV/HH[dB]) for that pixel, and the half-length of the line is the square-root of the dominant eigenvalue. Filled circles show the forward model with parameters as before in Figure 3.17. Other filled symbols explore the nearby model parameter space if one parameter is perturbed. Larger and smaller five-pointed stars represent a change in $V_{b,tot}$ by +25% and -25%, respectively. Larger and smaller six-pointed stars represent a change in ε_v to $\varepsilon_v = 45.6 + 13.7i$ and $\varepsilon_v = 17.1 + 5.8i$, respectively. Larger and smaller squares represent a change ε_g by +50% and -50%, respectively. Larger and smaller diamonds represent a change in h by +0.5cm and -0.5cm, respectively.

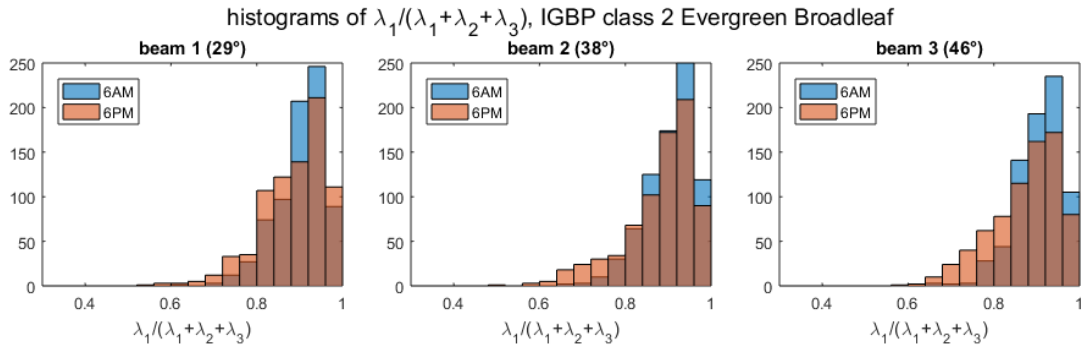

Figure 3.19. Histograms of pixel counts binned by the dominant eigenvalue as a fraction of the trace of the temporal covariance matrices, verifying that the dominant eigenvalue/eigenvector indeed capture most of the temporal variation. All eligible pixels of IGBP class 2 (evergreen broadleaf), not just the 10% random subsample, were included.

We see from Figure 3.18 that the spatial and temporal variations are not readily interpretable using the model. The first issue is that in this saturation regime of parameter space, some effects of changing ground vs. vegetation parameters are confounded and not easily separable. Only changes in the relative permittivity of the cylinders give changes that are somewhat in a different direction on the scatterplots, while changes in the other parameters are difficult to distinguish from one another. There does not appear to be a single physical factor that provides close agreement with either the spatial variations (visualized by the distribution of open circles) or the temporal variations (visualized by the lines through the open circles). This may be in part due to the afore-mentioned deficiencies in the model, or may indeed be a reflection of the lack of dominance of a single physical factor responsible for the changes in radar backscatter.

Our model is not entirely without merit, being able to explain the variations in polarimatric signature due to flooded/nonflooded conditions underneath the canopy in parts of the Amazon basin. It is known that some low-lying parts of the Amazon basin experience seasonal flooding due to the rainy season. This flooding is observable by radar [67] and measurable by interferometry [66]. Figure 3.20 shows the topography of the Amazon basin. The white box marks the region bounded by latitudes 2.5°S to 4.8°S and longitudes 57.9°W to 64.2°W. We compare Aquarius L-band scatterometer observations within this box with predictions of our forward model in Figure 3.21. Figure 3.21 plots scatterplots (left:

HV/HH[dB] vs HH[dB], right: VV/HH[dB] vs HH[dB]) similar to Figure 3.17, except that only the median values of pixels within the white box of Figure 3.20 are displayed, instead of the individual pixels. The three beam incidence angles are plotted separately (top: 29°, middle: 38°, bottom: 46°) and colors are now used to represent the month of year, as in the figure legend. Black filled circles show the forward model with parameters as before in Figure 3.17 and Figure 3.18, while open circles depict a simultaneous change in both ground parameters to $\varepsilon_g = 80$ and $h = 0\text{cm}$, representing a smooth flooded surface. While the forward model (black filled circles) is not entirely accurate as discussed earlier, the variations in polarimatric signature predicted, i.e. direction of change on the scatterplot, are in good agreement with the data, with both HV/HH[dB] and VV/HH[dB] changing inversely with HH[dB]. This is considering that in reality, the vegetation may also contribute a secondary component to the seasonal changes in backscatter. The magnitudes of change of radar backscatter observed, i.e. around 0.6dB for HH at 46° incidence angle and 1.2dB for HH at 29° incidence angle, are expectedly no greater than the model predictions (0.6dB at 46° incidence angle, 1.6dB at 29° incidence angle); parts of the radar footprints may not be within the inundation zones. The essence of our results are similar to modeling efforts by Wang et al. [92].

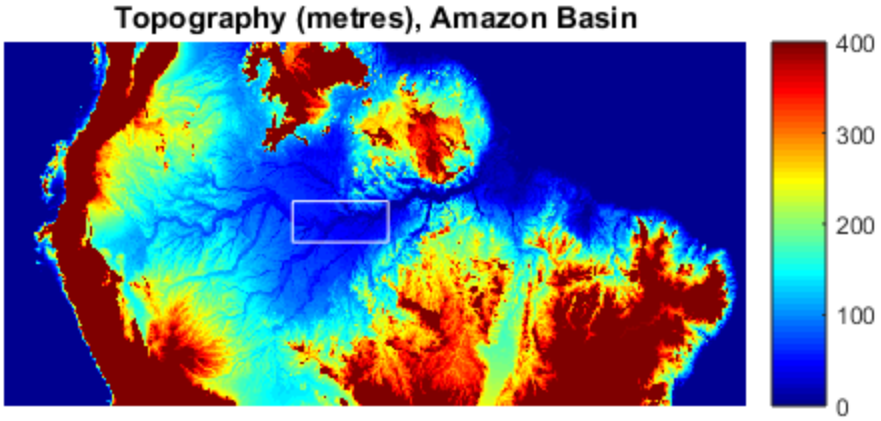

Figure 3.20. Topography of the Amazon Basin. The white box marks the region bounded by latitudes 2.5°S to 4.8°S, longitudes 57.9°W to 64.2°W. Many parts of this region are susceptible to seasonal flooding. Temporal variations in the data from this region are analyzed in Figure 3.21.

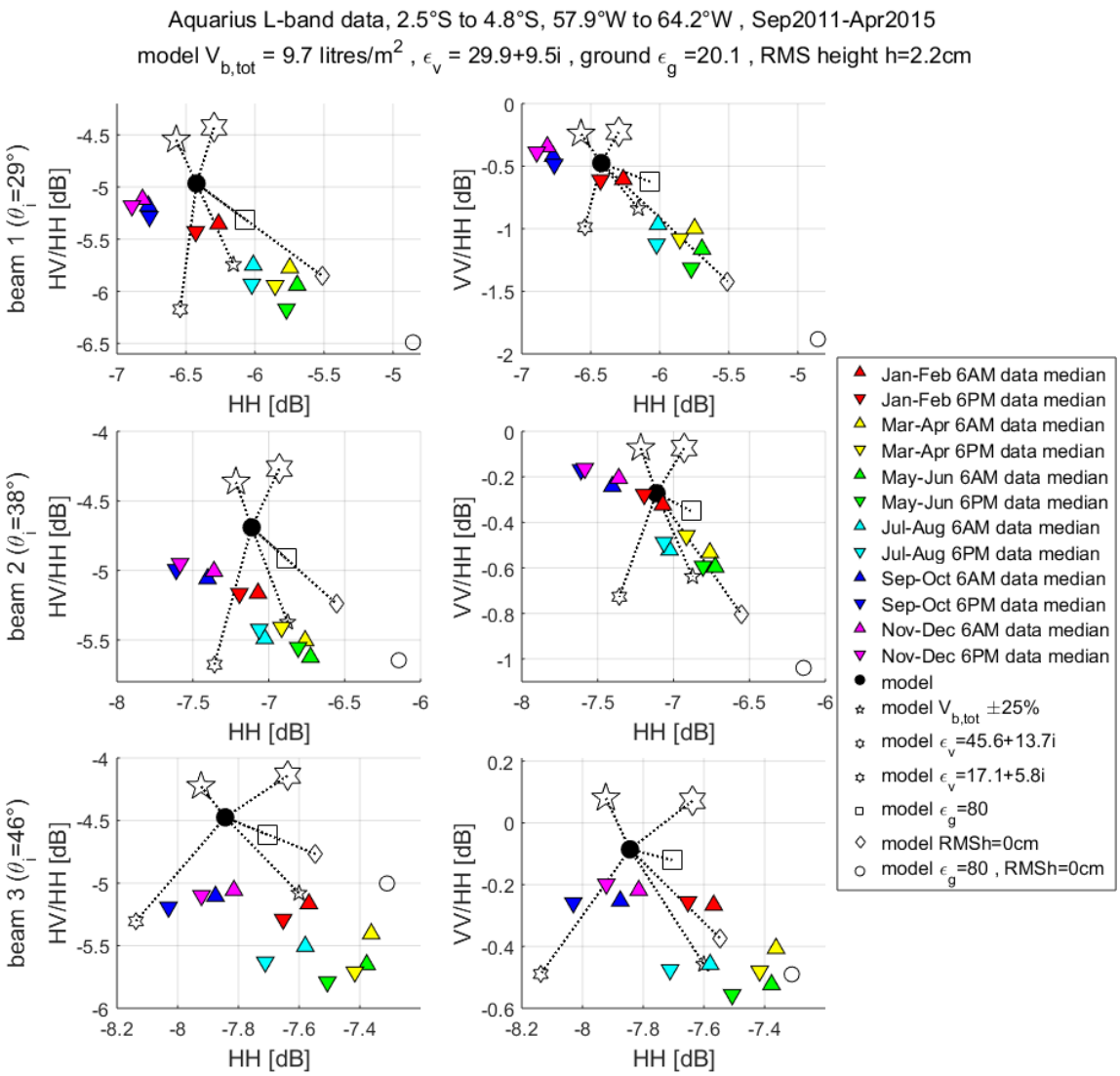

Figure 3.21. Temporal variations in the data from the region bounded by the white box in Figure 3.20, many parts of which are susceptible to seasonal flooding. Colored triangles are median values from the pixels within the box, for some part of each year (red: Jan-Feb; yellow: Mar-Apr; green: May-Jun; cyan: Jul-Aug; blue: Sep-Oct; magenta: Nov-Dec). Upward-pointing and downward-pointing triangles are for descending-pass (local time 6AM) and ascending-pass data (local time 6PM), respectively. Black filled circles show the forward model with parameters as before in Figure 3.17 and Figure 3.18. Black open symbols explore the nearby model parameter space if one parameter is perturbed. Larger and smaller five-pointed stars represent a change in total branch volume $V_{b,tot}$ by +25% and -25%, respectively. Larger and smaller six-pointed stars represent a change in vegetation cylinder relative permittivity ε_v to $\varepsilon_v = 45.6 + 13.7i$ and $\varepsilon_v = 17.1 + 5.8i$, respectively. Open squares depict ground relative permittivity $\varepsilon_g = 80$ (relative permittivity of water). Open diamonds depict ground roughness RMS height $h = 0$cm (perfectly smooth surface). Open circles depict a simultaneous change in both ground parameters to $\varepsilon_g = 80$ and $h = 0$cm, representing a smooth flooded surface.

In summary, the agreement between the model and data for IGBP class 2 (evergreen broadleaf forests) areas is significantly poorer than for IGBP class 1, but still not too far off quantitatively (within about 1dB). Spatial and temporal variations in the L-band radar backscatter were not readily interpretable using the model. However for the case of subcanopy flooding, the model did quantitatively predict changes in multi-polarization backscatter that were in agreement with the data observations.

3.5 Deciduous needleleaf forests (IGBP class 3)

Using the MODIS land cover map, the geographic distribution of the deciduous needleleaf forests (IGBP class 3) is shown in Figure 3.22. Much of this is the Eastern Siberian taiga: forests of primarily Dahurian and Siberian larches, with smaller areas of pines and other species. Figure 3.22 displays a random 20% selection of the EASE2 grid pixels.

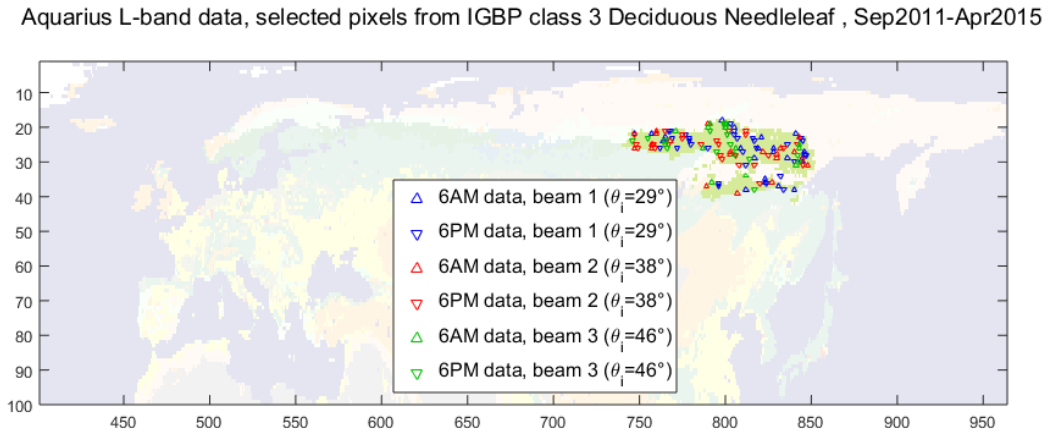

Figure 3.22. Geographic distribution of deciduous needleleaf forests (IGBP class 3) and randomly selected 20% of filtered Aquarius data EASE2 grid pixels, marked by open triangle symbols.

The scatterometer data for the selected pixels is displayed in Figure 3.23, which depicts two scatterplots: one of VV/HH[dB] vs HH[dB], and another of HV/HH[dB] vs HH[dB], similar to Figure 3.17 except for a change in IGBP class. The symbol descriptions are in the figure caption. Filled circles show the results from the forest forward model from Chapter 2. The parameters $m = 1$ and $\theta_0 = \pi/2$ (cosine-squared orientation distribution about the horizontal) in equation (2.6) were used, and the vegetation cylinder relative permittivity chosen as $\varepsilon_v = 29.9 + 9.5i$ as usual. Other model input parameters are: total volume in

branches $V_{b,tot} = 2.8 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, ground relative permittivity $\varepsilon_g = 6.2$, ground roughness RMS height $h = 2.1\mathrm{cm}$. These latter three parameters ($V_{b,tot}$, ε_g, h) were chosen by fitting to the pooled (ascending and descending) data, in the same way as in equation (3.1).

From Figure 3.23, it can be seen that the forward model gives reasonable overall agreement to the data, similar to the case for IGBP class 1 (evergreen needleleaf forests), albeit using fitted parameters. The fitted ground parameters fall within their expected ranges. The ground relative permittivity is lower than for IGBP class 1, which is in reasonable expectation considering the drier climate of Eastern Siberia. Following the same estimation procedure as for IGBP class 1 earlier, $V_{b,tot} = 2.8$ litres/m^2 corresponds to 63tons/ha of aboveground biomass. Referring to the study by Shepashenko et al. [84] for Siberian forest phytomass using Russian state forest inventories and regression analysis, Fig. 1 in that paper displays an estimate of around 3-4kgC/m^2 carbon for much of the corresponding region in Eastern Siberia. Using their conversion factor of 0.5 for phytomass to carbon content, this translates to 60-80 tons/ha. Bearing in mind that roots were including in their estimate, the agreement is close.

With regards to the variation between the descending-pass (local time 6AM) and ascending-pass (local time 6PM) data, there may be a slight difference, but it is small (compared to the spatial variation between pixels, and longer period temporal variations to be discussed later) when considered in an average sense over time and multiple pixels. Thus we consider it inconclusive and again pool the 6AM and 6PM data for now.

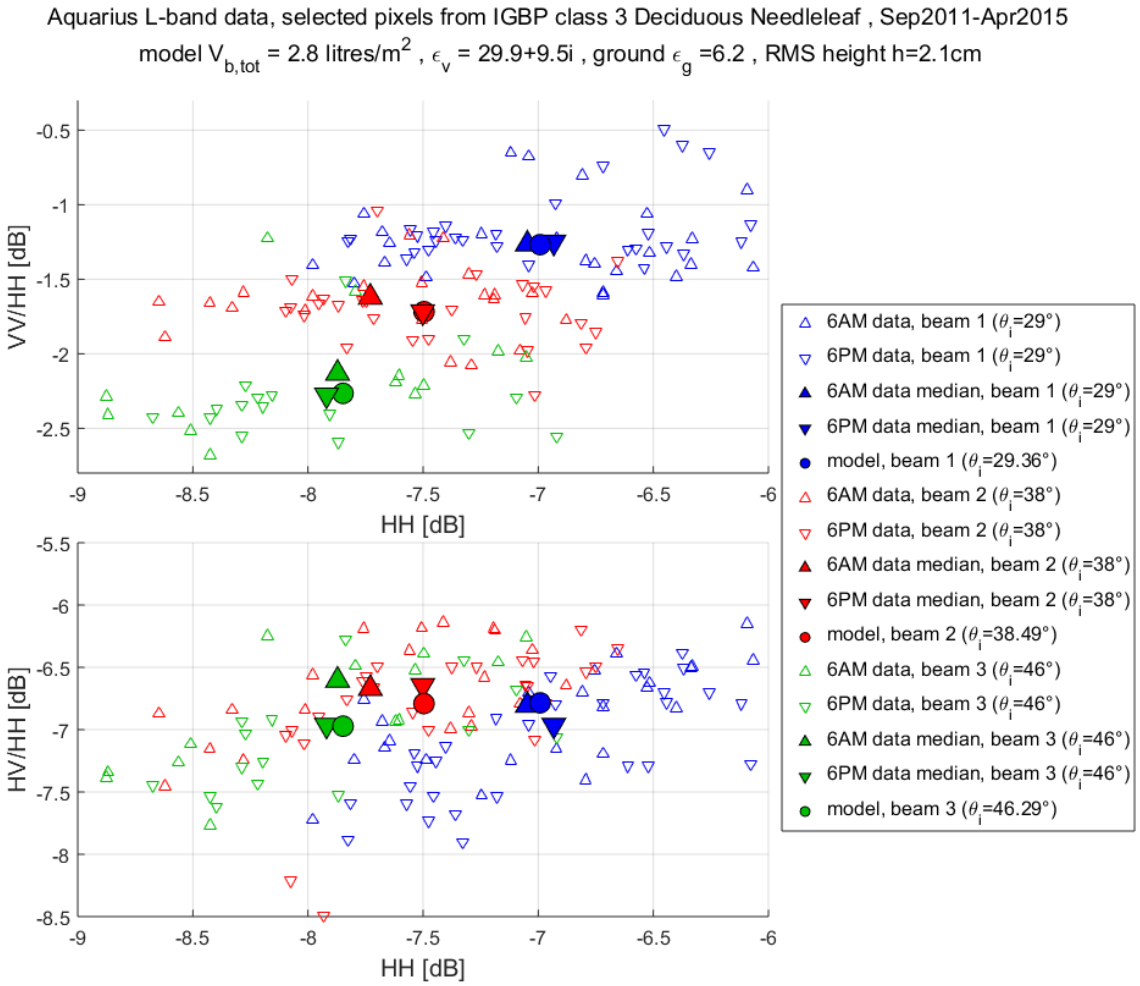

Figure 3.23. Aquarius L-band scatterometer data for EASE2 grid pixels classified as deciduous needleleaf forests (IGBP class 3). The three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°). Upward-pointing triangles are for descending-pass data (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM). Each open triangle represents data from one of the pixels shown in Figure 3.22, taking the mean over unfrozen periods from Sep2011-Apr2015 for each pixel. Filled triangles are the median over open triangles, but with median taken over all data pixels of this class instead of only the 20% displayed. Filled circles show the forest forward model from Chapter 2, with parameters $m = 1$ and $\theta_0 = \pi/2$ (preferential horizontal orientation) in equation (2.6), total volume in branches $V_{b,tot} = 2.8 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, ground relative permittivity $\varepsilon_g = 6.2$, ground roughness RMS height $h = 2.1\mathrm{cm}$.

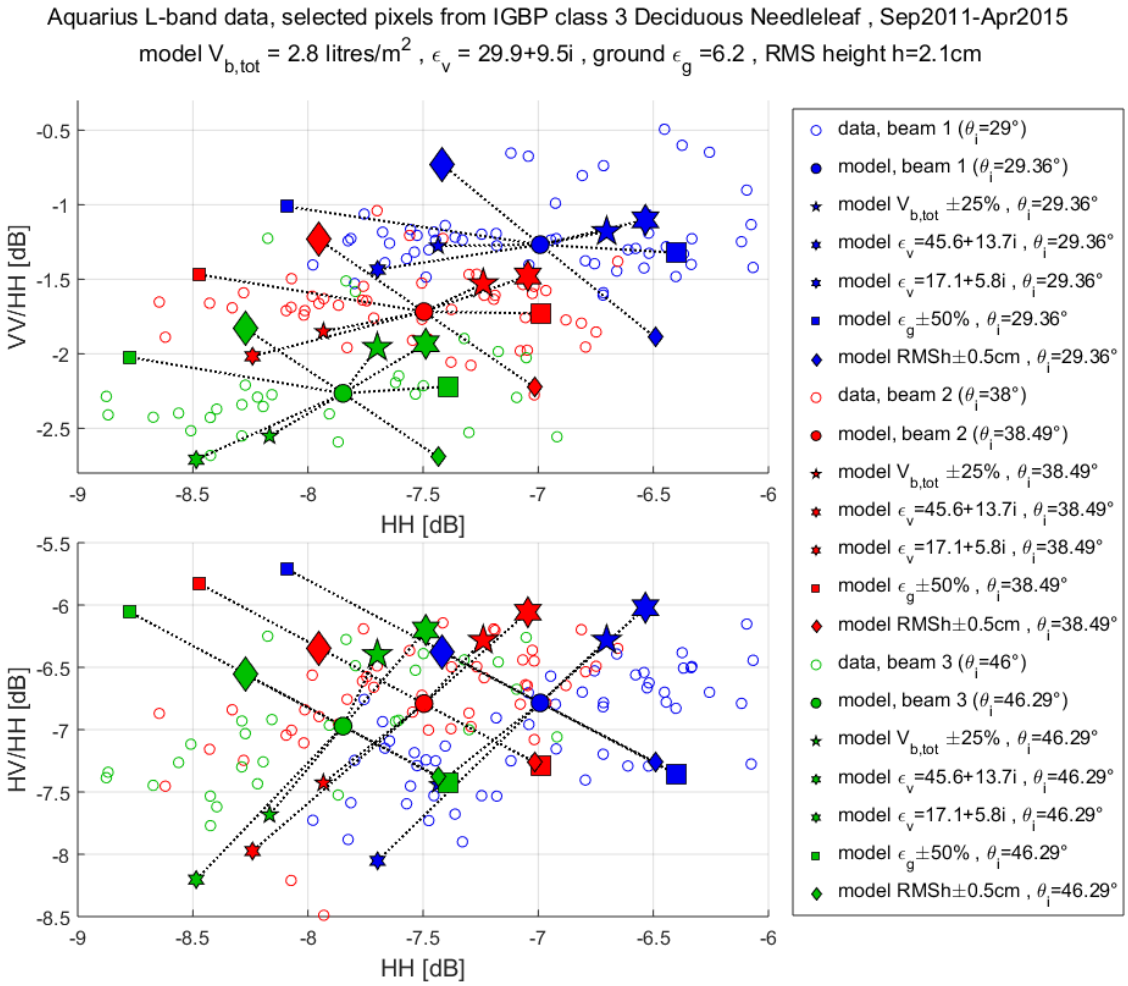

Figure 3.24. Same as Figure 3.23, but with no distinction made between descending- and ascending-pass data; all are marked with open circles. Filled circles show the forward model with parameters as before in Figure 3.23 (preferential horizontal orientation distribution of cylinders with $m = 1$ and $\theta_0 = \pi/2$ in equation (2.6), total volume in branches $V_{b,tot} = 2.8 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, ground relative permittivity $\varepsilon_g = 6.2$, ground roughness RMS height $h = 2.1\mathrm{cm}$). Other filled symbols explore the nearby model parameter space if one parameter is perturbed. Larger and smaller five-pointed stars represent a change in $V_{b,tot}$ by +25% and -25%, respectively. Larger six-pointed stars represent an increase in ε_v to $\varepsilon_v = 45.6 + 13.7i$, while smaller six-pointed stars represent a decrease to $\varepsilon_v = 17.1 + 5.8i$. Larger and smaller squares represent a change ε_g by +50% and -50%, respectively. Larger and smaller diamonds represent a change in h by +0.5cm and -0.5cm, respectively.

The model is then used to analyze the spatial variation in the radar backscatter data across different pixels. Figure 3.24 displays the same data as Figure 3.23, now with the descending-pass (local time 6AM) and ascending-pass (local time 6PM) data pooled and all marked with

open circles instead of triangles. The three different beam incidence angles are again represented with blue, red, and green colors (29°, 38°, 46°, respectively). Overlaid on the data are filled symbols showing the results from the forward model. The filled circles show the forward model with parameters as before. Other filled symbols explore the nearby model parameter space if one parameter is perturbed, as described in the caption. From Figure 3.24, we see that, unlike the case for evergreen needleleaf forests (IGBP class 1), it may be possible to attribute a greater role for vegetation parameters as a source of spatial variation than ground parameters. The decreased role of ground parameters for spatial variation within IGBP class 3 (deciduous needleleaf forests) may perhaps be due to drier soils and lower values of ε_g (thus weaker double-reflections), as well as the geographic distribution being localized primarily to Eastern Siberia only (thus perhaps less variance in ground roughness h). These guesses require more evidence to verify, however; ground parameters may still have a significant involvement in the spatial variations in the radar backscatter.

Next, temporal variations in the radar backscatter are considered. If we consider the total temporal covariance like in Figure 3.9 and Figure 3.10, we would find that the rank-1 approximation is also good for IGBP class 3 and that the primary physical parameter associated with the temporal variation during unfrozen periods (sampled weekly over Sep2011-Apr2015) is most probably the ground relative permittivity ε_g , i.e. changes in soil moisture. Furthermore, changes in vegetation parameters ($V_{b,tot}$ or ε_v) positively correlated with the change in ε_g , i.e. vegetation growth or increased vegetation water content correlated with increase in soil moisture, would appear to be a secondary but non-negligible contributor to changes in L-band radar backscatter. Here we do not plot figures analogous to Figure 3.9 and Figure 3.10, but directly move on to analyzing the temporal variations in terms of longer and shorter time-scale changes, analogous to Figure 3.11 and Figure 3.12. Again, we use two-month medians for "low-frequency" and the high-pass filter of equation (3.2) for "high-frequency" analysis.

Like in Figure 3.11, Figure 3.25 plots histograms of pixel counts binned by the fraction $\lambda_1 / (\lambda_1 + \lambda_2 + \lambda_3)$ in the eigendecomposition of the "high-frequency" temporal covariance

matrix. Again the rank-1 approximation is poorer than for the total temporal covariance, but may still be used. The scatterplots of Figure 3.26 are the same as Figure 3.12 except for a change in IGBP class. The three beam incidence angles are plotted in different rows (top: 29°, middle: 38°, bottom: 46°). The left column plots are HV/HH[dB] vs HH[dB], while the right column plots are VV/HH[dB] vs HH[dB]. Filled colored triangles display the two-month median data values; these are the "slow" or "low-frequency" fluctuations. Only May-Jun, Jul-Aug, and Sep-Oct are displayed because most of the November-April period corresponds to frozen conditions. "Fast" fluctuations are depicted by the solid gray lines. Their directions display the dominant eigenvector of the 3x3 "high-frequency" covariance matrix for that pixel, and the half-length of the displayed line is $\sqrt{\lambda_1}$, the square root of the corresponding dominant eigenvalue.

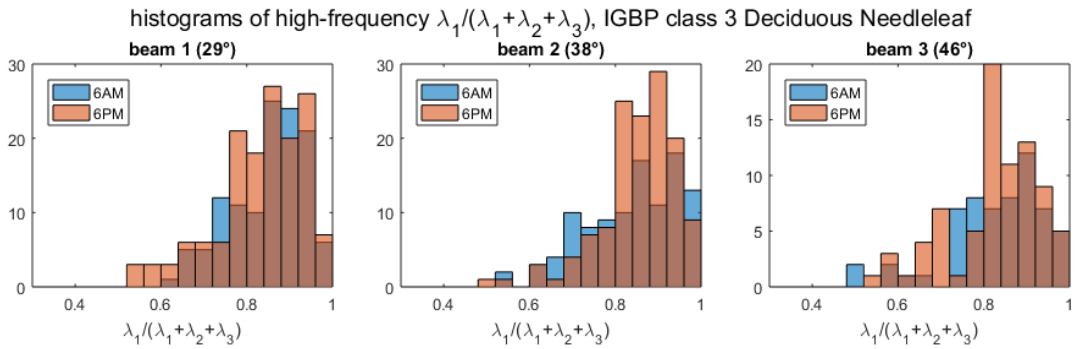

Figure 3.25. Histograms of pixel counts binned by the dominant eigenvalue as a fraction of the trace of the "high-frequency" temporal covariance matrices, for IGBP class 3 (deciduous needleleaf). The rank-1 approximation is poorer for the "high-frequency" covariance than for the total temporal covariance matrices.

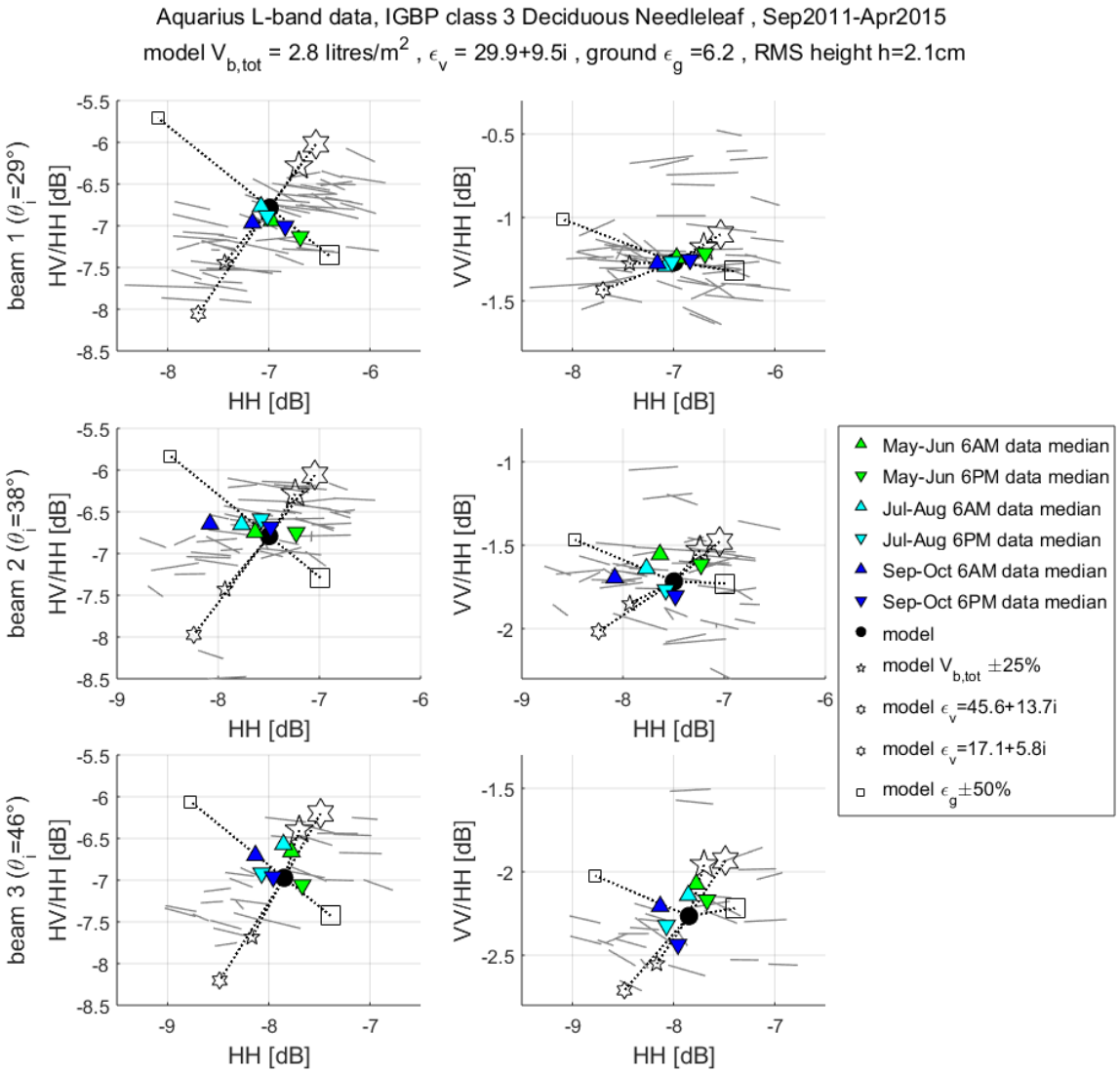

Figure 3.26. Same as Figure 3.12, showing "slow" and "fast" temporal variations in radar backscatter, but for IGBP class 3. The three rows correspond to the three beam incidence angles (top: 29°, middle: 38°, bottom: 46°). The left column plots are HV/HH[dB] vs HH[dB], while the right column plots are VV/HH[dB] vs HH[dB]. Filled triangles are the median over all pixels in IGBP class 3, data taken from time periods indicated by color: green (May-Jun), cyan (Jul-Aug), blue (Sep-Oct), thus displaying "low-frequency" temporal variations. Solid gray lines visually summarize the "high-frequency" temporal variations of each selected pixel in Figure 3.22. Specifically, the direction of this line is the dominant eigenvector of the covariance of the data (HH[dB], VV/HH[dB], HV/HH[dB]) for that pixel over the unfrozen period and after a high-frequency filter, and the half-length of the line is the square-root of the dominant eigenvalue. Filled circles show the forward model with input parameters as before in Figure 3.24. Other open symbols explore the nearby model parameter space if one parameter is perturbed. Larger and smaller five-pointed stars represent a change in $V_{b,tot}$ by +25% and -25%, respectively. Larger six-pointed stars represent an increase in ε_v to $\varepsilon_v = 45.6 + 13.7i$, while smaller six-pointed stars represent a

decrease to $\varepsilon_v = 17.1 + 5.8i$. Larger and smaller squares represent a change ε_g by +50% and -50%, respectively. Larger and smaller diamonds represent a change in h by +0.5cm and -0.5cm, respectively. Dotted lines merely serve to connect these open symbols with the filled circle.

From Figure 3.26, we see that the "fast" temporal changes are also probably due to changes in ground relative permittivity ε_g (i.e. soil moisture) with positively correlated changes in vegetation parameters ($V_{b,tot}$ or ε_v , i.e. vegetation growth or increased vegetation water content) as a secondary contribution. The "slow" temporal changes in radar backscatter are a little more ambiguous and interesting to interpret. It seems that at incidence angle of 29°, soil moisture changes dominate the slow changes, but at incidence angle of 46°, vegetation changes may become more important, consistent with a reduced sensitivity to ground parameters at larger incidence angles.

Let us summarize the comparison between model and data for IGBP class 3 (deciduous needleleaf forests) areas, taken at a coarse regional spatial scale. Overall the model and data show reasonable agreement. Spatial variations in the L-band radar backscatter are due to a combination of variations in both ground and vegetation factors. Temporal variations (sampled weekly during unfrozen periods) in the L-band radar backscatter are primarily due to changes in soil moisture, and secondarily due to changes in vegetation (either vegetation growth or increased vegetation water content, or both), and the changes in soil moisture and vegetation are positively correlated with each other. For temporal variations at slower time-scales of months, vegetation changes may become more dominant at higher incidence angles.

3.6 Deciduous broadleaf forests (IGBP class 4)

The geographic distribution of this class is in the eastern United States, parts of Europe, and northeast Asia, and the Chaco plain of South America. However due to human development, the only significant contiguous region in this class with data not excessively RFI polluted lies in the tropical dry forest of the Chaco plain of South America. As this land area is relatively small compared to some of the other IGBP land cover classes, a random 50% selection of the available filtered EASE2 grid are displayed. Using the MODIS land cover map, the geographic distribution of the deciduous broadleaf forests (IGBP class 4) is shown

in Figure 3.5 and Figure 3.27 for emphasis. The data pixels are marked by open triangle symbols on Figure 3.27, with upward-pointing triangles for descending-pass data (local time 6AM) and downward-pointing triangles for ascending-pass data (local time 6PM), while the three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°), like in the previous figures. Note that the radar footprints and swath widths are now no longer small compared to the land region being considered: it is apparent from Figure 3.27 that only about two full swath widths suffice to cover the entire longitude span of this region. As such, when we analyze the data from this IGBP class as whole, there is potentially a strong correlation between incidence angle and ascending/descending pass with other environmental factors that are changing from East to West across the region [96].

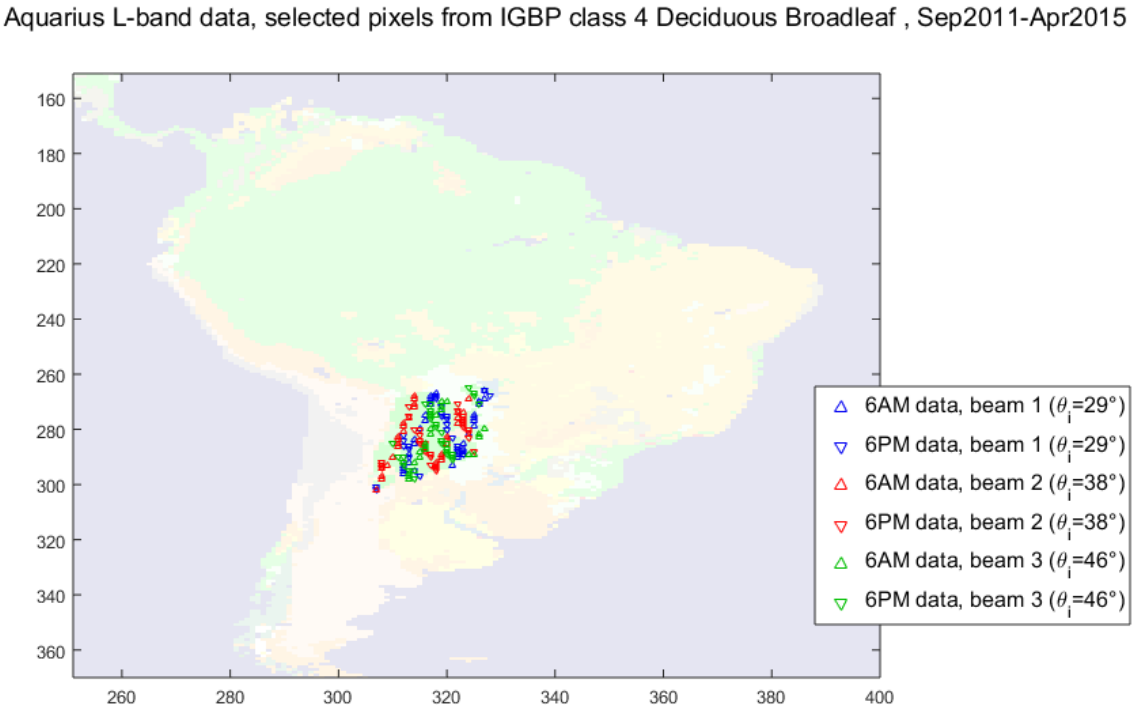

Figure 3.27. Geographic distribution of contiguous deciduous broadleaf forests (IGBP class 4) with sufficient good quality Aquarius scatterometer data is concentrated in the Chaco Plain region of South America. A random 50% selection of filtered Aquarius data EASE2 grid pixels are displayed, marked by open triangle symbols. The three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°). Upward-pointing triangles are for descending-pass data (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM).

We attempt to proceed with the analysis as for the other IGBP classes. The scatterometer data is displayed in Figure 3.28, which depicts two scatterplots: one of VV/HH[dB] vs HH[dB], and another of HV/HH[dB] vs HH[dB], similar to Figure 3.7 except for a change in IGBP class. The symbol descriptions are in the figure caption. Filled circles show the results from the forest forward model from Chapter 2. The parameters $m = 0$ (uniformly random orientation distribution) in equation (2.6) were used, and the vegetation cylinder relative permittivity chosen as $\varepsilon_v = 29.9 + 9.5i$ as usual. Other model input parameters are: total volume in branches $V_{b,tot} = 3.0 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, ground relative permittivity $\varepsilon_g = 4.3$, ground roughness RMS height $h = 2.5\mathrm{cm}$. These latter three parameters ($V_{b,tot}$, ε_g, h) were chosen by fitting to the pooled (ascending and descending) data, in the same way as in equation (3.1). The fitted parameters are somewhat similar to the other forest classes. The fitted ground relative permittivity $\varepsilon_g = 4.3$ seems unusually low. However it is also true that there is a long half-year dry season (winter-spring) with negligible rainfall in some areas [96], so a low value of ε_g might not be implausible.

From Figure 3.28, we can see that, as previously mentioned, the correlation between incidence angle and ascending/descending pass with other geographical factors is indeed a confounding issue. In particular, from the scatterplot of VV/HH[dB] vs HH[dB], the data for beam 2 shows two distinct clusters, one with VV/HH[dB] closer to about -0.3dB, and the other with VV/HH[dB] of about -1dB. The cluster with VV/HH[dB] around -0.3dB corresponds to the strip of western pixels (mostly of beam 2, descending pass 6AM data) in Figure 3.27. Geographically these are at a higher elevation and run up towards the Andean cordillera, and despite the terrain slope filter that had been applied, many of these pixels are border cases and some high-terrain slope lands of the Subandean Sierras are likely included in part of the radar footprints (which, at ~100km, are significantly larger than the pixel size of ~36km). This "contamination" is consistent with the data observations, as high-terrain slope areas have expected VV/HH ratio close to unity (i.e. or 0dB). Likewise, other confounding geographical factors may be responsible for the apparent disparity in median values between ascending/descending pass data. As such, we do not attempt to use the model to analyze the spatial variation in the radar backscatter data across different pixels.

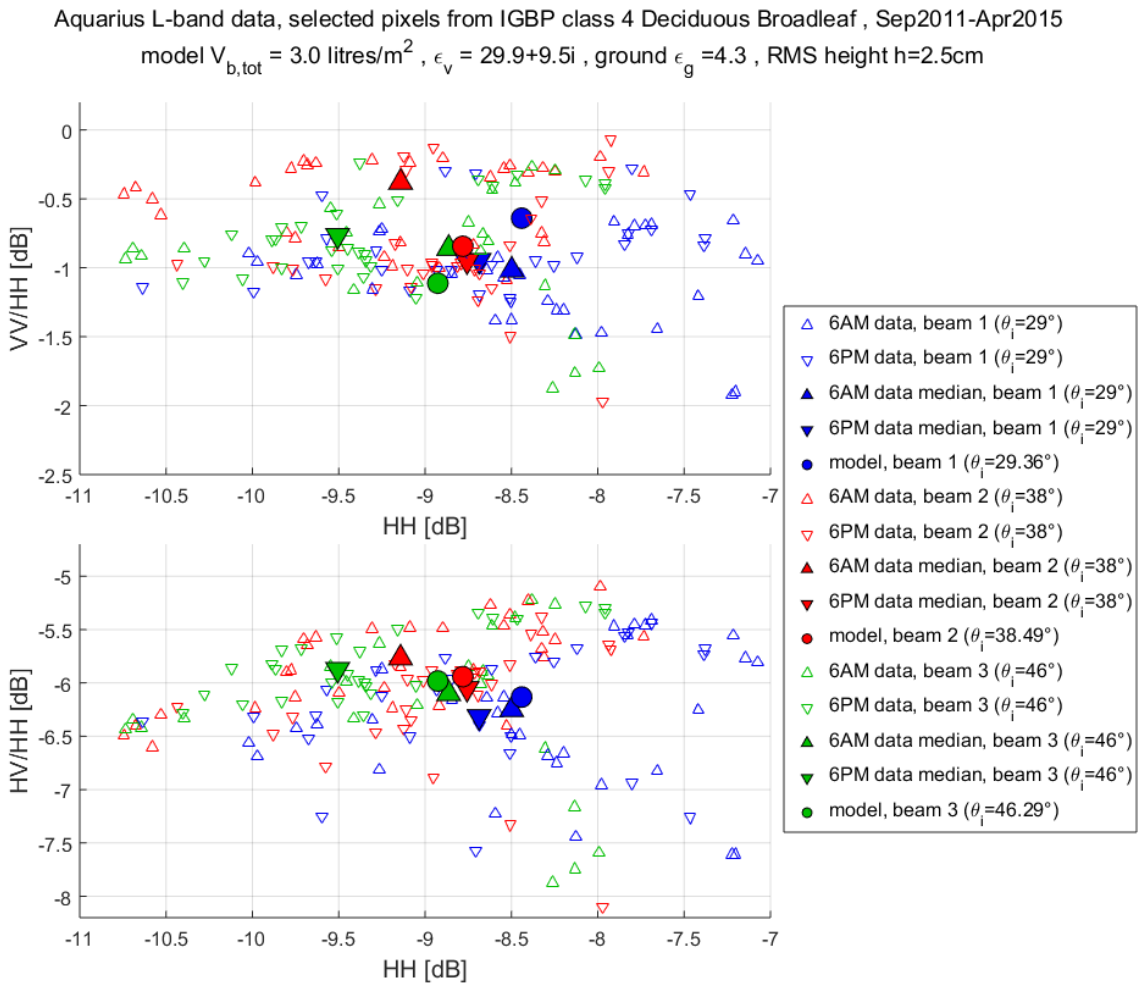

Figure 3.28. Aquarius L-band scatterometer data for EASE2 grid pixels classified as deciduous broadleaf forests (IGBP class 4) displayed in Figure 3.27. The three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°). Upward-pointing triangles are for descending-pass data (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM). Each open triangle represents data from one of the pixels shown in Figure 3.27, taking the mean over Sep2011-Apr2015 for each pixel. Filled triangles are the median over open triangles, but with median taken over all data pixels of this class instead of only the 50% displayed. Filled circles show the forest forward model from Chapter 2, with parameters $m = 0$ (uniformly random orientation distribution) in equation (2.6), total volume in branches $V_{b,tot} = 3.0 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, ground relative permittivity $\varepsilon_g = 4.3$, and ground roughness RMS height $h = 2.5\mathrm{cm}$

However, the temporal analysis shows clear results. We plot in Figure 3.30 the usual scatterplots with the two-month medians and "fast/high-frequency" dominant eigenvalue/eigenvector lines. In fact for this IGBP class, the first eigenvalue/eigenvector appear to be exceptionally dominant, as shown by Figure 3.29. From Figure 3.30, especially the scatterplots of VV/HH[dB] vs. HH[dB], we see that the primary physical parameter associated with both "fast" and "slow" temporal variations are most probably the ground relative permittivity ε_g , i.e. changes in soil moisture. The modelled direction of variation on the HV/HH[dB] vs. HH[dB] scatterplot variation for changes in ε_g is not exactly aligned with the temporal data, suggesting changes in vegetation parameters too as a secondary but non-negligible contributor to changes in L-band radar backscatter, in particular to HV. Furthermore, these changes in vegetation parameters ($V_{b,tot}$ or ε_v) are positively correlated with the change in ε_g, i.e. vegetation growth (or increased vegetation water content) correlated with increase in soil moisture.

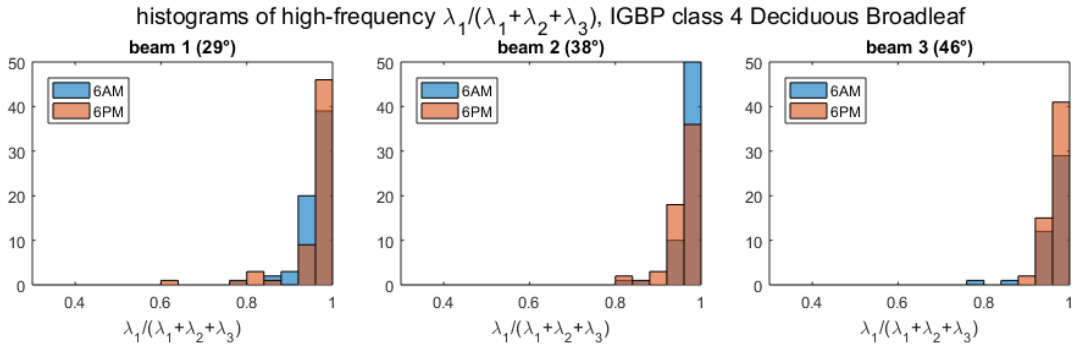

Figure 3.29. Histograms of pixel counts binned by the dominant eigenvalue as a fraction of the trace of the "high-frequency" temporal covariance matrices, verifying that the dominant eigenvalue/eigenvector indeed capture most of the temporal variation. All eligible pixels of IGBP class 4 (deciduous broadleaf), not just the 50% random subsample, were counted here. The rank-1 approximation is even better for the total temporal covariance matrices (not shown).

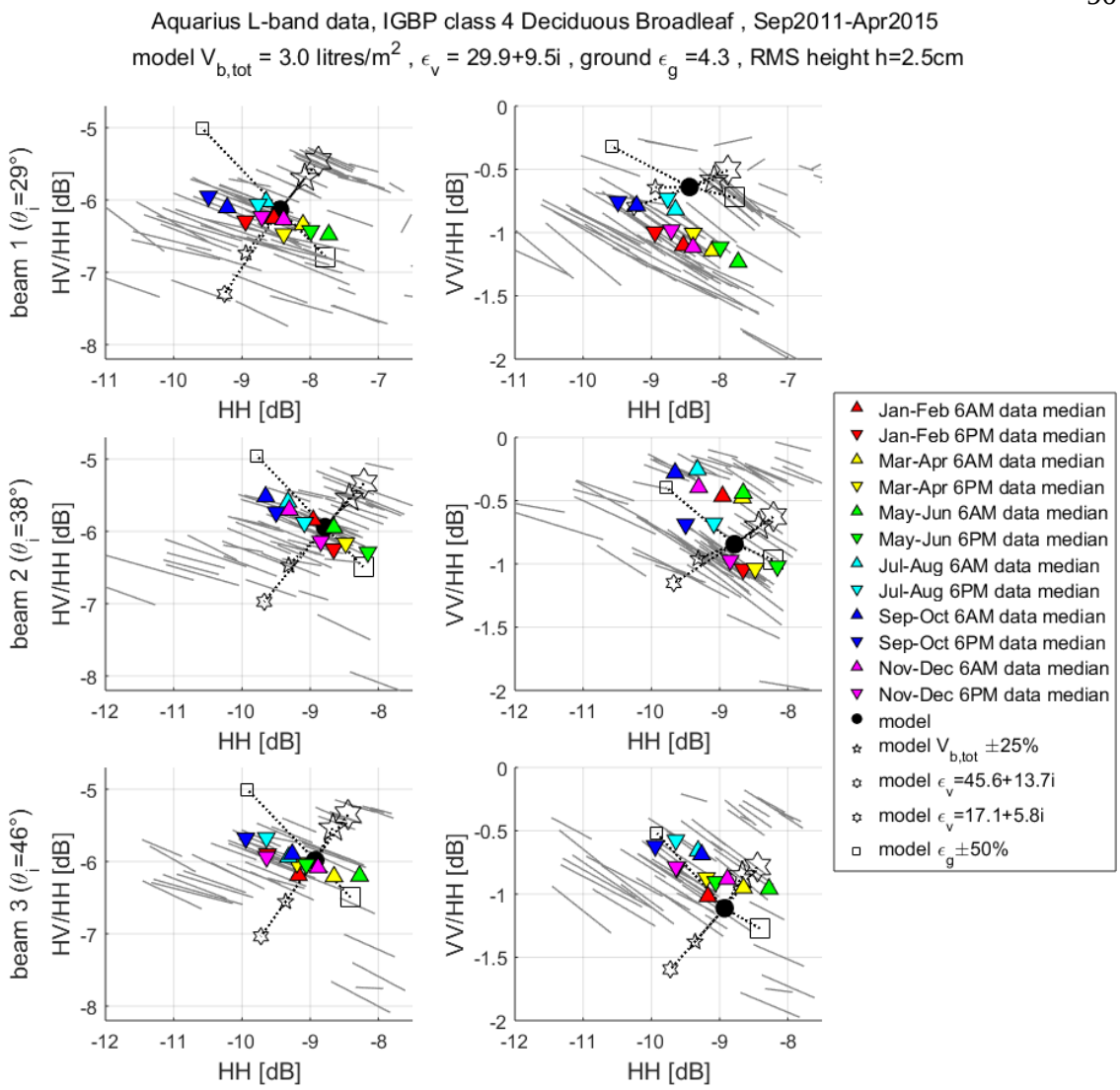

Figure 3.30. Same as Figure 3.12, showing "slow" and "fast" temporal variations in radar backscatter, but for IGBP class 4. The three rows correspond to the three beam incidence angles (top: 29°, middle: 38°, bottom: 46°). The left column plots are HV/HH[dB] vs HH[dB], while the right column plots are VV/HH[dB] vs HH[dB]. Filled triangles are the median over all pixels in IGBP class 4, data taken from time periods indicated by color: red (Jan-Feb), yellow (Mar-Apr), green (May-Jun), cyan (Jul-Aug), blue (Sep-Oct), magenta (Nov-Dec), thus displaying "low-frequency" temporal variations. Solid gray lines visually summarize the "high-frequency" temporal variations of each selected pixel in Figure 3.27. Specifically, the direction of this line is the dominant eigenvector of the covariance of the data (HH[dB], VV/HH[dB], HV/HH[dB]) for that pixel over the unfrozen period and after a high-frequency filter, and the half-length of the line is the square-root of the dominant eigenvalue. Filled circles show the forward model with input parameters as before in Figure 3.28. Other open symbols explore the nearby model parameter space if one parameter is perturbed. Larger and smaller five-pointed stars represent a change in $V_{b,tot}$ by +25% and -25%, respectively. Larger six-pointed stars represent an increase in ε_v to $\varepsilon_v = 45.6 + 13.7i$, while smaller six-pointed stars represent a decrease to $\varepsilon_v = 17.1 + 5.8i$. Larger and smaller

squares represent a change ε_g by +50% and -50%, respectively. Larger and smaller diamonds represent a change in h by +0.5cm and -0.5cm, respectively. Dotted lines merely serve to connect these open symbols with the filled circle.

3.7 Mixed forests (IGBP class 5)

The "mixed forests" class in the IGBP classification comprises lands dominated by trees but with interspersed mixtures or mosaics of the other four forests types, with none of the specific forest types exceeding 60% of the landscape. Using the MODIS land cover map, the geographic distribution of IGBP class 5 is shown in Figure 3.5 and Figure 3.31 emphasis. A random 20% selection of the EASE2 grid pixels is displayed in Figure 3.31. As can be seen, much of the land falling into this class is still in or near the boreal/hemiboreal zone. Indeed, it shall turn out that despite the greater heterogeneity of species, the radar backscatter characteristics from this class still bear similarity to IGBP class 1 (evergreen needleleaf forests) and IGBP class 3 (deciduous needleleaf forests). Data from non-coniferous temperate forests are underrepresented because these tend to be encroached significantly by areas of human settlement, greatly increasing the chance of being excluded by RFI flags.

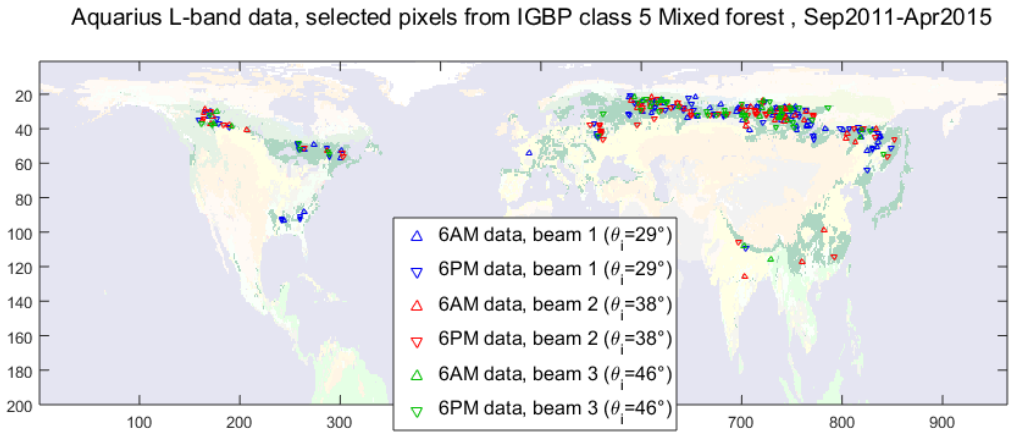

Figure 3.31. Geographic distribution of mixed forests (IGBP class 5) and randomly selected 20% of filtered Aquarius data EASE2 grid pixels, marked by open triangle symbols.

The scatterometer data for the selected pixels is displayed in Figure 3.32, which depicts two scatterplots: one of VV/HH[dB] vs HH[dB], and another of HV/HH[dB] vs HH[dB], similar to Figure 3.7 except for a change in IGBP class. The symbol descriptions are in the figure caption. Filled circles show the results from the forest forward model from Chapter 2. The

parameters $m = 1$ and $\theta_0 = \pi/2$ (cosine-squared orientation distribution about the horizontal) in equation (2.6) were used, and the vegetation cylinder relative permittivity chosen as $\varepsilon_v = 29.9 + 9.5i$ as usual. Other model input parameters are: total branch volume $V_{b,tot} = 3.5 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, ground relative permittivity $\varepsilon_g = 8.8$, ground roughness RMS height $h = 2.5\mathrm{cm}$. These latter three parameters ($V_{b,tot}$, ε_g, h) were chosen by fitting to the pooled (ascending and descending) data, in the same way as in equation (3.1).

From Figure 3.32, it can be seen that the forward model gives reasonable overall agreement to the data, albeit using fitted parameters. The fitted parameters are very similar to those for IGBP class 1 (evergreen needleleaf forests). This suggests that, in terms of the underlying physical factors affecting L-band radar backscatter, IGBP classes 1 and 5 (and class 3 as well) are rather similar, despite the heterogeneity of forest types by definition of IGBP class 5. In view of the proximity and overlap of the geographic regions for IGBP classes 1 and 5 (Figure 3.31), it is to be expected that signification components of "mixed forests" are still coniferous forests. In fact even the MODIS land cover map, which is taken as ground truth here, is likely to have some amount of confusion between the two classes [75].

With regards to the variation between the descending-pass (local time 6AM) and ascending-pass (local time 6PM) data, again similar to IGBP class 1, there might be an overall trend of L-band radar backscatter being slightly higher from ascending-pass (local time 6PM) data compared to descending-pass (local time 6AM) data; but this difference, when considered in an average sense over time and multiple pixels, is small (compared to the spatial variation between pixels and longer period temporal variations to be discussed later), so we proceed to pool the 6AM and 6PM data as usual.

The model is then used to analyze the spatial variation in the radar backscatter data across different pixels. Figure 3.33 displays the same data as Figure 3.32, now with the descending-pass (local time 6AM) and ascending-pass (local time 6PM) data pooled and all marked with open circles instead of triangles. Overlaid on the data are filled symbols showing the results from the forward model with parameters as before. Other filled symbols explore the nearby model parameter space if one parameter is perturbed, as described in the caption. From

Figure 3.33, we see that the ground roughness RMS height h is required to explain spatial variations in the VV/HH ratio, while vegetation parameters are require to explain spatial variations in HV. Thus both vegetation and ground parameters are involved in the spatial variations in the radar backscatter.

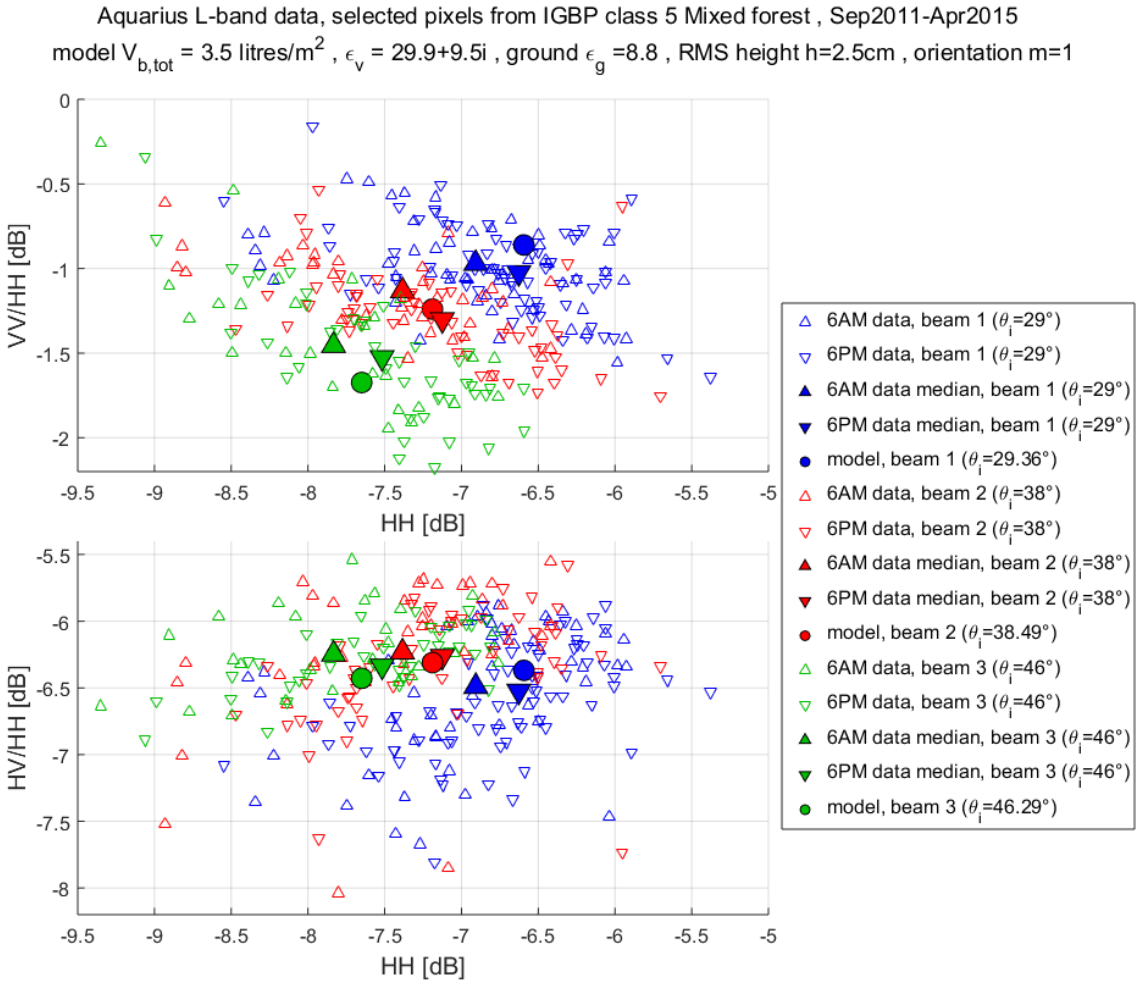

Figure 3.32. Aquarius L-band scatterometer data for EASE2 grid pixels classified as mixed forest (IGBP class 5). The three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°). Upward-pointing triangles are for descending-pass data (local time 6AM) and downward-pointing triangles are for ascending-pass data (local time 6PM). Each open triangle represents data from one of the pixels shown in Figure 3.31, taking the mean over unfrozen periods from Sep2011-Apr2015 for each pixel. Filled triangles are the median over open triangles, but with median taken over all data pixels of this class instead of only the 20% displayed. Filled circles show the forest forward model from Chapter 2, with parameters $m = 1$ and $\theta_0 = \pi/2$ (preferential horizontal orientation) in equation (2.6), total volume in branches $V_{b,tot} = 3.5 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, ground relative permittivity $\varepsilon_g = 8.8$, ground roughness RMS height $h = 2.5\mathrm{cm}$.

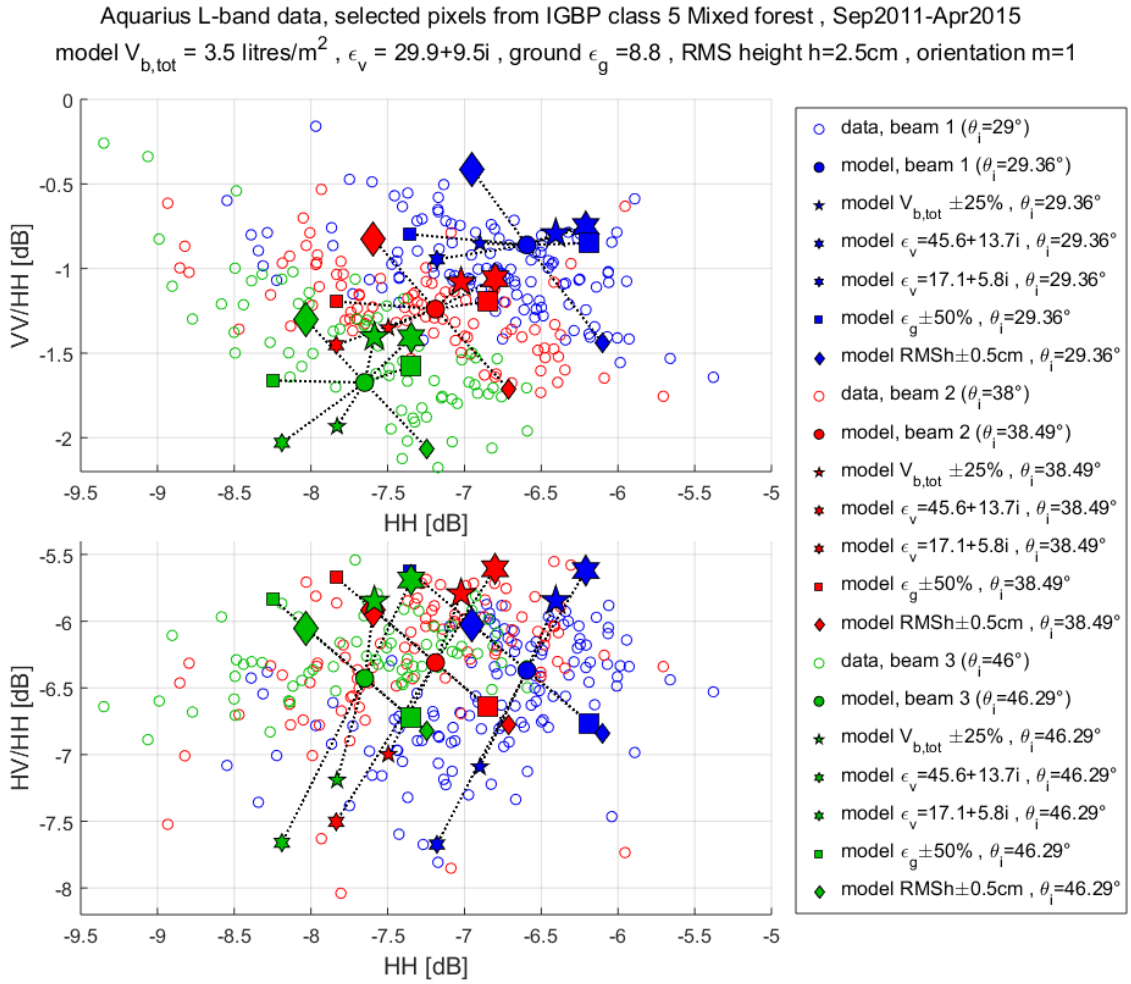

Figure 3.33. Same as Figure 3.32, but with no distinction made between descending- and ascending-pass data; all are marked with open circles. Filled circles show the forward model with parameters as before in Figure 3.32. Other filled symbols explore the nearby model parameter space if one parameter is perturbed. Larger and smaller five-pointed stars represent a change in $V_{b,tot}$ by +25% and -25%, respectively. Larger six-pointed stars represent an increase in ε_v to $\varepsilon_v = 45.6 + 13.7i$, while smaller six-pointed stars represent a decrease to $\varepsilon_v = 17.1 + 5.8i$. Larger and smaller squares represent a change ε_g by +50% and -50%, respectively. Larger and smaller diamonds represent a change in h by +0.5cm and -0.5cm, respectively.

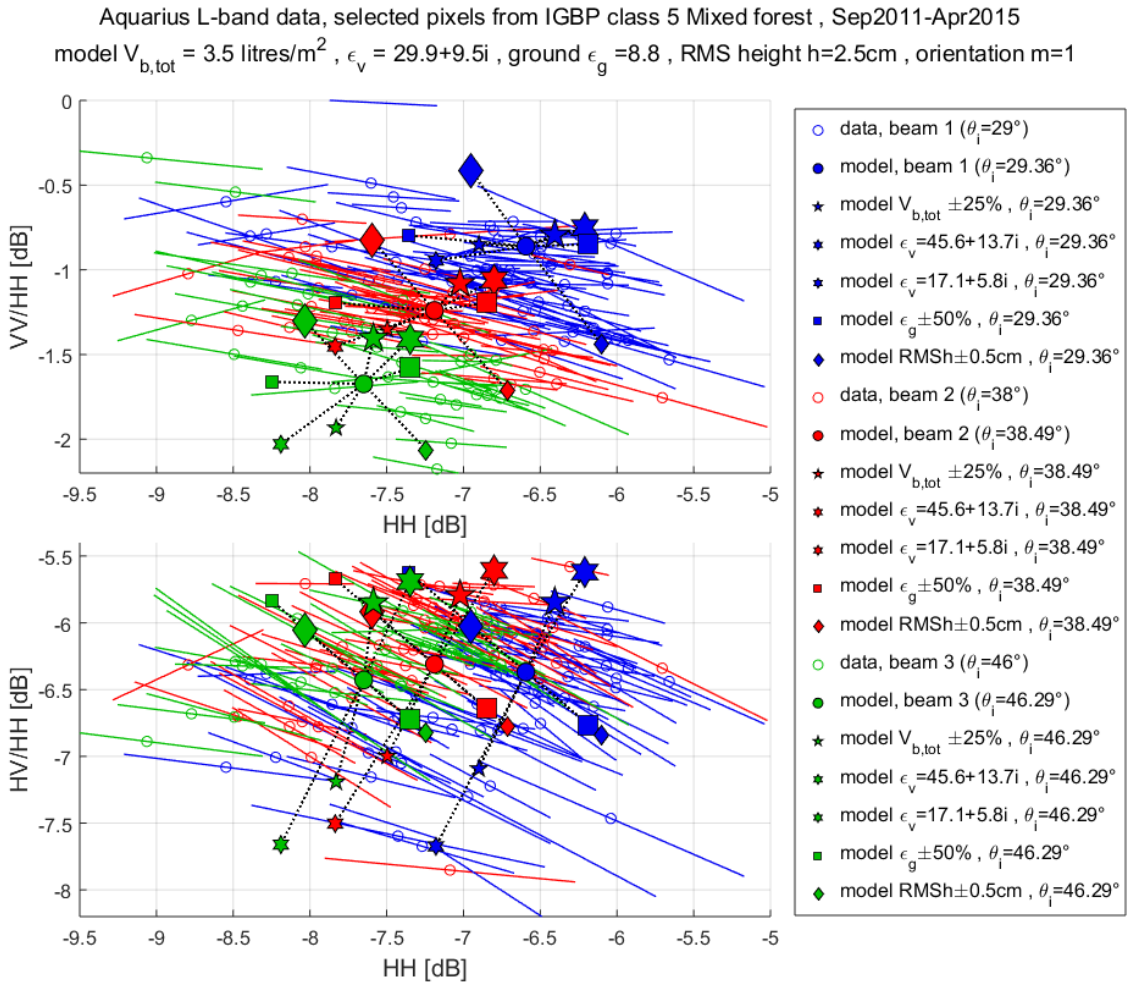

Figure 3.34. Same as Figure 3.33, but also showing temporal variations in the data. Each open circle, corresponding to one of the pixels shown in Figure 3.31 and representing the mean taken over unfrozen periods from Sep2011-Apr2015, now also comes with a line through it that visually summarizes the temporal variation. Specifically, the direction of this line is the dominant eigenvector of the covariance of the data (HH[dB], VV/HH[dB], HV/HH[dB]) for that pixel over the unfrozen period, and the half-length of the line is the square-root of the dominant eigenvalue.

Next, temporal variations in the radar backscatter are considered in Figure 3.34, which contains the usual scatterplots with the dominant eigenvalue/eigenvector lines for the total temporal covariance, similar to Figure 3.10. The validity of the rank-1 approximation to the 3x3 total temporal covariance is checked in Figure 3.35. Figures are not shown for further analysis into "fast/high-frequency" and "slow/low-frequency" temporal variations for IGBP class 5, in part because the rank-1 approximation turns out to be poor for the "fast/high-frequency" covariance and the results are less conclusive. From Figure 3.34, we see that the

primary physical parameter associated with the temporal variation is again most probably the ground relative permittivity ε_g, i.e. changes in soil moisture. In the median case, the values of ground relative permittivity corresponding to ±1 standard deviation of variation in HH[dB] are approximately $\varepsilon_g = 5.4$ and $\varepsilon_g = 15.2$. Though details vary depending on soil type, these are reasonable values for drier and wetter soils. Additionally, there is greater spatial heterogeneity in the temporal variation trends compared to the other classes. This may speculatively be attributable to IGBP class 5 covering a greater spread of geographic and ecological zones, and being intrinsically an "umbrella" class.

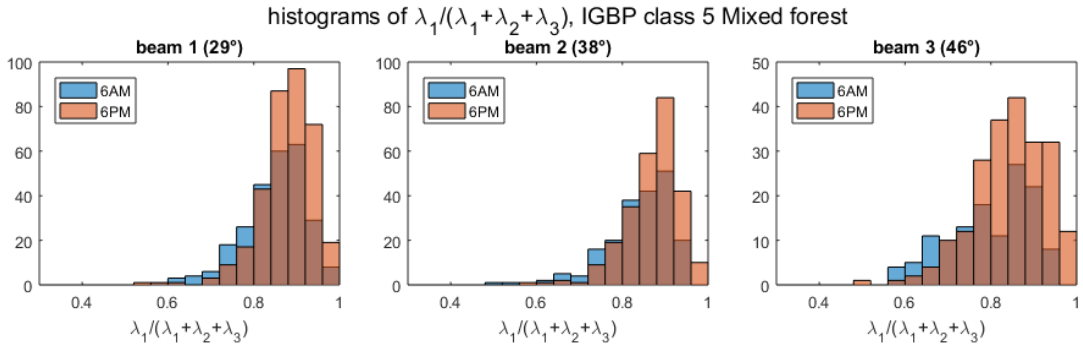

Figure 3.35. Histograms of pixel counts binned by the dominant eigenvalue as a fraction of the trace of the temporal covariance matrices, verifying that the dominant eigenvalue/eigenvector indeed capture most of the temporal variation. Note that all eligible pixels of IGBP class 5 (mixed forest), not just the 20% random subsample, were counted in these histograms.

To address the question of whether or not an orientation distribution of cylinders that is uniformly random might instead be a better fit to the data, we plot Figure 3.36. Figure 3.36 is the same as Figure 3.33 except that a uniformly random orientation distribution of cylinders ($m = 0$ in equation (2.6)) was used in the forward model, and equation (3.1) refitted to the data to give $V_{b,tot} = 4.2 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, $\varepsilon_g = 8.8$, and $h = 2.25$ cm. Following the same estimation procedure as for IGBP class 1, here $V_{b,tot} = 4.2$ litres/m^2 corresponds to 94.5tons/ha of total aboveground biomass instead of 78.8tons/ha for $V_{b,tot} = 3.5 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$.

An argument can be made from Figure 3.36 that the uniformly randomly oriented distribution may also be a reasonable fit to the data. Both biomass values seem to be reasonable estimates;

as mentioned in the previous chapter, we cannot expect very accurate estimates for total above-ground biomass for dense vegetation using only L-band normalized radar backscatter cross-sections, hence we had parameterized using $V_{b,tot}$ instead. The main differences between the two orientation distributions lie in the way in which backscatter at all polarizations depend on incidence angle. For $m = 1$ (cosine-squared distribution about the horizontal), the ratio HH/VV increases (i.e. VV/HH decreases) with increasing incidence angle. For $m = 0$ (uniformly randomly oriented cylinders) however, the direct canopy backscatter has no preferential polarization for either VV or HH, so the total VV/HH ratio is only weakly dependent on incidence angle, primarily through the scattering terms involving the ground. HV is more strongly dependent on incidence angle for $m = 1$ than for $m = 0$ due to the change in projected distribution perpendicular to the beam direction. Other than incidence angle, there are also some slight differences between the $m = 1$ and $m = 0$ models in terms of the directions of variation on the scatterplots associated with the vegetation parameters. The reality might be closer to some mixture between the $m = 1$ and $m = 0$ models.

To further illustrate that there is indeed sensitivity to orientation distributions from the radar data, Figure 3.37 plots the best fit parameters from equation (3.1) if the forward model used a uniformly random orientation distribution of cylinders ($m = 0$ in equation (2.6)) for IGBP class 3 instead. Evidently, a uniformly random orientation distribution provides a poorer fit than cosine-squared orientation distribution about the horizontal for IGBP class 3. Though the disparity is not as large as for IGBP class 3, IGBP class 1 is also a better fit by the preferentially horizontal orientation distribution compared to the uniformly random distribution (figures not displayed). On the other hand, a uniformly random orientation distribution was more appropriate for IGBP classes 2 and 4 as compared to a preferential orientation distribution (figures not displayed).

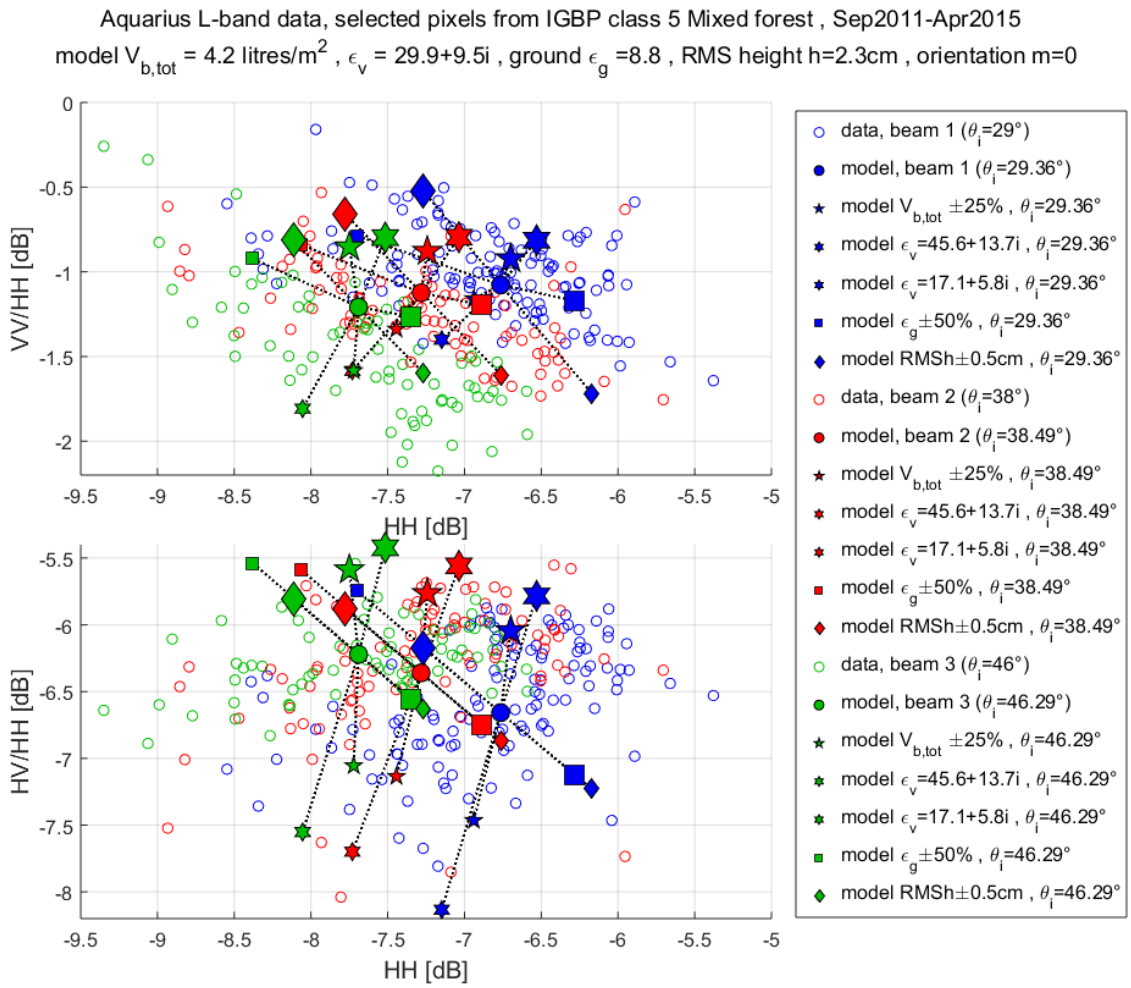

Figure 3.36. Same as Figure 3.33, except with a uniformly random orientation distribution of cylinders ($m = 0$ in equation (2.6)) for the forward model, and equation (3.1) refitted to find best fit parameters as shown: total volume in branches $V_{b,tot} = 4.2 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, ground relative permittivity $\varepsilon_g = 8.8$, ground roughness RMS height $h = 2.3$cm.

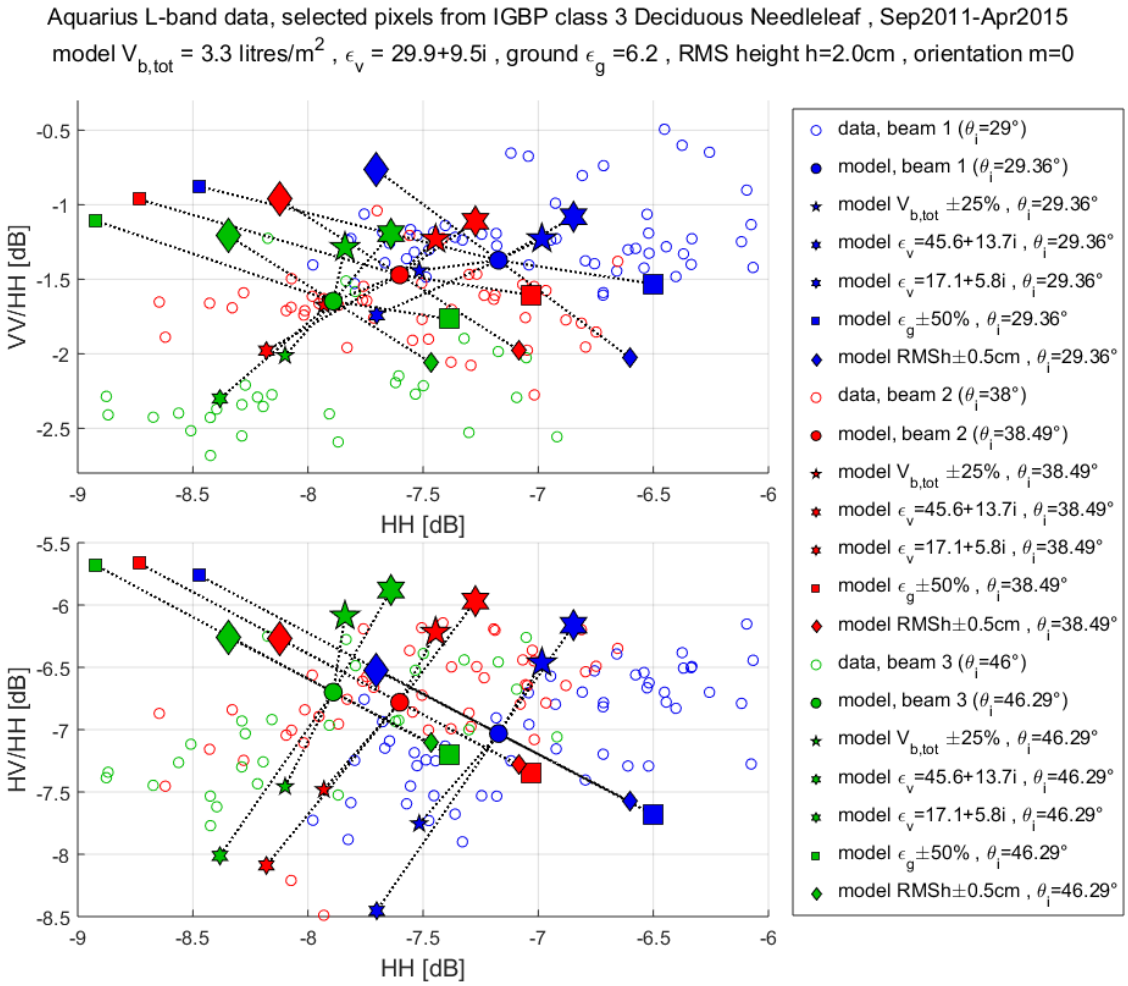

Figure 3.37. Same as Figure 3.24, except with a uniformly random orientation distribution of cylinders ($m = 0$ in equation (2.6)) for the forward model, and equation (3.1) refitted to find best fit parameters as shown: total volume in branches $V_{b,tot} = 3.3 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, ground relative permittivity $\varepsilon_g = 6.2$, ground roughness RMS height $h = 2.0$cm. Comparing with Figure 3.24, it is evident that a uniformly random orientation distribution of cylinders is less appropriate in this case.

In summary, the L-band radar multi-polarization backscatter characteristics for IGBP class 5 (mixed forests) areas, taken at a coarse global spatial scale, are broadly similar to those for IGBP class 1 (evergreen needleleaf). The orientation distribution parameters $m = 1, \theta_0 = \pi/2$ (preferentially horizontal orientation distribution) and $m = 0$ (uniformly random orientation distribution) both fit the data comparably, in reasonable expectation with the heterogeneity of forests within this class.

3.8 Frozen vs. unfrozen conditions

In Sections 3.3 to 3.7, the Aquarius L-band radar data had been filtered for frozen conditions and only data corresponding to unfrozen conditions were analyzed for spatial and temporal patterns. This was because, if otherwise unfiltered, the large change between frozen and unfrozen conditions would have dominated and swamped other effects. In this section, the L-band radar multi-polarization backscatter would be compared between frozen and unfrozen conditions. The primary IGBP forest classes involved are class 1 (evergreen needleleaf), class 3 (deciduous needleleaf), and class 5 (mixed forest), as these cover the boreal/hemiboreal/temperate zones.

For this frozen/unfrozen comparison, data with corresponding temperature <0°C during a time period from two weeks before to two weeks after were flagged as "frozen", while data with corresponding temperature >5°C during a time period from two weeks before to two weeks after were flagged as "unfrozen". As before, the temperature value used was the temperature of the 0-10cm subsurface layer as provided by the NCEP GFS GDAS model product provided along with the Aquarius dataset. Only EASE grid pixels with at least 10 weeks of such "frozen" and 10 weeks of such "unfrozen" data over the Sep2011-Apr2015 period were included for consideration in the analysis. Data from a random 10% selection of these EASE2 grid pixels are displayed in Figure 3.38, Figure 3.39, and Figure 3.40 for IGBP classes 1,3, and 5 respectively. In each of these three figures, the three different beam incidence angles are again represented by color: blue (29°), red (38°), and green (46°), and data for descending-pass (local time 6AM) and ascending-pass (local time 6PM) are represented by upward-pointing and downward-pointing triangles, respectively. Each pair of open/filled triangles linked by a dotted line summarize data from an EASE2 grid pixel, with the open triangle representing the mean over unfrozen periods, while the filled triangle is the mean over frozen periods, from Sep2011-Apr2015. Filled circles show the forward model with parameters as before in Figure 3.8, Figure 3.24, and Figure 3.33. Filled squares show the forward model but with the vegetation cylinder relative permittivity changed to $\varepsilon_v = 4.5 + 1.1i$ and ground relative permittivity changed to $\varepsilon_g = 4.0$ instead. These parameters attempt to represent the relative permittivities for frozen branches and frozen soil,

respectively, following experimental results from El-rayes and Ulaby [97]; for soil we have neglected the imaginary part of the relative permittivity as before.

The comparison is similar for all three figures so we shall discuss them all together. In agreement with the data, the model predicts a decrease in HH of about 3-5dB from frozen to unfrozen conditions, and a smaller effect at smaller incidence angle of 29°. The model predicts a slight decrease in the VV/HH ratio, especially at the larger incidence angle of 46°, while the data shows either negligible change in the VV/HH ratio or a very small decrease. Finally the model predicts a large decrease in the HV/HH ratio, by about 6dB, while the data only shows a decrease in HV/HH by about 3dB; though the model prediction that this decrease is greater at 29° incidence angle is in qualitative agreement with the data. A possible explanation accounting for all these observations is that overall the model has, with simply a naïve change in dielectric relative permittivities, essentially managed to capture the important physical differences, but there are additional important details to the frozen/unfrozen transition. Consider that in winter, not only are the tree branches frozen, they are often also "coated" with some amount of ice, effectively increasing cylinder radii. Hanging icicles also change the effective cylinder size and orientation distributions. These may explain slightly higher VV values and much higher HV values than expected from simply changing the dielectric relative permittivities values.

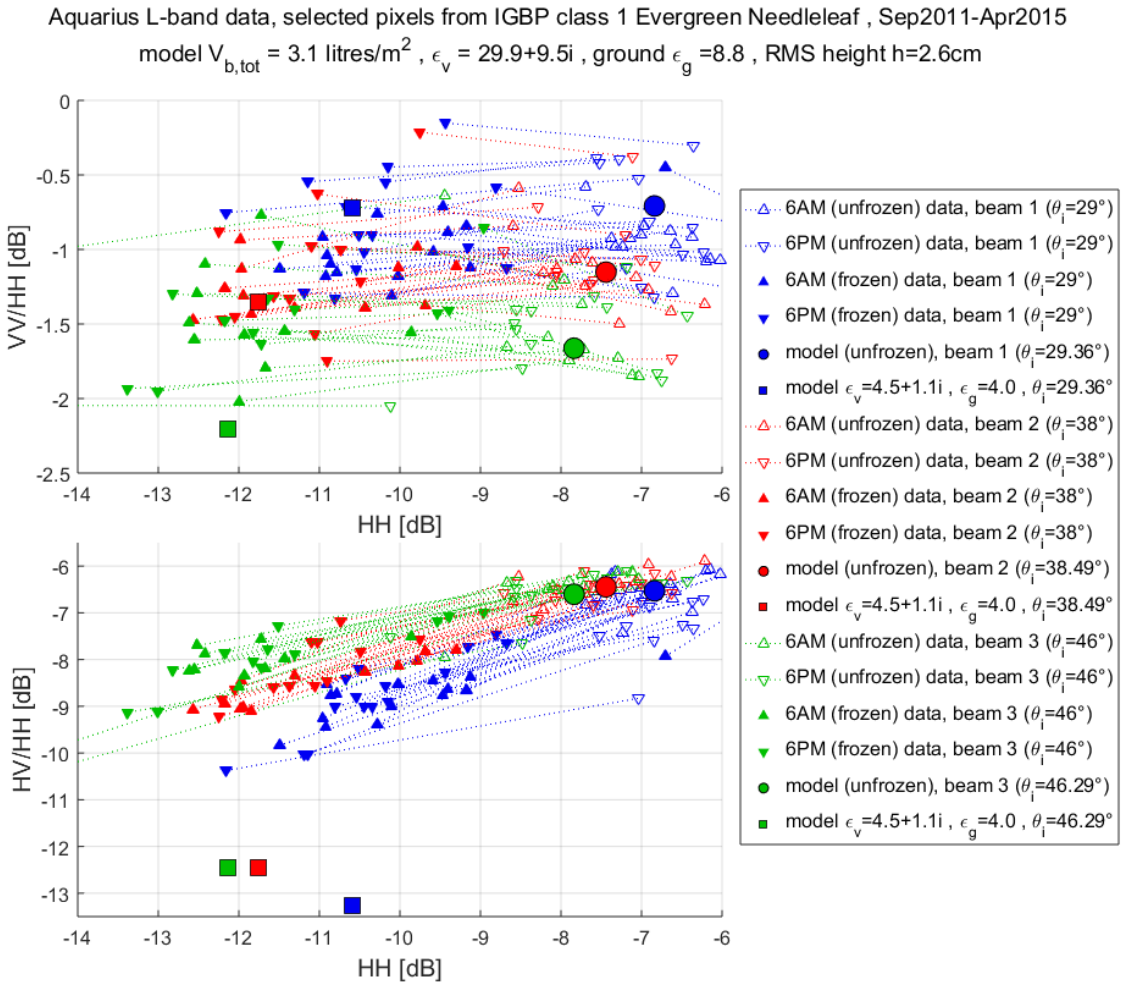

Figure 3.38, Aquarius L-band scatterometer data for a 10% random selection of EASE2 grid pixels classified as evergreen needleleaf (IGBP class 1) and satisfying various filter conditions as described in the text. The three different beam incidence angles are represented by color: blue (29°), red (38°), and green (46°). Upward- and downward-pointing triangles are for descending-pass data (local time 6AM) and ascending-pass data (local time 6PM) respectively. Open triangles represent the mean over unfrozen periods, while filled triangles are the mean over frozen periods, from Sep2011-Apr2015. Each pair of open and filled triangles linked by a dotted line summarize data from an EASE2 grid pixel. Filled circles show the forward model with parameters as before in Figure 3.8 (horizontal cosine-squared orientation distribution of cylinders, total volume in branches $V_{b,tot} = 3.1 \times 10^{-3}\mathrm{m}^3/\mathrm{m}^2$, vegetation cylinder relative permittivity $\varepsilon_v = 29.9 + 9.5i$, ground relative permittivity $\varepsilon_g = 8.8$, ground roughness RMS height $h = 2.6\mathrm{cm}$). Filled squares show the same forward model but with parameters $\varepsilon_v = 4.5 + 1.1i$ and $\varepsilon_g = 4.0$ instead.

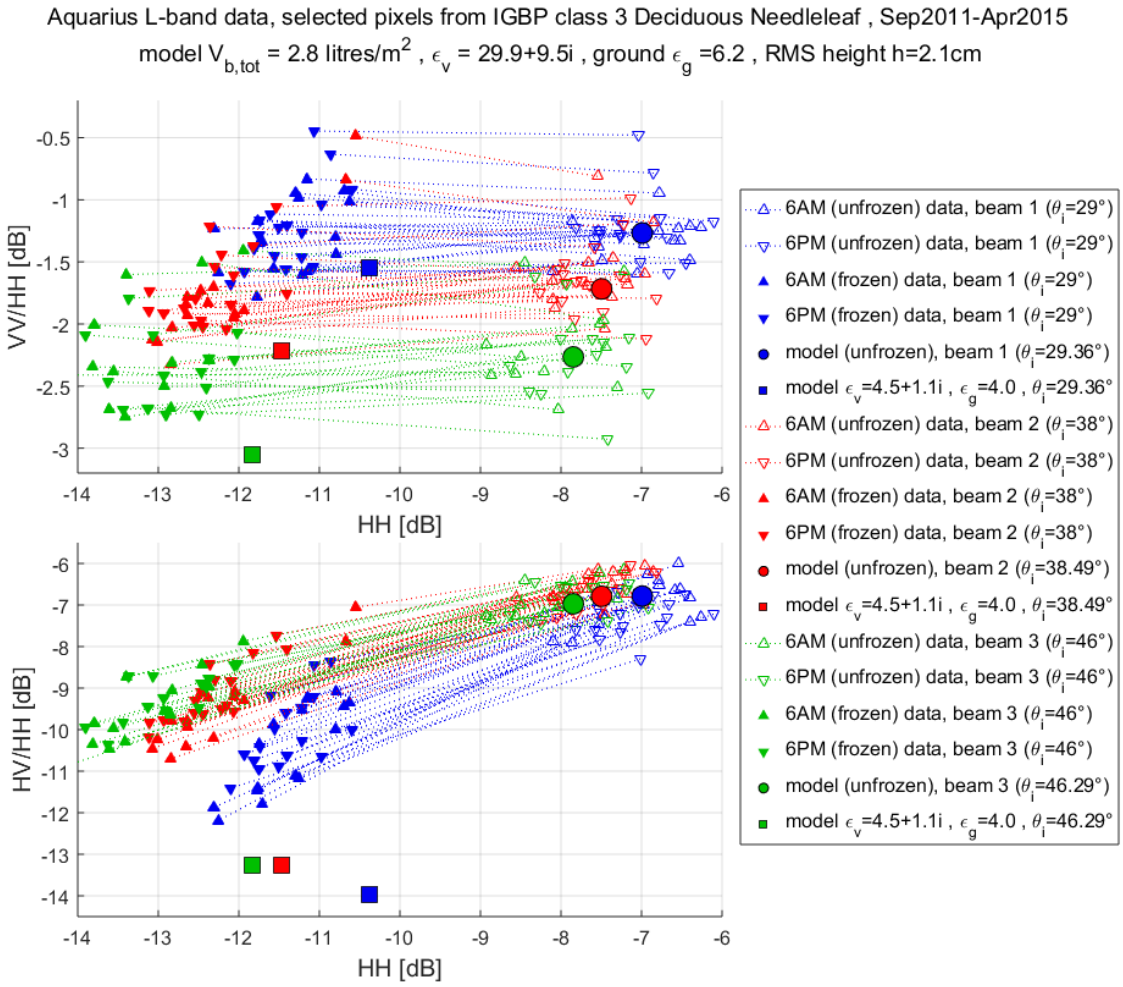

Figure 3.39. Same as Figure 3.38, but for IGBP class 3 (deciduous needleleaf forests).

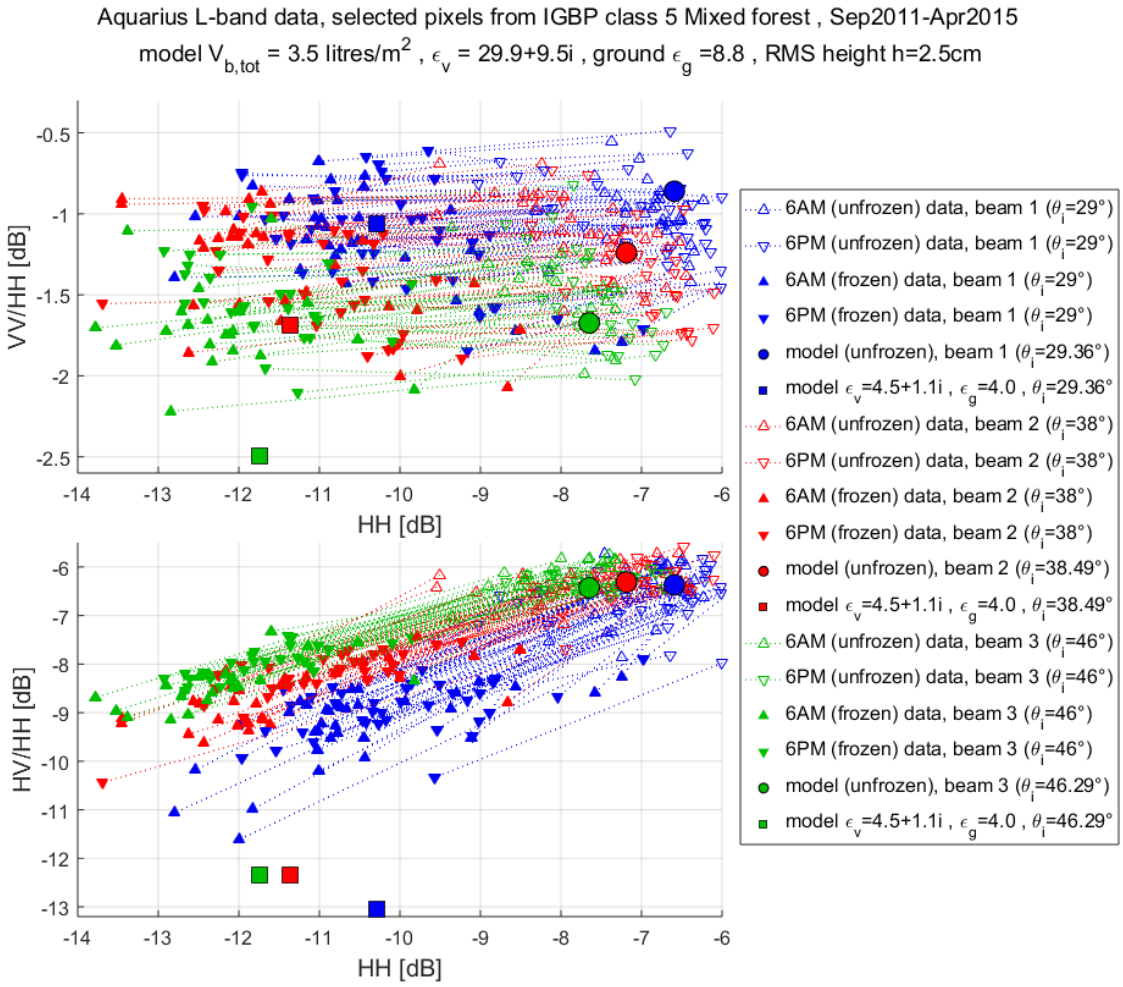

Figure 3.40. Same as Figure 3.38, but for IGBP class 5 (mixed forests).

3.9 Microwave vegetation optical depth

The microwave vegetation optical depth (VOD) is often a parameter of interest for both active and passive microwave studies of vegetated land areas, in applications where the vegetation state or underlying soil moisture are of interest. In this section, we report the VOD values from our forward model using the input parameters as described in Sections 3.3-3.7.

Equation (2.9) from Chapter 2 introduced the simplified average one-way optical thickness of the canopy layer as

$$\tau_{cn}(\theta_i) = \frac{n_{cn}\big(\langle \kappa_{h,cn}(\theta_i)\rangle + \langle \kappa_{v,cn}(\theta_i)\rangle\big)(Z_2 - Z_1)}{\cos\theta_i}. \qquad (3.5)$$

Strictly speaking, the one-way optical thickness of the canopy layer is polarization dependent if the orientation distribution of the cylinders is not uniform:

$$\tau_{h,cn}(\theta_i) = \frac{2n_{cn}\langle\kappa_{h,cn}(\theta_i)\rangle(Z_2 - Z_1)}{\cos\theta_i} \tag{3.6a}$$

$$\tau_{v,cn}(\theta_i) = \frac{2n_{cn}\langle\kappa_{v,cn}(\theta_i)\rangle(Z_2 - Z_1)}{\cos\theta_i} \tag{3.6b}$$

and likewise for the trunk layer:

$$\tau_{h,trk}(\theta_i) = \frac{2n_{trk}\langle\kappa_{h,trk}(\theta_i)\rangle Z_1}{\cos\theta_i} \tag{3.7a}$$

$$\tau_{v,trk}(\theta_i) = \frac{2n_{trk}\langle\kappa_{v,trk}(\theta_i)\rangle Z_1}{\cos\theta_i}. \tag{3.7b}$$

A more common convention in the literature has the definition of optical depth corrected by the factor of $\cos\theta_i$. In this convention, the total microwave vegetation optical depth (summing contributions from canopy and trunk layers) from our model is

$$VOD_{h,tot}(\theta_i) = \Big(\tau_{h,cn}(\theta_i) + \tau_{h,trk}(\theta_i)\Big)\cos\theta_i \tag{3.8a}$$

$$VOD_{v,tot}(\theta_i) = \Big(\tau_{v,cn}(\theta_i) + \tau_{v,trk}(\theta_i)\Big)\cos\theta_i\,. \tag{3.8b}$$

These values are tabulated in Table 3.1 for our forward model using the input parameters as described in Sections 3.3-3.7 for each of the IGBP classes 1-5.

IGBP class	θ_i	$\tau_{h,cn}$	$\tau_{v,cn}$	$\tau_{h,trk}$	$\tau_{v,trk}$	$VOD_{h,tot}$	$VOD_{v,tot}$
1	28	0.561	0.503	0.086	0.130	0.564	0.552
	39	0.628	0.524	0.112	0.171	0.579	0.544
	49	0.714	0.556	0.143	0.216	0.592	0.533
2	28	1.535	1.535	0.269	0.408	1.572	1.694
	39	1.710	1.710	0.352	0.538	1.614	1.759
	46	1.936	1.936	0.447	0.679	1.647	1.807
3	28	0.500	0.449	0.076	0.116	0.502	0.492
	39	0.560	0.467	0.100	0.153	0.516	0.485
	46	0.637	0.496	0.127	0.193	0.528	0.476
4	28	0.478	0.478	0.084	0.127	0.490	0.527
	39	0.532	0.532	0.110	0.167	0.502	0.548
	46	0.603	0.603	0.139	0.212	0.513	0.563
5	28	0.630	0.565	0.096	0.146	0.632	0.619
	39	0.705	0.588	0.126	0.192	0.650	0.611
	46	0.802	0.624	0.160	0.243	0.665	0.599

Table 3.1. θ_i-uncorrected optical thicknesses and θ_i-corrected Vegetation Optical Depths (VOD) from L-band radar backscatter model for IGBP classes 1-5.

From Table 3.1, our model assigns total microwave VOD values of about 0.5-0.6 for IGBP forest classes 1,3,4,5 at both h-pol and v-pol; most of the VOD is contributed by the canopy layer. The exception is IGBP class 2, evergreen broadleaf forests, which are much thicker, at VOD values of about 1.6-1.7. Konings et al. [98] used Aquarius radiometer data to estimate global VOD values, obtaining values of about 1.1-1.3 for tropical jungles and between 0.5-0.9 for other corresponding forest areas, with some variance depending on algorithm used. Rötzer et al. [99] extended work by Konings et al. [98] by regressing Aquarius scatterometer HV against the radiometer-derived VOD to estimate VOD from scatterometer HV, obtaining similar values, but slightly lower by about 0.1-0.2. VOD values derived from the L-band radiometer of the Soil Moisture and Ocean Salinity (SMOS) satellite with the aid of Leaf Area Index (LAI) auxiliary inputs by Kerr et al. [100] and Rahmoune et al. [101] are about 1.0-1.2 for Amazon jungles and between 0.7-0.9 for summertime boreal forests. Shortcomings in our model for IGBP class 2 may be part of the reason for our significantly higher VOD values in comparison to these other estimates.

Chapter 4

DIURNAL EFFECTS ON L-BAND RADAR BACKSCATTER OVER GLOBAL FORESTS USING SMAP

4.1 Chapter introduction and overview

Given the sensitivity of microwaves to vegetation water content via the dielectric permittivity of vegetation, and the diurnal cycle of transpiration and related processes, the potential of radar remote sensing as a tool for monitoring vegetation water status has long been recognized and studied since Ulaby and Batlivala [23], Brisco et al. [24], Way et al. [25], Weber and Ustin [26]; see also Steele-Dunne et al. [27] for a more comprehensive review of the literature. Indeed, diurnal fluctuations in radar backscatter from vegetation have been observed in various reported experiments, with varying degrees of evidence for association with vegetation water status and the specific mechanisms driving the radar observations – these shall be discussed in greater detail in Section 4.4. However our understanding of the relationship between vegetation water status and radar backscatter, in particular with regards to the diurnal fluctuations (since many variables change on a diurnal basis), is far from complete.

One might expect L-band radar, with a wavelength close to being on resonance with tree branches, to contribute valuable information towards this understanding. In Chapter 3, we had analyzed L-band multi-polarization radar backscatter over global forests, using data from the spaceborne Aquarius scatterometer and our forward model from Chapter 2. One issue that surfaced in Chapter 3 was that though local time 6AM and 6PM data were available from Aquarius descending and ascending passes, they were seldom over the same footprints, and were at different azimuth angles. If there are differences in radar backscatter due to beam azimuth angle, this might confound possible inferences from Aquarius about diurnal effects. Another L-band spaceborne mission, SMAP (Soil Moisture Active Passive), has a conically scanning antenna, allowing us to separate these effects. SMAP would have continued the

global L-band radar coverage after Aquarius, but unfortunately the radar was lost after a few months of operation. Nevertheless, we shall use whatever available data exists.

The sections in this chapter will be organized as follows. Section 4.2 will introduce the SMAP dataset to be used for our analysis. Section 4.3 will show that there are indeed significant differences due to the beam azimuthal angle, possibly associated with the orientation of vegetation due to plant phototropism. Section 4.4, before presenting results concerning diurnal fluctuations observed by SMAP, would also review in greater detail some of the relevant results from the literature. One of the results in this chapter is that the co-pol L-band radar backscatter observed in late spring-summer over the northern boreal forests was higher at 6PM than 6AM, which is not what one might expect based on previous studies. Based on our modelling, increased canopy extinction at 6AM was a possibility, but this is unproven and its true underlying physical cause undetermined.

4.2 SMAP dataset

The SMAP mission [102] was launched in 2015 with the primary goal of measuring soil moisture from space. The spacecraft orbit is sun-synchronous (local 6AM descending node equator crossing) at about 685km altitude, with global coverage in 2-3 days on an 8-day repeat orbit. The instrument carried an L-band radiometer operating at 1.41GHz and an L-band radar at 1.22-1.3GHz, both sharing a rotating reflector antenna in a conically scanning configuration; the local incidence angle on the ground is about 40°. The radar measured normalized radar backscattering cross-sections HH, HV, VH, and VV, with real-aperture footprints of 30-40km (the so-called “low-res” mode). SMAP had a high-resolution SAR mode that had a variable azimuth resolution due to the rotating antenna, but in this study we shall only use the real-aperture (“low-res”) radar data. Further instrument details can be found in the SMAP handbook [4] and references therein; in particular, Figure 19 from the SMAP handbook shows a schematic of the conically scanning configuration, reproduced here in Figure 4.1.

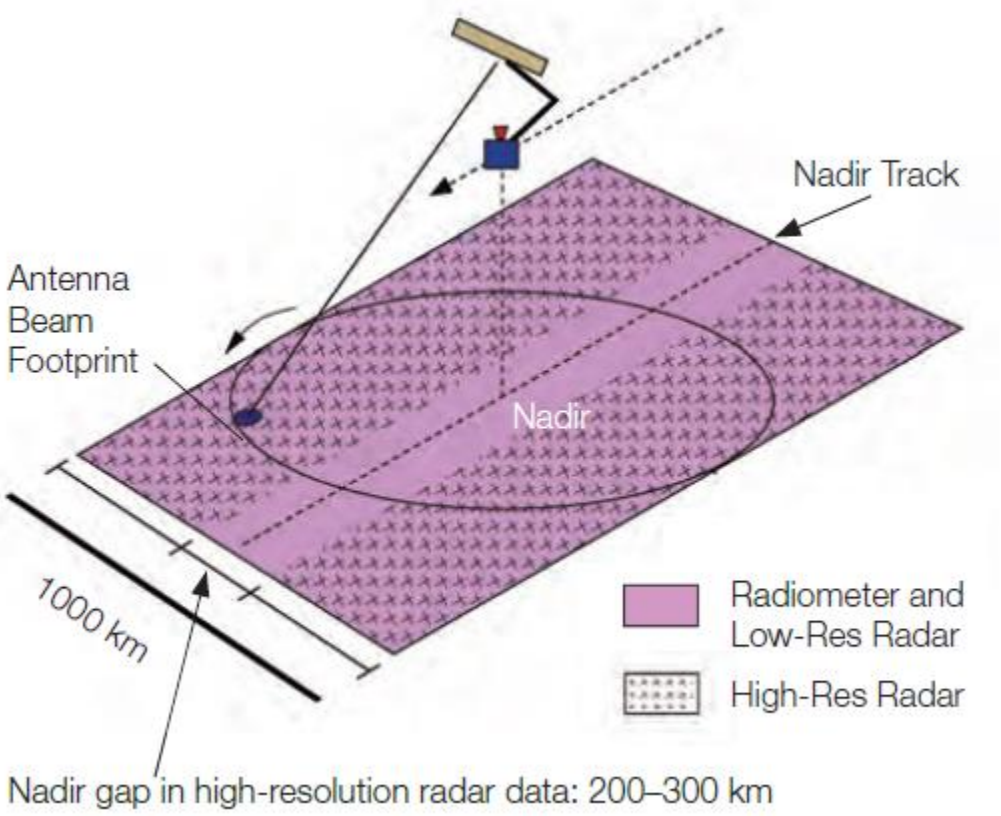

Figure 4.1. The conically scanning configuration of SMAP, as shown in Figure 19 of the SMAP Handbook [4]. Credit: NASA 2014.

We used the SMAP Level 1B Radar data product version 3, obtained from NASA EOSDIS Alaska Satellite Facility (ASF) [103, 104, 105]. The data record ended on July 7, 2015 due to an unexpected failure of the SMAP radar. Using the 36km EASE-Grid 2.0 [71], the data was regridded according to whichever EASE2 grid pixel the centre of its radar footprint was nearest to. The data was also sorted by beam azimuth angle into 8 bins each 45° wide, centred on the directions north, northeast, east, southeast, south, southwest, west, and northwest. Wherever there was more than one data value corresponding to a pixel or azimuth bin, data was summarized by taking the median value. Data was then sorted by local time instead of UTC; by design of the sun-synchronous orbit, the local time was either approximately 6AM (descending pass) or 6PM (ascending pass). These rearrangements yielded 406x964x8x87 arrays of backscatter values for 6AM and 6PM data at each polarization HH,HV,VH,VV, for

87 days of data from April 12, 2015 to July 7, 2015; not all values in these arrays are filled because not every point on the globe was revisited every day. The cross-polarized backscatter values are further summarized by taking the average of HV and VH values, henceforth simply referred to as HV.

As we already know that changes in freeze-thaw state give a large contrast in radar backscatter, while we are now instead interested in studying the smaller variations during unfrozen periods, it is thus necessary to exclude data taken under frozen conditions from our analysis. The SMAP Level 3 Freeze-Thaw Passive Product was used for this purpose [106]. This product uses the passive radiometer measurements from SMAP to flag the freeze/thaw status of each EASE2 grid pixel. To be conservative, we masked out data from any pixel on a given date as long as it had a "frozen" flag within two weeks of that date, and also (arbitrarily) restrict most of our analysis of the dataset to begin from May 17, 2015 to minimize risk of data contamination from frozen conditions in the Northern hemisphere.

Using the SMAP dataset and with the afore-mentioned processing steps, a global false color map of the radar backscatter over land is displayed in Figure 4.2, by taking pixel-wise median values over the entire time period and all beam azimuth angle bins. As described in the figure caption, information from the polarization ratios VV/HH and HV/HH are mapped to a color wheel (lower left corner of Figure 4.2) while the HH[dB] normalized backscatter cross-sections are mapped to brightness values. Figure 4.2 can be compared with the MODIS land cover classification map (IGBP) of Figure 3.5. As expected, deserts and sparsely vegetated areas have low backscatter and thus appear dark, or dark blue/purple due to low HV/HH ratios and the stronger VV than HH of surface backscatter. On the other hand, thickly vegetated areas appear bright. The different polarization signatures from the different forest classes reveal themselves clearly, e.g. the tropical jungles in the saturation limit have the highest HV/HH ratio and are colored bright green. Coniferous forests with significant preferential horizontal orientations of branches and double-bounce scattering contributions have lower VV/HH ratios, thus appearing red; other forest or savanna regions where backscatter is not saturated by the canopy layer would also appear red due to the stronger

HH than VV from double-bounce scattering contributions. Mountainous areas have HH and VV backscatter approximately equal and thus appear as a pale color or some shade of gray.

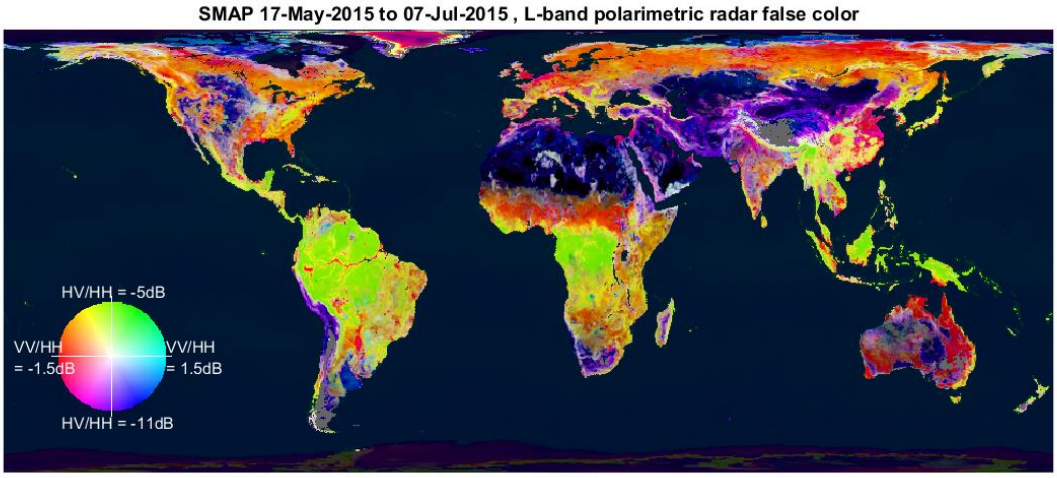

Figure 4.2. False-color map of the world (land only) on an EASE2 projection grid using L-band radar backscatter from SMAP data. For each pixel, median values were computed for HH[dB], VV/HH[dB], HV/HH[dB] from May 17, 2015 to July 7, 2015 and over the 8 beam azimuth direction bins, with frozen conditions excluded as described in the text. These median HV/HH[dB] values and VV/HH[dB] values are mapped via a polar coordinate transformation to the color wheel depicted the lower left corner. The color wheel saturates to red at VV/HH values of -1.5dB, cyan at VV/HH values of +1.5dB, light green at HV/HH values of -5dB, and purple at HV/HH values of -11dB. The origin of the color wheel (gray colors) is at HH=VV and HV/HH=-8dB. Median HH[dB] values are mapped to brightness values, with maximum brightness at values of HH>=-5dB, and minimum brightness (i.e. black) at values of HH<=-30dB. Pixels with no unmasked data (due to frozen conditions, data quality, etc.) are colored gray (e.g. Himalayas).

4.3 Beam azimuth effects

Using the SMAP data binned by azimuth angle as described earlier, we can make comparisons about the radar backscatter taken from the same location at almost the same time. Due to the 98° inclination (i.e. the flight path is almost north-south when far away from the poles) of the sun-synchronous orbit and the conically scanning configuration, there are many more opportunities for near-coincident (both spatially and temporally) comparisons of north-pointing vs south-pointing beams than other pairings. For pairings that are not opposite in direction, the alignment direction of mountain ranges tend to provide obvious differences,

due to foreshortening, shadowing, layover, etc. Comparing opposite direction pairings help to cancel these effects.

Figure 4.3 and Figure 4.4 display the ratio of radar backscatter between the north- and south-pointing beam, for HV polarization. Specifically, the median (over the time period May 17, 2015 to July 7, 2015) of data (in dB) from the south-pointing beam azimuth bin is subtracted from the median for north-pointing beam data (in dB). Figure 4.3 is for descending pass data only (local time 6AM), while Figure 4.4 is for ascending pass data only (local time 6PM). Red areas show higher radar backscatter from the north-pointing beam, while blue areas show higher radar backscatter from the south-pointing beam.

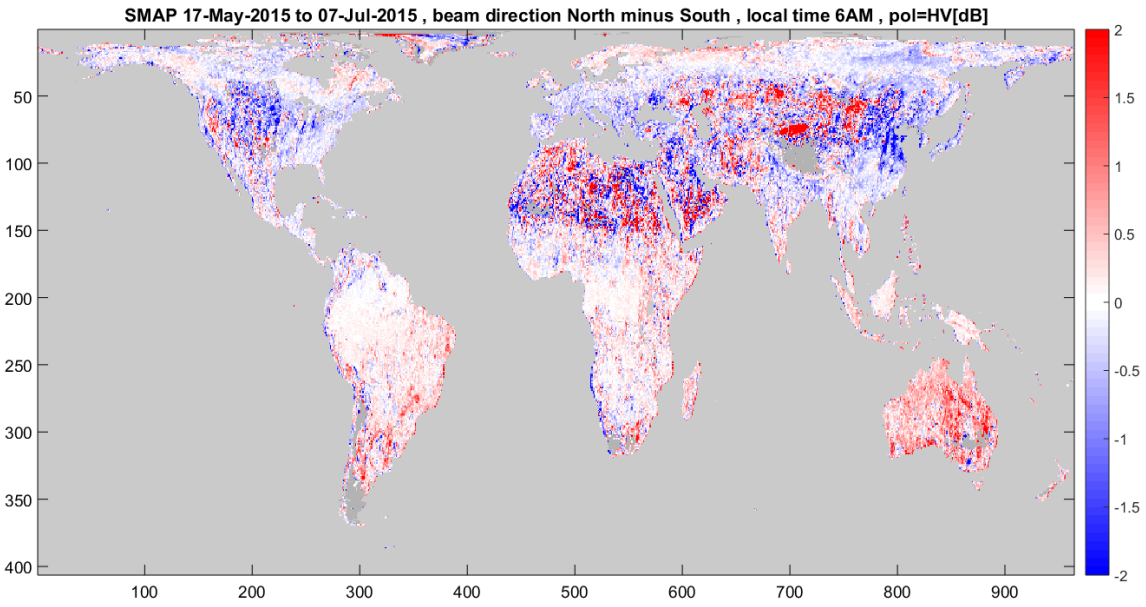

Figure 4.3. Differences in SMAP L-band radar backscatter due to beam azimuth angle. The median value of data (in dB) from the south-pointing beam azimuth bin is subtracted from the median value of data (in dB) from the north-pointing beam azimuth bin, for HV polarization, descending passes (local time 6AM). Red areas show higher radar backscatter from the north-pointing beam, while blue areas show higher radar backscatter from the south-pointing beam; note the color scale on the right in dB.

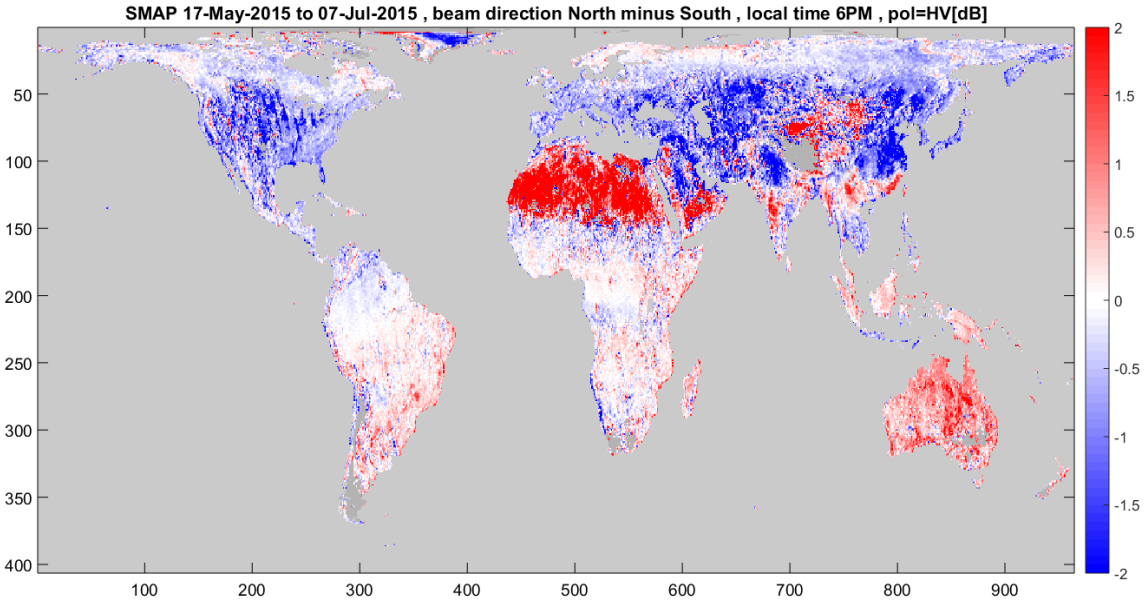

Figure 4.4. Same as Figure 4.3, but for ascending passes (local time 6PM).

A glance at Figure 4.3 and Figure 4.4 reveals that there can be significant differences in radar backscatter due simply to the azimuth angle of the radar beam. The differences in azimuth angle are generally more pronounced at local time 6PM than 6AM, and these differences are roughly in the same direction (i.e. similar color at 6PM and 6AM), with some exceptions. The largest fractional differences (i.e. largest absolute differences in dB) are over the Sahara and the deserts and steppes of Asia, but note that in some desert regions the HV backscatter values can be so low (<-30dB or even -40dB) that great caution is required for interpretations (ionosphere Faraday rotation and/or its correction artifacts may become more visible in the afternoons). Stephen [107] and Stephen and Long [108] showed that in broad sand dune fields called ergs, which are formed and constantly reshaped by prevailing winds, the geometry and alignment of their large- and small-scale surface features (dunes and ripples) may be responsible for azimuthal differences. For most of the vegetated areas, we see a broad distinction between the Northern hemisphere and the Southern hemisphere. In the Northern hemisphere, there is higher radar backscatter when the radar beam points south than when the radar beam points north, with the converse for the Southern hemisphere (there are some exceptions e.g. in parts of India and southern China). The differences in dB are larger over

the grasslands, steppes, and agricultural areas, than for the forests. There are also azimuthal differences for HH and VV polarization (not displayed), but the differences are smaller and the patterns less obvious.

For the global SMAP observations, a possible explanation we propose to explain many of the broad patterns is vegetation alignment due to the sun, i.e. phototropism. A preferential orientation of cylindrical vegetation structures towards the direction of the sun – southwards in the Northern hemisphere, and northwards in the Southern hemisphere – could present greater radar cross-sections in a direction perpendicular to the preferential alignment. With SMAP radar incidence angles of about 40° from the vertical, in the Northern hemisphere, a northward-pointing beam would be "more parallel" to the vegetation cylinders while a southward-pointing beam would be "more perpendicular"; and conversely for the Southern hemisphere.

Bartalis et al. [109] had also studied azimuthal differences, using the 13.4GHz scatterometer on QuikSCAT, focusing on strong local azimuthal anisotropies, identifying local topography and other human development features as some of the causes. As beam azimuthal effects are not the primary focus of this study, a deeper investigation of the causes of azimuthal anisotropy in the global L-band radar backscatter is not pursued at this moment and left as the subject of future work. The primary purpose here is to illustrate these azimuthal effects as a confounding factor if we had used Aquarius data alone for diurnal investigations. For the purposes of the diurnal analysis in the next section, the radar backscatter values used shall be the median values taken over the eight azimuthal angle bins.

4.4 Diurnal effects

This section studies the diurnal variation in L-band radar backscatter as seen by SMAP. Though we expect this diurnal signal to be related to vegetation water status, there is not currently a very clear picture in understanding the full relationship, as seen from a review of some previous experiments reported in the literature, summarized below.

Brisco et al. [24] in experiments on a wheat field observed higher radar backscatters at night and peaking at dawn, with lows in the afternoons, for C and L bands at HH, VV, and HV

polarizations, but whether these were more correlated with soil or vegetation moisture depended on the age of the plants.

Sparks et al. [110] found that for lodgepole pines, volumetric water content in the sapwood during April-October were in the range 0.45-0.55m^3/m^3, higher in early April following snowmelt, and then relatively constant from June-October, with little relationship to wood temperature. Diurnal variations (peaking at night) were approximately 0.02m^3/m^3 during periods of active transpiration, and significant precipitation could lead to an increase by 0.05-0.06m^3/m^3 for 2-3 days.

Experiments on walnut trees over 2-3 days in August 1987 [25, 26, 111] measured a large diurnal fluctuation in the L-band dielectric constant, peaking in the early morning at around 6AM and reaching a minimum around noon and through the afternoon; these times and dielectric values had dependence upon the depth of the measurement probe into the tree. L-band radar backscatter at 55° incidence showed response to both the walnut trees and the soil moisture, with HH and HV appearing more sensitive to the soil moisture, while HV and VV seemed more sensitive to the vegetation; dips of between 0.5-1.5dB were observed each afternoon (but there was also daily irrigation, which confounded the diurnal signal somewhat). C-band HV and VV backscatter however dipped in the morning and peaked in the afternoon, while C-band HH peaked in the morning and dipped in the afternoon.

McDonald et al. [112] conducted experiments on two Norway spruce trees, measuring dielectric constants of the xylem. Note that the xylem layer, though deeper, could be significantly thicker than the outer thinner phloem and cambium layers where the tree bole dielectric constants peaked. Temporal variations in the xylem dielectric constant at P-band was typically inversely correlated with vapor pressure deficit (VPD, the partial water vapor pressure difference between free water in the plant leaf tissue and the atmosphere, a quantity closely related to evapotranspiration), but when VPD was low e.g. after rain or during cloudy days, the correlation with dielectric constant was lost. There was significant variation between different trees in the same microclimate and different parts of the same tree. Also, there could be as much as a few hours lag in the real and imaginary parts of the dielectric

constant, relative to each other, as well as relative to the VPD. The dielectric constant fluctuations were smaller than for the walnut trees.

Monteith and Ulander [113], in an experiment at a stand of mature Norway spruce showed that wind could be a factor in diurnal variations in boreal forests; high winds may bend the vertical stems of trees and increase HV backscatter. Rainfall sometimes caused increase in HH at L-band, but little change at P-band. VPD and L-band radar backscatter were positively correlated.

On a global scale, Friesen et al. [85] found using ERS-1/2 scatterometers (5.3GHz VV polarization) on a sun-synchronous orbit that 10:30AM passes had significantly higher backscatter than 10:30PM passes by 1-2dB particularly in vegetation-covered areas. In some regions, the seasonality of these diurnal differences do not follow the annual soil moisture cycle, but instead is strongest during the transition period between wet and dry seasons; Friesen identified this as due to a shift in the diurnal cycle of vegetation water content in response to water stress [27]. Paget et al. [114] analyzed QuickSCAT and RapidScat data (13.4GHz co-pol scatterometers) and found a median diurnal backscatter variation of 1.05dB globally, with details depending on region; in particular, the Amazon and Congo had variations of 0.5-0.8dB with a maximum at 6AM and minimum at 6PM, while the Upper Danube region had 1.2dB variations with a maximum at midnight and minimum at noon, observed by RapidScat only, but not QuikSCAT. The Amazon diurnal variations were also observed by Satake and Hanado [94] using TRMM (Ku-band, HH polarization) radar data; they proposed morning dew as a possible cause.

From the literature review, a full consensus does not readily emerge, but we might broadly expect that: transpiration can deplete water stored in plants through the day, reducing vegetation water content and dielectric constant, which are then replenished over the night through root water uptake. Thus L-band backscatter values over vegetated areas in general might be expected to be higher at 6AM than 6PM, in particular for HV polarization. Let us now check these expectations against the SMAP data.

Global maps of the ratio in radar backscatter (i.e. difference in dB) between the ascending (local time 6PM) and descending (local time 6AM) passes are displayed in Figure 4.5 (HH), Figure 4.6 (VV), and Figure 4.7 (HV). Specifically, for each polarization, the median over the time period (17May2015-07Jul2015) and 8 azimuth angle bins was computed for ascending and descending pass data separately to yield two global radar backscatter maps in dB, and then the descending pass map (local time 6AM) subtracted from the ascending pass map (local time 6PM). Higher backscatter in the evening is shown in red, while lower backscatter in the evening is in blue (refer to figures for color scale of the dB difference values).

Several initial remarks can be made regarding the global diurnal maps. Regarding the large fractional differences (i.e. large absolute differences in dB) in HV over the Sahara, the very low absolute levels of backscatter (less than -30dB, or even on the order of -40dB) should be borne in mind; many other sources of noise may be coming into play. Also, examination of the data over the oceans reveals that some ionosphere Faraday rotation and/or its correction artifacts may still be visible. For the vegetated areas, as seen in all three polarization maps Figure 4.5 to Figure 4.7, the intuition built from the literature that transpiration-induced reductions in daytime vegetation water content cause lower backscatter values at 6PM than 6AM seems to indeed be roughly verified over many areas – the thinner vegetation grasslands, steppes, and croplands of North America and Asia, the savannas of Africa and South America, and dense jungles of the Amazon. That HV changes more than the co-polarized returns is also consistent with these diurnal changes being vegetation-driven. Even though we have not presented a quantitative forward model for non-forested areas in Chapter 2 and Chapter 3, we carry the qualitatively intuition from our forest model that, where there is sufficient vegetation cover (yet not too much as to be in the saturation limit) for double-bounce mechanisms to be important, positive correlation between HV/HH fluctuations and HH fluctuations are suggestive of vegetation changes, while negative correlation between HV/HH fluctuations and HH fluctuations are suggestive of ground factors.

However the majority of the boreal forests of the Northern hemisphere show a surprising result, with higher backscatter at 6PM than 6AM, particularly at HH and VV polarization,

less so at HV. This seems to contradict what we would have expected from transpiration-induced water stresses in the afternoon. It is in concordance only with some of the measurements by Monteith and Ulander [113] that also showed higher instead of lower backscatter in the afternoon. Looking back at the Aquarius data (e.g. Figure 3.12, Figure 3.26 and similar figure for IGBP class 5 not shown), though possibly confounded by azimuthal effects, we see a similar result. The geographic specificity also suggests that this is not merely an artifact or instrument bias.

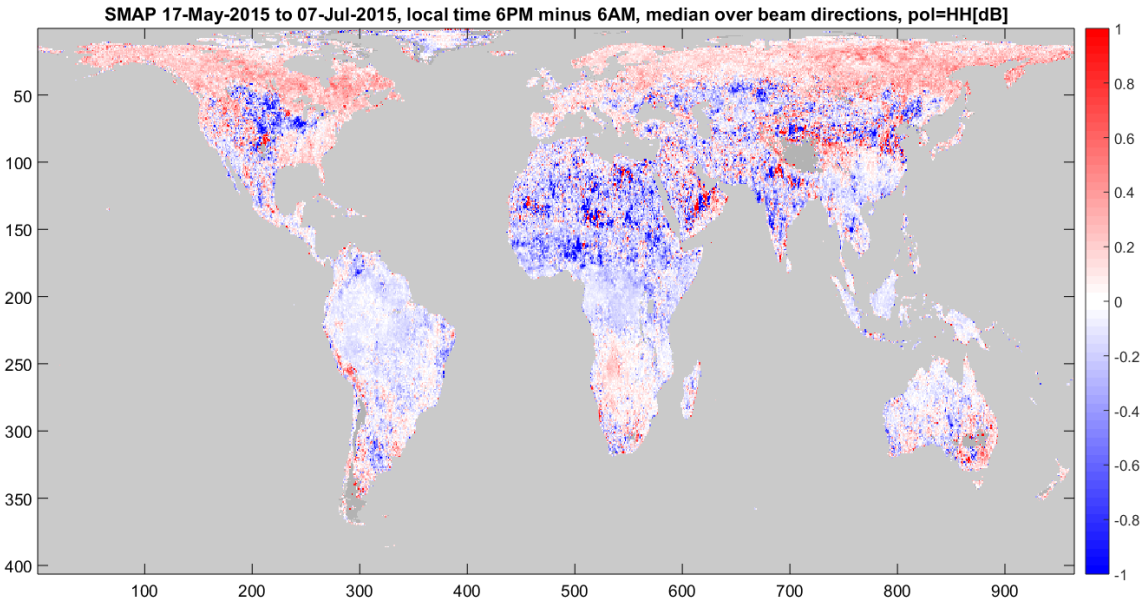

Figure 4.5. SMAP global map of ratio (difference in dB) in L-band radar backscatter between evening (local time 6PM) vs. morning (local time 6AM), at HH-polarization, taking median value over azimuth angles and the time period 17May2015-07Jul2015. Higher backscatter in the evening is shown in red, while lower backscatter in the evening is in blue. On the right is the color scale for the dB difference values. Note the broad zone of red in the northern boreal forests.

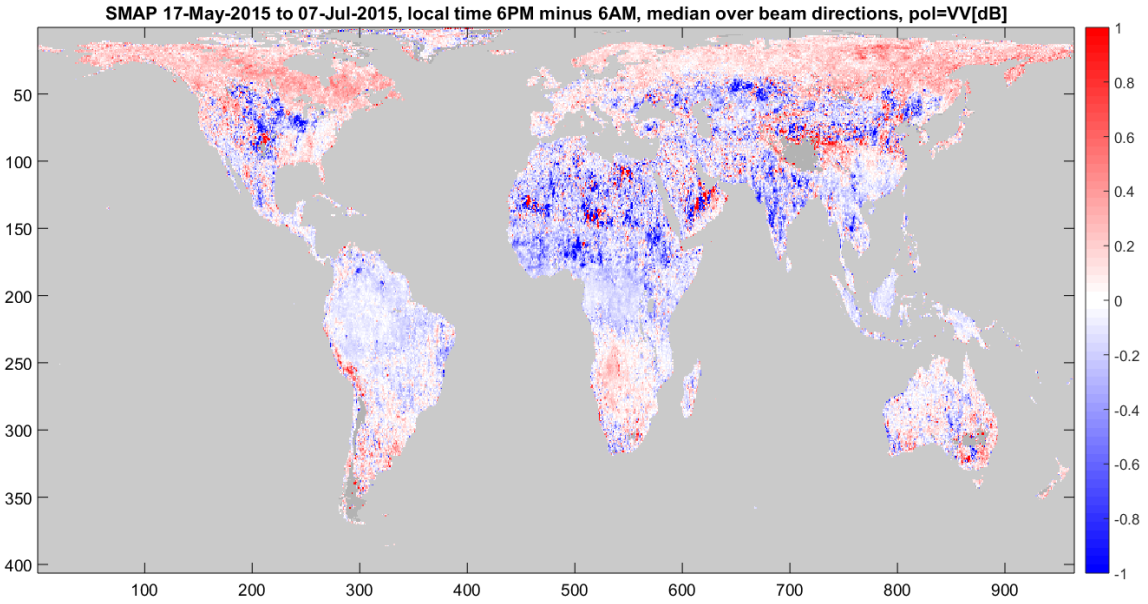

Figure 4.6. Same as Figure 4.5, but for VV polarization.

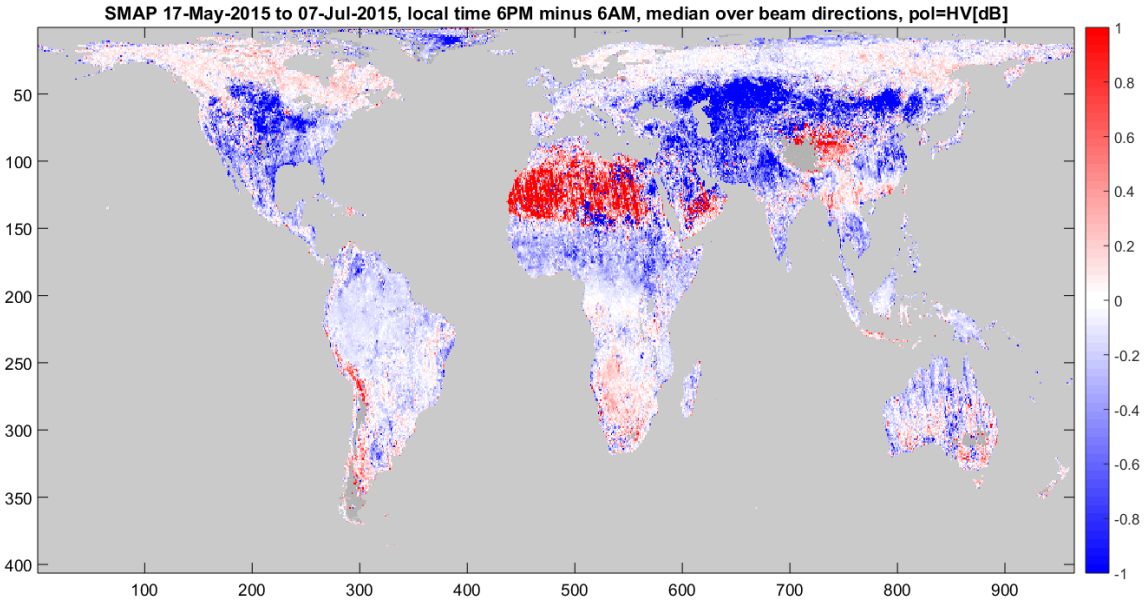

Figure 4.7. Same as Figure 4.5, but for HV polarization.

This surprising observation, over a majority of the area in the Northern hemisphere boreal zone, motivated further analysis. In an attempt to gain further insight, we take a look at the

time-series for the SMAP L-band radar backscatter for several example pixels. 5 pixels were chosen, labelled A, B, C, D, E, with their geographic locations displayed in Figure 4.8. Figures 4.9a to 4.9e display the respective SMAP radar time series for these five pixels, alongside precipitation information. The NASA Global Precipitation Measurements (GPM) was used for the precipitation information; specifically, the half-hourly Integrated Multi-satellitE Retrievals for GPM (IMERG) data product version 05B [115] was used, and further binned into hourly precipitation rates. In Figures 4.9a to 4.94.9e, all the time series share the same horizontal axes, which is in local time. 6AM backscatter data is plotted as blue upright triangles, while 6PM backscatter data is plotted as red inverted triangles. Data before May 17, 2015 (that we had somewhat arbitrarily chosen as our cutoff date) is colored black.

Though only for five pixels, a number of observations from these examples are noteworthy. Firstly, from the data before May 17, where the scene is frozen at 6AM but thawed at 6PM, as discussed in Section 3.8, the backscatter is very low and the diurnal contrast very large, to be distinguished from the smaller signal displayed in Figure 4.5 to Figure 4.7, which had been masked for frozen conditions. The surge in backscatter due to spring melt is also evident, decaying long after May 17 into the summer. From May 17-July 7, the intriguingly higher 6PM vs. 6AM co-pol backscatter being investigated truly seems to be a diurnal affair, instead of being an artifact of a few strong but sporadic differences. This diurnal signal itself also seems to decay along with the decay of the spring melt surge; unfortunately the dataset ended prematurely, preventing observation of whether the diurnal differences last throughout the rest of the summer. For pixel B, there is no discernable diurnal signal from mid-May onwards. Though Chapter 3 identified soil moisture changes as a primary factor for temporal changes in L-band radar backscatter, note the inconsistency of response of radar backscatter to precipitation, across time and across the five pixels; the only consistency is the spring melt decay.

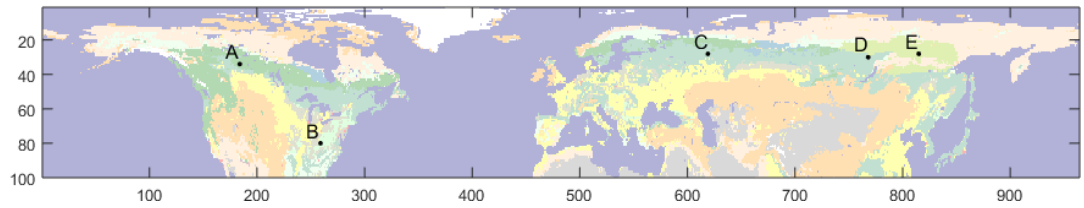

Figure 4.8. World map from latitude 30°N to 84°N on EASE2 grid, showing locations of pixels A, B, C, D, E chosen for display of SMAP L-band radar backscatter time-series in Figures 4.9a to 4.9e.

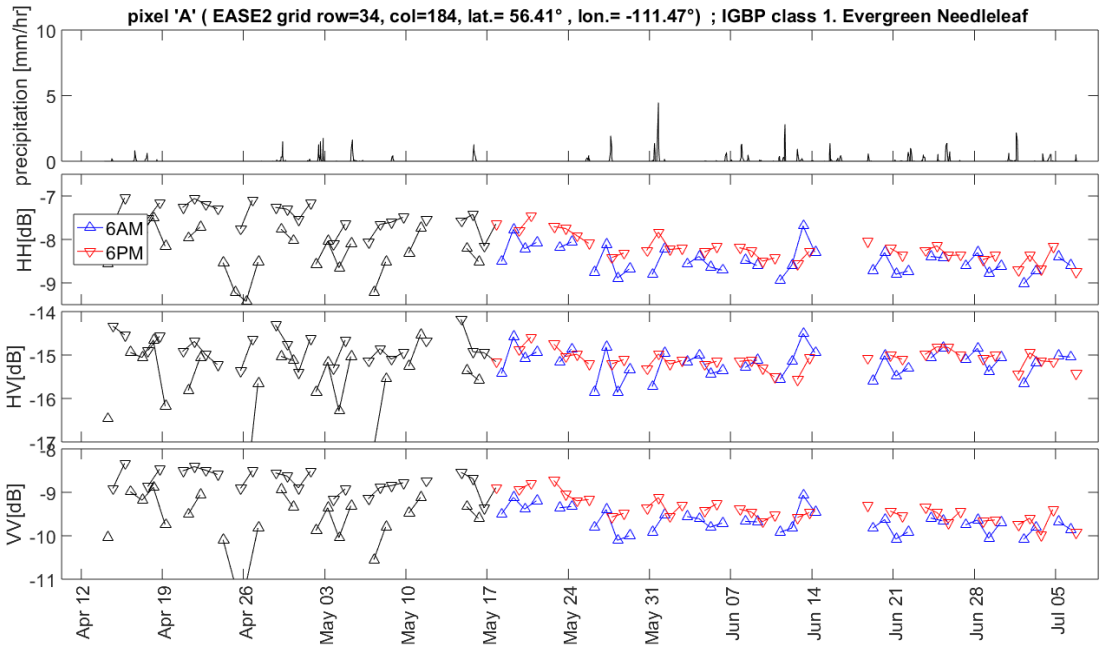

Figure 4.9a. SMAP L-band radar backscatter time-series for pixel A. Upright triangles in blue are 6AM data, inverted triangles in red are 6PM data. Data before May 17 are colored black. Precipitation time-series is hourly, from GPM IMERG data.

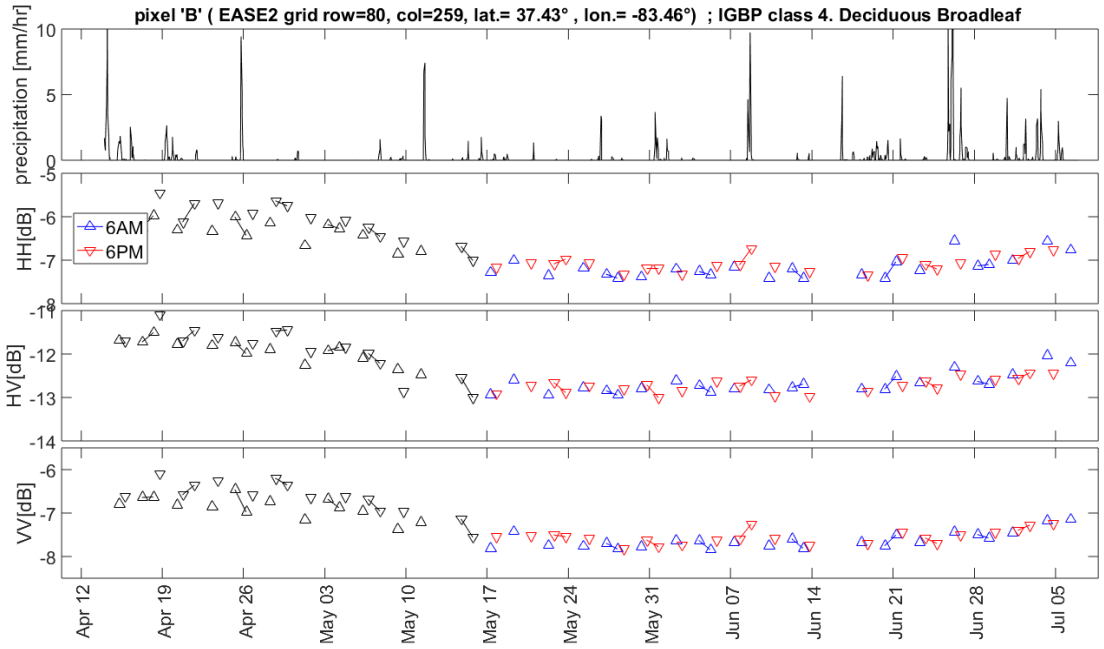

Figure 4.9b. Same as Figure 4.9a, but for pixel B.

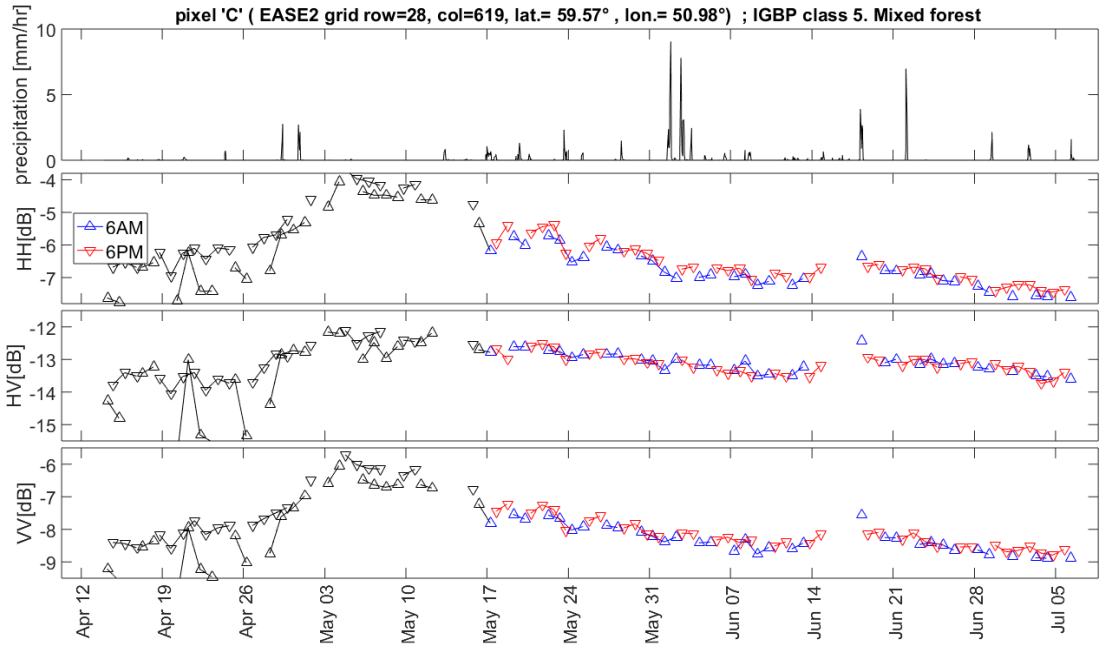

Figure 4.9c. Same as Figure 4.9a, but for pixel C.

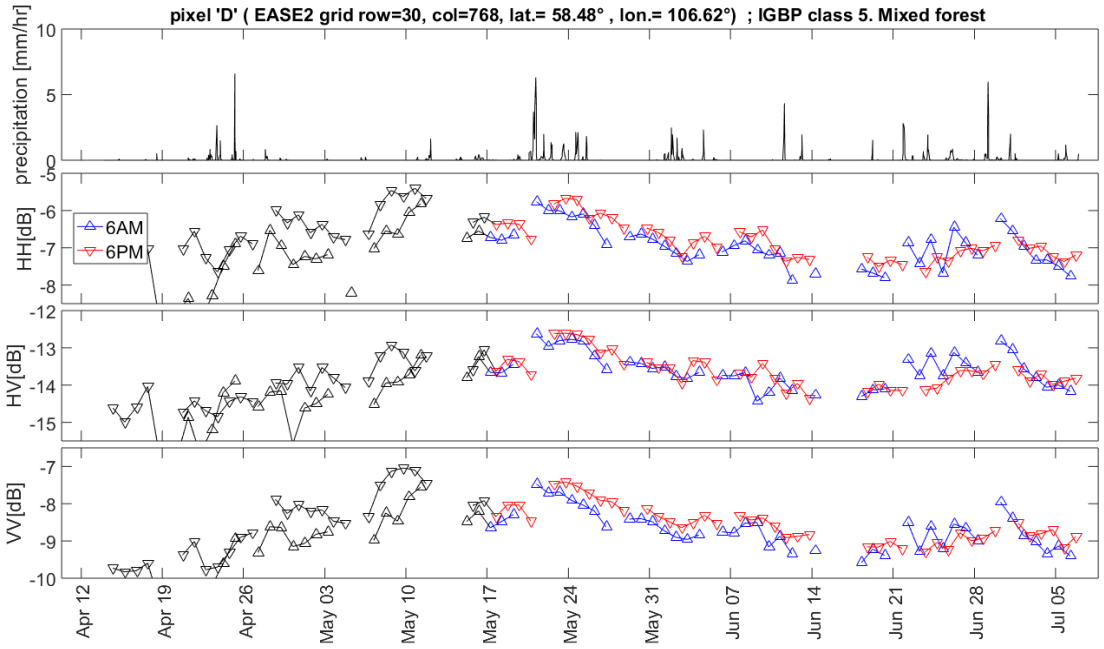

Figure 4.9d. Same as Figure 4.9a, but for pixel D.

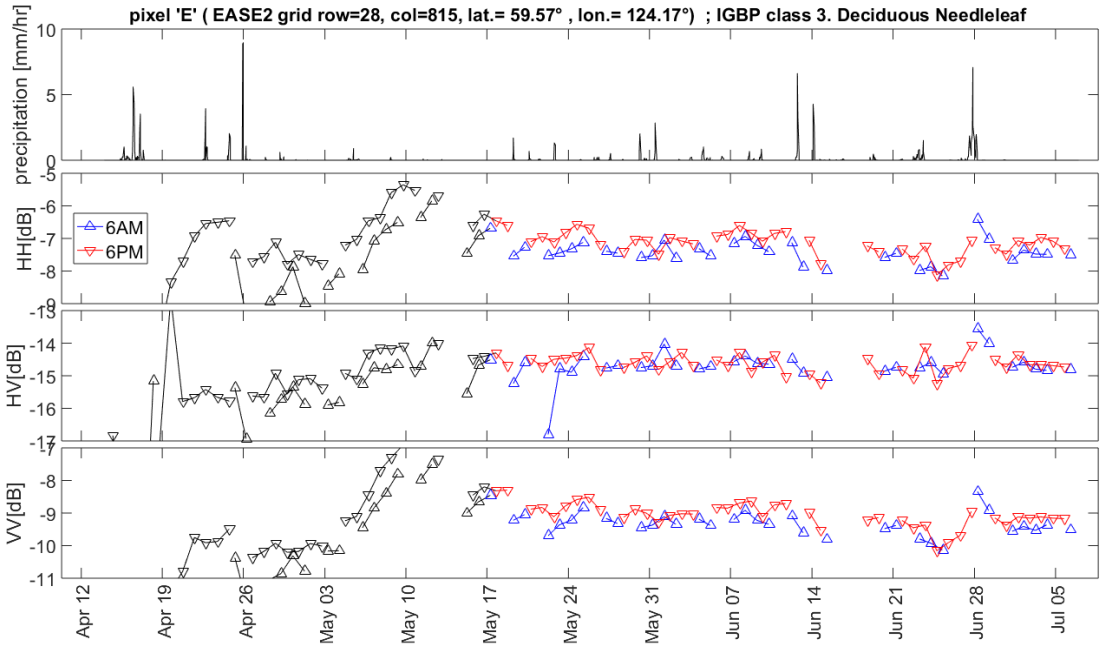

Figure 4.9e. Same as Figure 4.9e, but for pixel E.

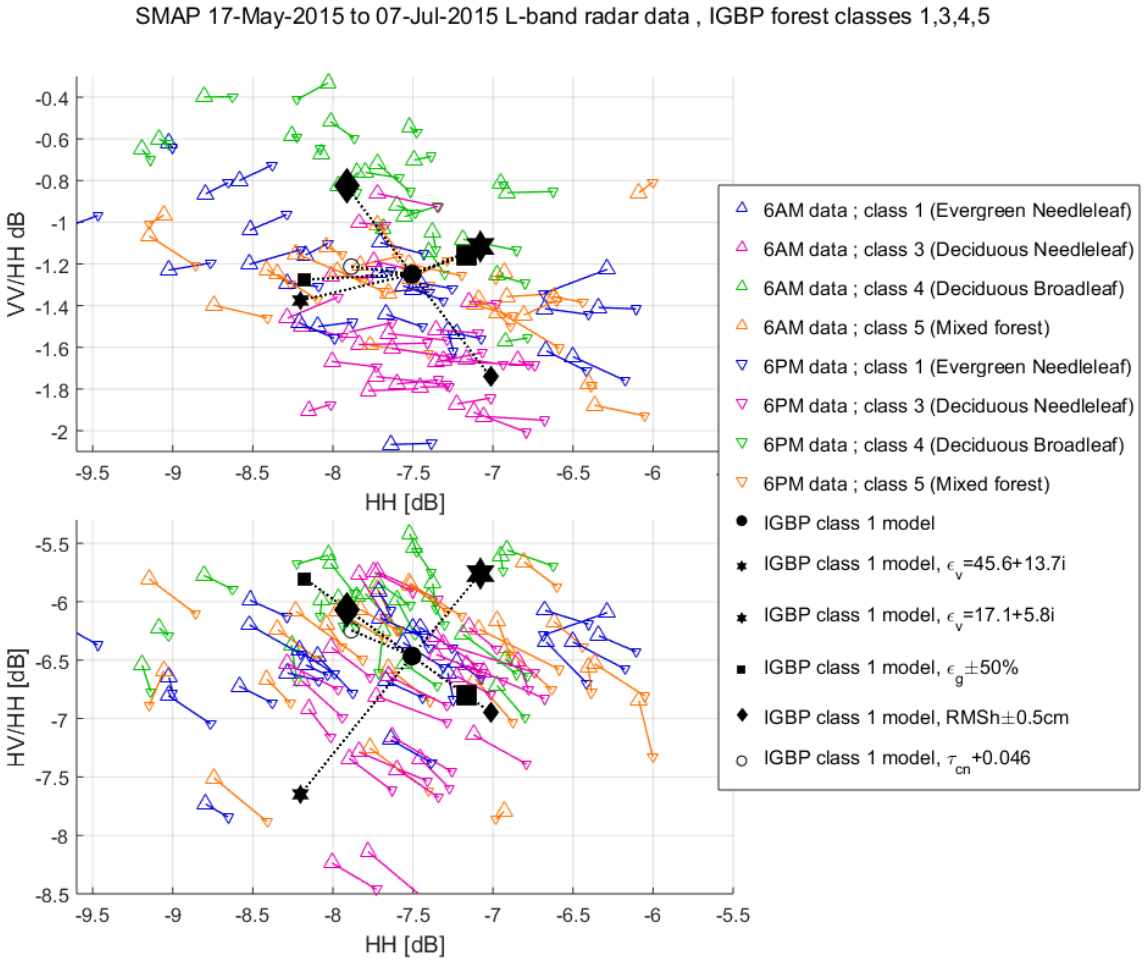

Figure 4.10. SMAP L-band radar data for some EASE2 grid pixels from IGBP classes 1, 3, 4, and 5; 25 pixels from each class, north of 30°N and not of high terrain slope, were chosen randomly for this display. Upright triangles: 6AM data; inverted triangles: 6PM data. Colors blue, magenta, green, and orange represent IGBP classes 1, 3, 4, 5; see figure legend. Each pair of upright and inverted triangles linked by a solid line summarize data from the same EASE2 pixel; median values were taken over the 8 azimuthal angle bins, and over the time period 17May2015-07Jul2015. Filled black circles show the forward model with the input parameters of Section 3.3 (i.e. fitted to the Aquarius data for IGBP class 1). Other filled symbols explore the nearby model parameter space if one parameter is perturbed, as in Section 3.3. The open circle represents the artificial addition of 0.2dB of one-way extinction (0.4dB two-ways) to the canopy layer.

Next, we apply the multi-polarization methods of Chapter 3 to further analyze these diurnal changes; see Figure 4.10. The format of Figure 4.10 is similar to some of the figures in Chapter 3, but for diurnal changes instead of longer-term temporal changes, and combined for IGBP classes 1, 3, 4, and 5. 25 pixels from each class, north of 30°N and not of high

terrain slope, were chosen randomly for this display. At first glance, the diurnal changes resemble the directions of ground moisture changes on the multi-polarization scatterplots as in Chapter 3. However, it seems surprising for soil moisture to consistently be higher at 6PM than 6AM over such a vast geographic region. Bear in mind that frozen conditions had already been excluded from the data being analyzed. Inspection of several examples of in-situ soil moisture time-series did not reveal any consistent significant differences between 6AM and 6PM.

One alternative possibility to consider is the effect of temperature on the dielectric permittivity of soil. Wagner et al. [116] measured the temperature dependence of soil dielectric constant and noted an approximately 5% increase in dielectric constant at 1GHz from 10°C and 20°C for a silty clay loam with 17.4% volumetric water content. This amount of change in ε_g however seems insufficient to explain all the diurnal variations.

Another alternative is that, despite all attempts to exclude frozen conditions from the data being analyzed, within the large radar footprints there remain small isolated pockets of frozen ground that partially thaw out every afternoon and refreeze every night. However, this seems unlikely to be the explanation, because if so, we would expect the total radar backscatter to increase corresponding to the diminishing area of these pockets as the summer progresses.

A third alternative to consider is not changes in the ground, but instead an increase in the extinction of the vegetation canopy. The modelled effect of an increase by 0.2dB of one-way-extinction (0.4dB two-ways) to the canopy layer, in both horizontal and vertical polarizations, is represented by the open circle in Figure 5.4-6. We see that the effect on the HH, HV, and VV backscatter due to this additional extinction bears similarity to that from a decrease in ε_g, because the backscatter contributions from scattering pathways involving the ground are reduced more than for direct backscatter contributions from the vegetation. We have yet to attribute this modelled extinction to an exact physical cause. If it is due to scatterers, they should have low albedo such that they do not contribute significantly as additional scattering. Small scatterers might foot the bill, e.g. an increase in water content and dielectric permittivity of needle-like leaves, or dew/guttation drops on them, at 6AM

compared to 6PM. If it is due to water condensed from the atmosphere, the droplet sizes have to be large enough to significantly increase the extinction efficiency (relative to water vapor). As yet, we do not have sufficient evidence to prove any of these cases, though the impact of leaf moisture primarily through its transmissivity rather than scattering had also been suggested before [27]. There are probably further possibilities that we have failed to consider here. At the moment, there is insufficient conclusive evidence for what the true cause of this curious diurnal signal is.

In this chapter we noted many complications from the picture presented in Chapter 3, which had identified soil moisture as a primary factor for temporal changes in L-band radar backscatter over forests at the timescale of weeks-months. The response of L-band radar backscatter to precipitation (at the shorter timescale of days) was inconsistent across time and across several different forested areas; only the response to the longer-timescale dry-down from the snow melt peak was consistent. Differences in backscatter depending on azimuthal angle have been noted, likely due to plant phototropism. Diurnal variations due to plant transpiration were expected to show higher backscatter at 6PM than 6AM over vegetated areas, which was true in many regions, but co-pol backscatter observed in late spring-summer over the northern boreal forests was higher at 6PM than 6AM; increased canopy extinction at 6AM was a possibility, but this was unproven and its true underlying physical cause undetermined.

A full comprehension of the temporal response of L-band radar backscatter over forested areas, on a global scale and at various timescales – from diurnal to seasonal – likely involves complications beyond our forward model. It is also likely that just L-band radar backscatter cross-sections HH, HV, VV may be insufficient to fully comprehend the geophysical changes in forested areas on some timescales.

Chapter 5

ALGORITHMS FOR SOIL MOISTURE RETRIEVAL FROM FORESTS USING L-BAND SCATTEROMETRY

5.1 Chapter overview

As mentioned in Section 4.2, the primary goal of the SMAP mission was to measure global soil moisture using spaceborne L-band radar and radiometer remote sensing. However, it was acknowledged that this would be a difficulty in areas covered by dense vegetation; as such the 0.04cm^3/cm^3 RMSE (root-mean-squared error) validated performance of the SMAP soil moisture product is only applicable to regions having vegetation water contents below ~5kgm^{-2} [4] . This excludes much of the forested areas of IGBP classes 1-5, which remain a challenge. In a study by Rahmoune et al. [117], using a soil moisture retrieval algorithm for the SMOS L-band radiometer data and evaluated over 14 nodes of the SCAN/SNOTEL soil moisture network in the United States, 0.07-0.12cm^3/cm^3 RMSE was obtained. In Chapter 3, we had attributed some of the temporal changes in L-band radar backscatter over forested areas to changes in the ground dielectric permittivity, i.e. soil moisture. In principle we should thus be able to use L-band radar scatterometry for soil moisture retrieval under forest canopies, despite the thick vegetation.

In this chapter, we shall explore soil moisture estimation under forests using the SMAP L-band radar. Section 5.2 presents our proposed soil moisture retrieval algorithms, which include purely empirical approaches (e.g. linear regression), and a semi-empirical approach that incorporates our forward model. Section 5.3 compares the results from our soil moisture retrieval algorithm against ground measurements at a number of stations from the SCAN, SNOTEL, USCRN, and COSMOS soil moisture networks in the United States.

5.2 Proposed soil moisture retrieval algorithms

At sufficiently high incidence angles θ_i, e.g. about 40° or more, and when the vegetation is dense but the radar backscatter is not yet completely saturated, the direct backscatter from the ground is very small and the sensitivity to soil moisture is primarily through the double-

reflection mechanism, in particular for the HH backscatter. Recall equation (2.1) of our forward model of the radar backscatter as the sum of various components, from Chapter 2:

$$\sigma^0 = \sigma^0_{cn} + \sigma^0_{gnd,direct} + \sigma^0_{cn-gnd,db} + \sigma^0_{trk-gnd,db}. \quad (5.1)$$

If we can neglect the direct ground backscatter in our forward model, the dielectric relative permittivity of the ground ε_g enters the equation only through the Fresnel reflectivity factor within the double-bounce terms. In the h-polarization, we can thus write the backscatter as

$$\sigma^0_{hh} = \sigma^0_{cn}\left(V_{b,tot}, \varepsilon_v, \theta_i\right) + A\left(V_{b,tot}, \varepsilon_v, \theta_i\right) \exp(-4k^2h^2\cos^2\theta_i) \left|R_h\left(\varepsilon_g, \theta_i\right)\right|^2 \quad (5.2)$$

where

$$\left|R_h\left(\varepsilon_g, \theta_i\right)\right|^2 = \left|\frac{\cos\theta_i - \sqrt{\varepsilon_g - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\varepsilon_g - \sin^2\theta_i}}\right|^2 \quad (5.3)$$

is the Fresnel reflectivity at h-polarization with ground dielectric relative permittivity ε_g and an incidence angle θ_i from the normal. The direct backscatter from the canopy $\sigma^0_{cn}\left(V_{b,tot}, \theta_i, \varepsilon_v\right)$ is independent of the soil moisture, and likewise for the factor $\left(V_{b,tot}, \varepsilon_v, \theta_i\right)$, which wraps up the dependences of the double-bounce terms on the vegetation; the ground roughness also enters only through the factor of $\exp(-4k^2h^2\cos^2\theta_i)$. If we further make the assumption that the vegetation parameters stay constant, then the temporal variations in co-polarized backscatter are linearly dependent on the Fresnel reflectance:

$$\text{HH} = \sigma^0_{cn} + A\exp(-4k^2h^2\cos^2\theta_i) \left|R_h\left(\varepsilon_g, \theta_i\right)\right|^2. \quad (5.4)$$

Next, we consider the relationship between the Fresnel reflectance and the soil moisture. Mironov et al. [43] provide formulae (see Appendix C) to relate the volumetric soil moisture and ground dielectric relative permittivity at L-band, using only the soil temperature and clay fraction as additional inputs. The dependence on temperature is quite weak over most of the range of interest, so for convenience and ease of inversion we shall simply fix the temperature to be T=15°C for computations. Using Mironov's model and the Fresnel equations, the h-

polarization Fresnel reflectance as a function of volumetric soil moisture is displayed in Figure 5.1.

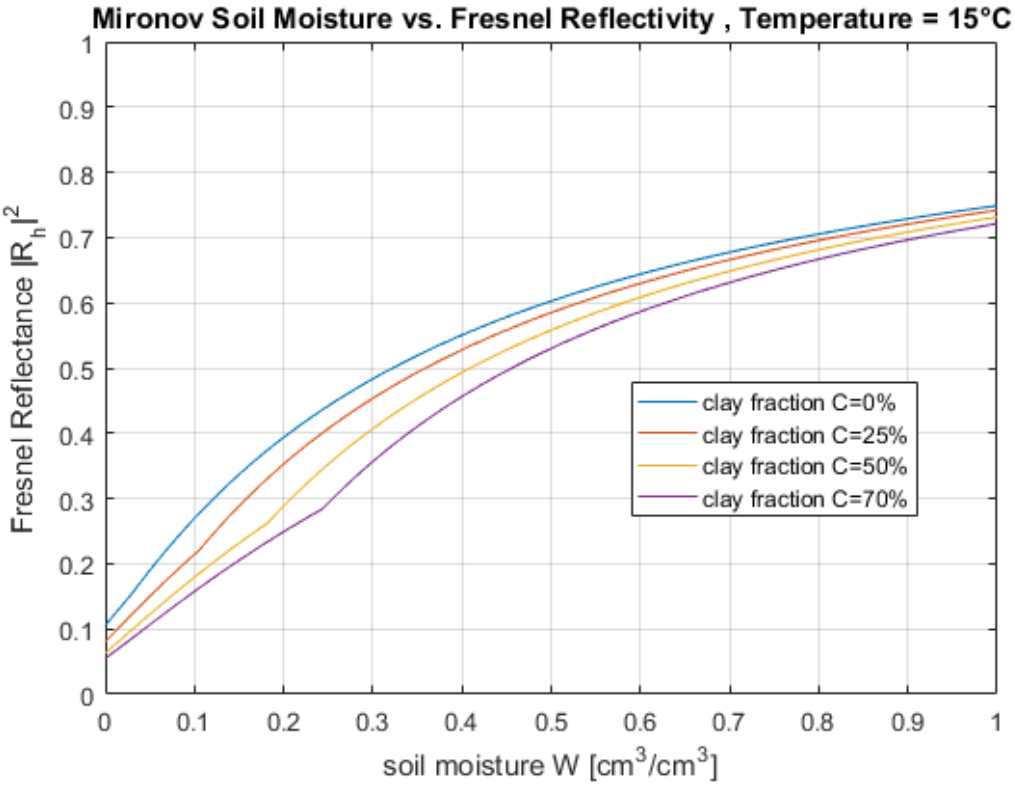

Figure 5.1. Graph of h-polarization Fresnel reflectance $|R_h|^2$ vs. volumetric soil moisture W at incidence angle $\theta_i = 40°$, for different soil clay fractions C and at temperature T=15°C, using the Mironov model [43] (details in Appendix C) to relate soil moisture and ground dielectric relative permittivity.

As can be seen from Figure 5.1, the Fresnel reflectance is a non-linear function of the soil moisture, however since the rate of change of slope is gradual, it is reasonable to approximate the slope of the graph as a constant value of 1 for volumetric soil moisture values W between 0-0.5cm^3/cm^3, for all clay fractions between 0-70%. The purpose of this approximation is to yield a simple linear relation between the radar backscatter and the volumetric soil moisture. We can write this linear relation in the form

$$\mathrm{HH} - \mathrm{HH}_{ref} = \frac{d\mathrm{HH}\left(V_{b,tot}, \theta_i, \varepsilon_v, h\right)}{d\mathrm{W}}\left(\mathrm{W} - \mathrm{W}_{ref}\right) \tag{5.5}$$

where HH_{ref} and W_{ref} are some corresponding reference values for the radar backscatter and volumetric soil moisture, and the slope in the linear relationship is a function of the various parameters but independent of soil moisture. Any dependence on clay fraction is subsumed into HH_{ref} and W_{ref}. We keep in mind the assumption that the vegetation and other input parameters are assumed constant, and soil moisture is the driver of changes in radar backscatter.

Soil moisture W can thus be retrieved from L-band radar backscatter HH by inverting the linear equation (5.5). Linear regression is a straightforward empirical method to minimize the root-mean-squared error between the retrieved soil moisture and the ground truth values. Ideally what is required would be a representative training set of corresponding L-band radar backscatter HH and in-situ ground-truth soil moisture values for each pixel; "representative" in the sense that the training set should be from the same distribution as the evaluation set on which the soil moisture retrieval is to be performed. Using linear regression, the parameters a and b in the equation

$$\mathrm{W} = a + b\mathrm{HH} \tag{5.6}$$

would be fitted and used subsequently for retrieval. The slope b and intercept a of the line would need to be found separately for each pixel, or group of pixels believed to share the same values. This linear regression approach is similar to that by Burgin et al. [118], Burgin and van Zyl [119] and Kim and van Zyl [120], except that here we use the radar backscatter in linear units instead of decibels. If the vegetation or parameters other than soil moisture change, then a and b would need to change too. Here, we assume they are static and unchanged across all the training and evaluation time periods.

Equation (5.6) requires a suitable amount of "training" data for the linear regression fit to find the slope b and intercept a. This may not always be available. In the absence of training values for the soil moisture W, but with a sufficient history of HH measurements, we can adopt the approach suggested by Wagner et al. [121] and Kim and van Zyl [120] and set

$$\mathrm{W} = \mathrm{W}_{dry} + \frac{\left(\mathrm{W}_{wet} - \mathrm{W}_{dry}\right)}{\left(\mathrm{HH}_{wet} - \mathrm{HH}_{dry}\right)}\left(\mathrm{HH} - \mathrm{HH}_{dry}\right) \tag{5.7}$$

where W_{wet} and W_{dry} are some specific a-priori known soil moisture values corresponding to anticipated maximum conditions for wetness and dryness, e.g. let W_{wet} be the field capacity or total water capacity, and W_{dry} be the wilting point. HH_{wet} and HH_{dry} are the corresponding values of the radar backscatter; with a sufficient history of HH measurements, Wagner et al. [121] suggested, using 6 years of ERS scatterometer data (C-band VV) over Ukraine, that it may be reasonable to assume that the highest observed value of HH corresponds to HH_{wet}, and the lowest observed value of HH corresponds to HH_{dry}. We shall refer to equation (5.7) as Wagner's approach. In our implementation, we shall let W_{wet} be the field capacity and W_{dry} be the wilting point, and HH_{dry} and HH_{wet} be the 10$^{\text{th}}$ and 90$^{\text{th}}$ percentiles of the observed HH values in the training set. This necessary hinges on the debatable assumption that the full range of wetness/dryness conditions are explored in the training set. Values for the field capacity and wilting point are obtained from a global 1km-resolution map of soil hydraulic properties produced by Zhang et al. [122].

If neither training values for the soil moisture W nor a long history of HH measurements are available, we offer a third alternative, which is to write

$$\mathrm{W} = \mathrm{W}_{avg} + B\left(\mathrm{HH} - \mathrm{HH}_{avg}\right) \tag{5.8}$$

where W_{avg} is the mean soil moisture, and HH_{avg} is the mean backscatter value. HH_{avg} may be estimated from a smaller set of training HH measurements that may not encompass the full range of wetness/dryness conditions; for robustness, we estimate it using the empirical median instead of the mean. In the absence of ground truth training data for W_{avg}, we propose to estimate it by setting the ground permittivity to be the value in Chapter 3 from fitting our forward model to the Aquarius data for that particular IGBP land cover class, and converting it to a soil moisture value using the Mironov soil dielectric model. Soil texture information, in particular the clay content required for the Mironov model, were obtained at 3km resolution from the SMAP ancillary data [123] [124]. For IGBP class 4, most of the

Aquarius data analyzed in Section 3.6 were from the Chaco dry forest in South America and thus we do not consider it representative of the deciduous broadleaf forests in USA, so instead we also use the values for IGBP class 5 for them.

In equation (5.8), the slope B can also be estimated from our forward model by

$$\mathrm{B} = \rho^2 \left[\frac{d\mathrm{HH}}{d\mathrm{W}}\right]^{-1} = \rho^2 \left[\frac{d\mathrm{HH}}{d(|R_h|^2)}\right]^{-1} \tag{5.9}$$

where we have used our approximation from Figure 5.1 that $\frac{d(|R_h|^2)}{d\mathrm{W}} = 1$. We then obtain from our forward model (using Aquarius-fitted parameters from Chapter 3) values of $\frac{d\mathrm{HH}}{d(|R_h|^2)}$. Again, for IGBP class 4 (deciduous broadleaf forests) in the United States, we use the values from IGBP class 5 instead (with the choice of the uniformly randomly oriented cylinders instead of horizontal cosine-squared orientation for the canopy layer). The values we use are tabulated in Table 5.1.

IGBP Class	$\frac{d\mathrm{HH}}{d(\lvert R_h\rvert^2)}$	$\left[\frac{d\mathrm{HH}}{d(\lvert R_h\rvert^2)}\right]^{-1}$
1 (Evergreen Needleleaf)	0.1502	6.658
4 (Deciduous Broadleaf)	0.1928	5.187
5 (Mixed forest)	0.1928	5.187

Table 5.1. Values of $\frac{d\mathrm{HH}}{d(|R_h|^2)}$ and its inverse, using our forward model and Aquarius-fitted parameters from Chapter 4, to be applied to equation (5.9).

In equation (5.9), ρ is the Pearson correlation coefficient between the soil moisture W and the radar backscatter HH. It appears that ρ still needs to be estimated empirically so we have not saved ourselves the need for a suitably-sized training dataset, but we might potentially workaround this by plugging in some prior or pooled estimate. Due to the input from the forward model, the approach of equation (5.8)-(5.9) is semi-empirical in contrast to linear regression and Wagner's approach which are both fully-empirical.

The three algorithms – linear regression, Wagner's approach, and semi-empirical approach – shall be evaluated using in-situ soil moisture measurements and Aquarius and SMAP radar backscatter data in the next section. Note that these algorithms do not guarantee that the retrieved soil moisture values would be physically reasonable. Thus we take the additional step of truncating the range of retrieved soil moisture values W to between the values for the saturated and residual water contents. These values are also obtained from the global soil hydraulic properties map by Zhang et al. [122]. Amongst the three algorithms, linear regression makes the most use of (and is most reliant upon) training information, while the semi-empirical approach uses the least training information.

5.3 Comparison and evaluation with in-situ soil moisture measurements

For the evaluation of our soil moisture retrieval algorithms and comparison with in-situ ground measurements, we use data from the SCAN [125], SNOTEL [86], USCRN [126], and COSMOS [127] [128] soil moisture networks in the United States, obtained through the International Soil Moisture Network (ISMN) database [87, 88]. There shall be two parts to the evaluation. The first part shall be on Aquarius data from 2011-2014 (HH-pol backscatter, with median taken over azimuth angle). 2 stations from SCAN, 2 stations from SNOTEL, and 3 stations from USCRN were identified and selected to fulfil the following criteria: lie in EASE2 grid pixels classified within IGBP forest classes 1, 4, and 5; not flagged for high terrain slope (see Section 3.2); have a fair amount of available and soil moisture data and corresponding Aquarius radar backscatter data values for unfrozen conditions. This last criterion was rather limiting and was the primary factor for only using 7 stations: limited Aquarius data have beam 2 overpasses coinciding with the station locations, and also having sufficient number data points not corrupted by RFI; in fact simply applying the overall quality flag provided with the data product would have been too stringent. Therefore, here we do not apply that overall quality flag and perform outlier detection manually by eye. Ascending pass Aquarius data (local 6PM) was also included for consideration. For the second part of the evaluation on SMAP data, 11 stations from SCAN, 10 stations from SNOTEL, 12 stations from USCRN, and 5 stations from COSMOS were used; their locations are marked on Figure 5.2, and they include the 7 stations used for the evaluation on Aquarius

data. These stations have been selected to lie in EASE2 grid pixels classified within IGBP forest classes 1,4, and 5, that are not flagged for high terrain slope (see Section 3.2), and also having a fair amount of available and valid soil moisture data and corresponding radar backscatter data values. The soil moisture measurements were made at a depth of 5cm, with the exception of COSMOS, which is sensitive to soil moisture from the surface down to some effective depth on the order of tens of centimeters depending on the soil moisture. For the SMAP comparison, soil moisture data from each station, from 13 April 2015 to 7 July 2015 at 6AM local time each day, was used.

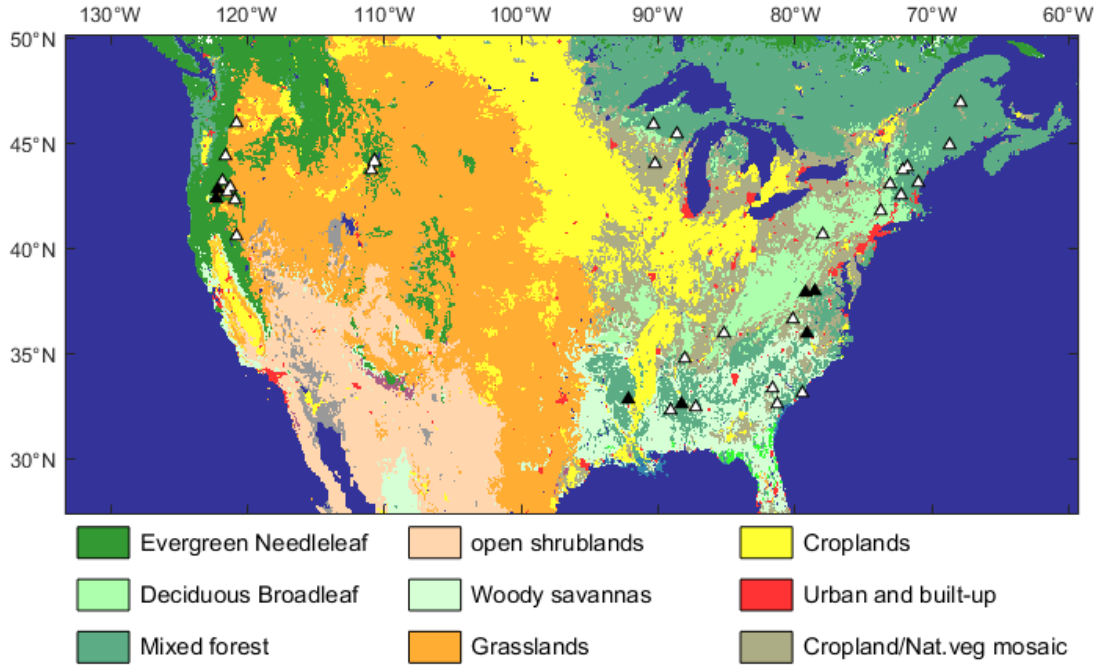

Figure 5.2. IGBP land cover classification map with triangles showing locations of selected ground stations from various soil moisture networks in the United States used for comparison with our soil moisture retrieval algorithm. Open triangles are for evaluations based on SMAP data, while filled black triangles are for evaluations based on both Aquarius and SMAP data.

The results for the Aquarius comparison against the 7 ground stations are shown in Table 5.2 and Figure 5.3-Figure 5.4, with each column of Table 5.2 corresponding to one ground station. For each station (and corresponding Aquarius HH backscatter), the data time series is divided into an earlier part for training and a later part for testing/evaluation, with

approximately equal number of points in each half. The training and testing time periods are shown in Table 5.2, and the data points themselves are plotted in the scatterplots of Figure 5.3 for each of the 7 stations. Training sets spanned a period of at least a year, though the available data may not always be uniformly/densely distributed over time. The Aquarius-derived soil moisture values on the test set are scattered against the in-situ values in Figure 5.4. The root-mean-squared differences (RMSD) between the in-situ soil moisture values and those derived from the Aquarius HH radar backscatter (using the three algorithms described previously in Section 5.2) are also tabulated in Table 5.2. Note that for the semi-empirical method, the correlation coefficient ρ in equation (5.9) was found from the training set.

On the training set, the root-mean-squared differences (RMSD) are approximately in the range 0.05-0.10cm^3/cm^3 for the linear regression algorithm, 0.05-0.13 cm^3/cm^3 for the Wagner approach, and 0.05-0.14 cm^3/cm^3 for the semi-empirical approach. On the test set, the RMSD values are approximately in the range 0.07-0.11 cm^3/cm^3 for the linear regression algorithm, 0.06-0.18 cm^3/cm^3 for the Wagner approach, and 0.06-0.14 cm^3/cm^3 for the semi-empirical approach. Overall amongst the three algorithms, linear regression gave the lowest RMSD values on the test set, while the Wagner and semi-empirical algorithms were comparable in RMSD. This is no surprise since if the assumptions of linear regression are valid, e.g. if the training and test set are indeed from the same distribution, linear regression theoretically minimizes the expected RMSD. From Table 5.2 and Figure 5.3, we see that though there is correlation in between the Aquarius L-band radar backscatter and the in-situ soil moisture, it is not very tight. When the correlation is poor, the linear regression and semi-empirical algorithms rely less on the radar measurement and guess the soil moisture based on the prior, yielding scatterplots in Figure 5.4 that look almost a horizontal line. The Wagner approach of equation (5.7) implicitly assumes a good correlation and might yield wild guesses and a higher RMSD than the other two algorithms in such cases.

	Station ID#	344	1000	1012	1347	1346	2088	2114
	IGBP class	1	1	5	5	4	4	5
Training	period	Aug'11-Aug'13	Aug'11-Jun'13	Apr'12-May'13	Apr'13-Apr'14	Apr'12-Jun'13	Apr'13-Apr'14	Apr'12-Apr'13
	data #pts	41	21	17	14	15	7	10
	Corr. Coeff. ρ	0.65	0.86	0.36	0.51	0.55	0.19	0.27
	Lin. reg. RMSD	0.1024	0.0622	0.0688	0.0483	0.0751	0.0702	0.0490
	Wagner RMSD	0.1194	0.1061	0.1003	0.0543	0.1016	0.1286	0.0747
	Semi-empiric. RMSD	0.1034	0.1193	0.0755	0.0561	0.0770	0.1358	0.0522
Testing	Period	Aug'13-Dec'14	Jun'13-Nov'14	May'13-Jul'14	Apr'14-Jul'14	Jun'13-Jun'14	Apr'14-Jun'14	Apr'13-Jun'14
	data #pts	42	22	17	15	16	8	11
	Lin. reg. RMSD	0.0997	0.0787	0.0925	0.0736	0.0882	0.1058	0.1110
	Wagner RMSD	0.1126	0.1416	0.0935	0.0671	0.1007	0.0638	0.1767
	Semi-empiric. RMSD	0.0986	0.1363	0.0869	0.0644	0.1013	0.0942	0.1131
Total	Period	Aug'11-Dec'14	Aug'11-Nov'14	Apr'12-Jul'14	Apr'13-Jul'14	Apr'12-Jun'14	Apr'13-Jun'14	Apr'12-Jun'14
	data #pts	83	43	34	29	31	15	21
	Corr. Coeff. ρ	0.66	0.75	0.40	0.65	0.60	0.59	0.10

Table 5.2. Summary of results comparing soil moisture retrieved from Aquarius HH backscatter vs. in-situ measurements at 7 selected ground stations. See main text for discussion.

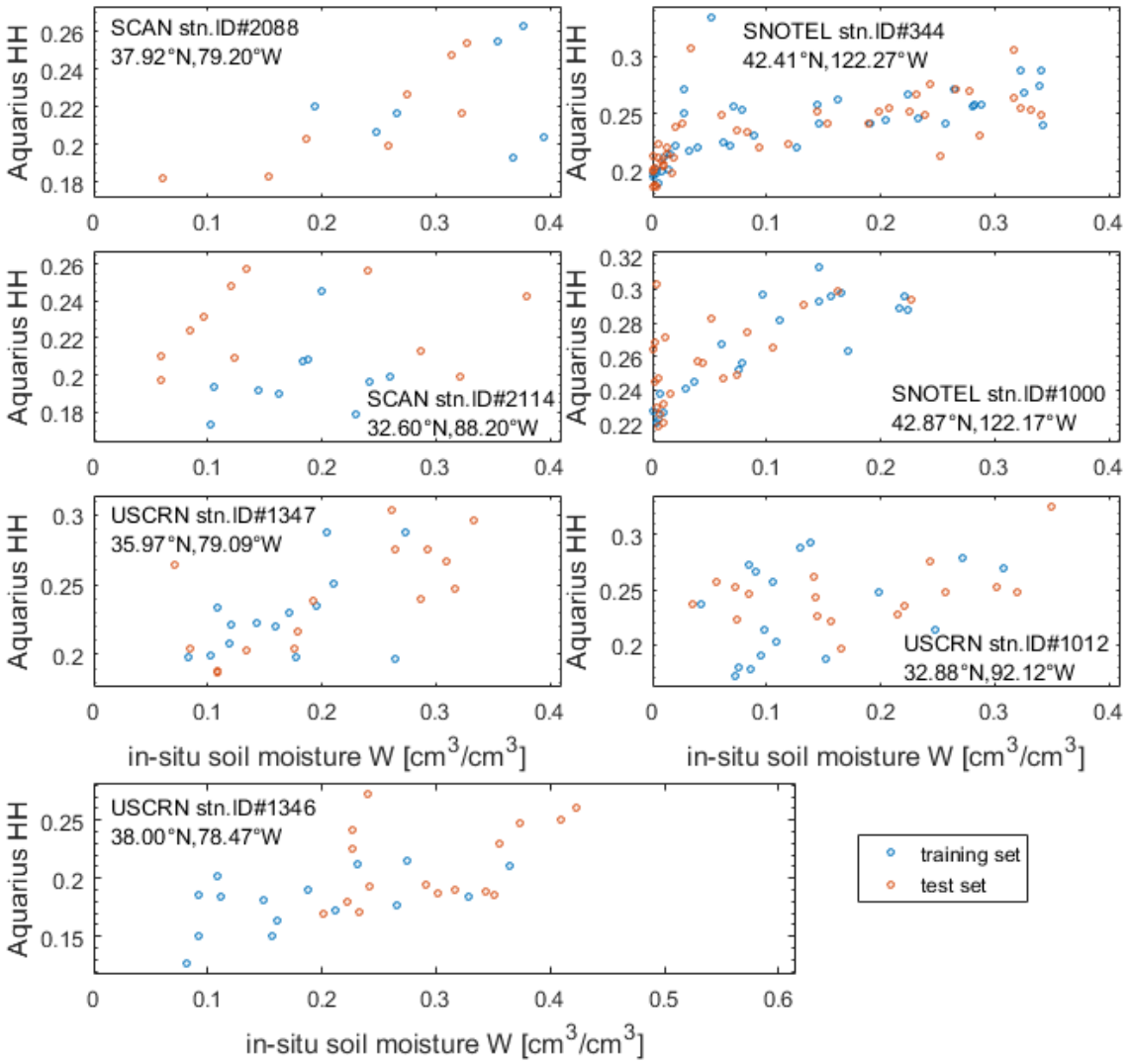

Figure 5.3. Scatterplots of 2011-2014 Aquarius beam 2 HH-pol radar backscatter (vertical axes, in linear units) data against corresponding in-situ volumetric soil moisture values (horizontal axes) at 7 selected ground stations. The data are divided into training (blue pts) and testing (red pts) sets; the training and testing time-periods are shown in Table 5.2. See main text for discussion.

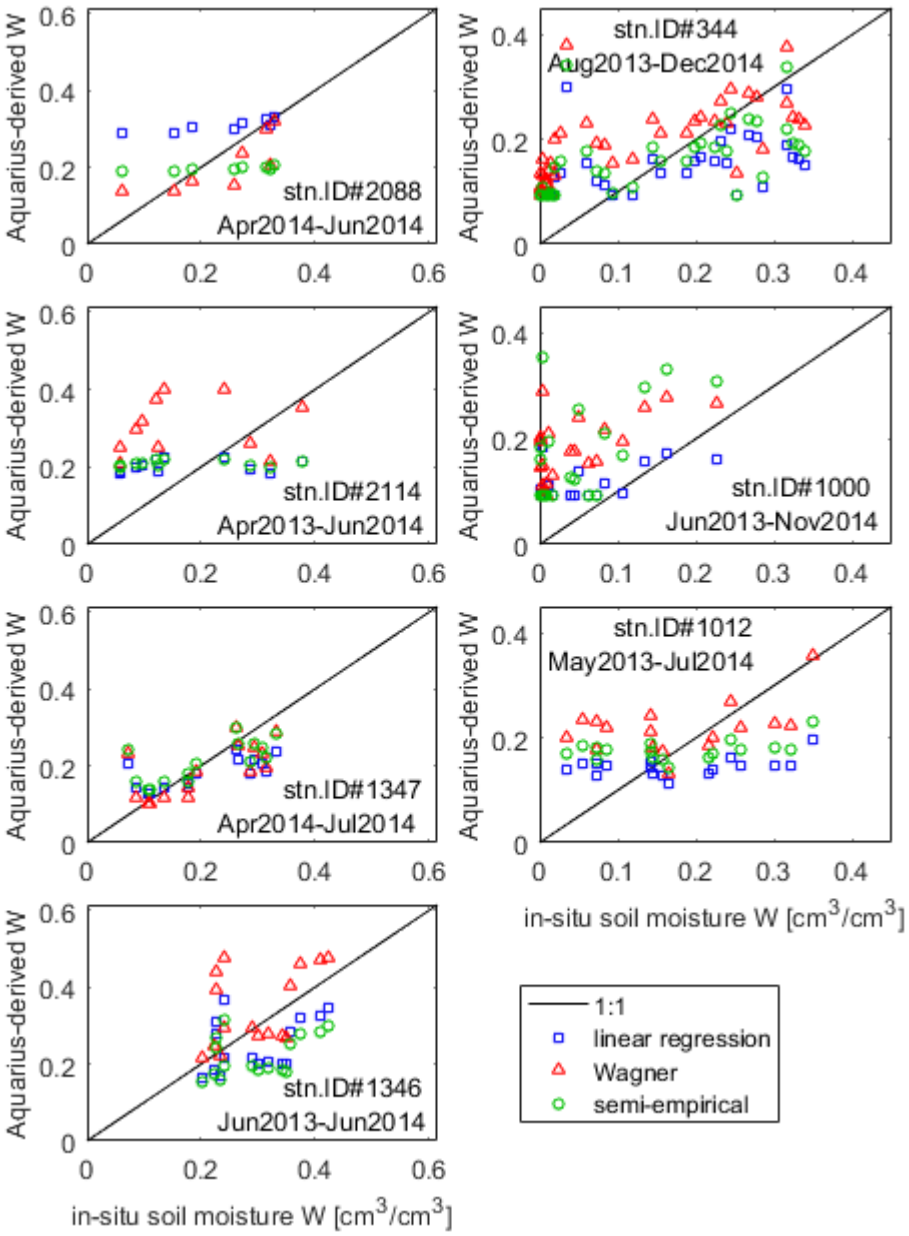

Figure 5.4. Scatterplots of retrieved volumetric soil moisture values (vertical axes) against corresponding in-situ values (horizontal axes) on the test set, for each of the 7 selected ground stations, and each of the three algorithms described in Section 5.2. Blue squares are the results from linear regression (equation 5.6), red triangles from Wagner's approach, and green circles from the semi-empirical method. See main text for discussion.

Next, we evaluate the algorithms on the SMAP data. We now pool the training and test sets from Aquarius together, and use that as a training set. The correlation coefficients from this larger training set are shown in the last row of Table 5.2. The algorithms are then applied to the SMAP HH backscatter data, and these retrieved soil moisture results (paired with corresponding in-situ values at the seven ground stations) form the test set. The SMAP test set comprises data from April 13, 2015 to July 7, 2015. Only 6AM (local time) data was used, for consistency comparing with the SMAP soil moisture products. The results of the evaluation of the three algorithms described in Section 5.2 (linear regression, Wagner, semi-empirical) and trained on Aquarius data, against the in-situ volumetric soil moisture values at the seven stations, are tabulated in Table 5.3, with each column corresponding to one station. The correlation coefficients in the third row are merely a re-display from the last row of Table 5.2. The correlation coefficients in the fourth row are the correlation coefficients between the SMAP HH backscatter and the in-situ soil moisture values. The RMSD between the in-situ values and the outputs from each of the algorithms are in the subsequent rows. Also in Table 5.3 are RMSD values between the in-situ soil moisture values and the outputs from the existing soil moisture products provided by SMAP. "SMA" is the soil moisture product using the SMAP radar data and the baseline algorithm by Kim et al. [129]. "SMP" is the soil moisture product using the SMAP radiometer data and the algorithm by O'Neill et al. [130]. We used Level-3, version 3 products, obtained from the NASA National Snow and Ice Data Center Distributed Active Archive Center (NSIDC) [130, 131]. Note that these algorithms and products are not validated nor expected to be accurate for areas with vegetation water content >5kg/m2 which includes the forested areas being studied here. ρ_W is the correlation coefficient between the in-situ values and the values from the soil moisture product. Complementing Table 5.3 is Figure 5.5, which (analogous to Figure 5.4) displays scatterplots of the retrieved soil moisture values vs. corresponding in-situ values on the SMAP test set, for each of the 7 selected ground stations, and each of the three algorithms described in Section 5.2, as well as the SMAP products.

Station ID#		344	1000	1012	1347	1346	2088	2114
IGBP class		1	1	5	5	4	4	5
Aquarius correlation coeff. ρ		0.66	0.75	0.40	0.65	0.60	0.59	0.10
SMAP correlation coeff. ρ		0.77	0.89	0.42	0.72	0.68	0.79	0.15
Data #pts		37	41	36	40	37	40	16
Linear regression RMSD		0.0704	0.0413	0.1450	0.0735	0.0882	0.0713	0.1820
Wagner RMSD		0.0730	0.1149	0.0777	0.0737	0.1051	0.0995	0.1811
Semi-empirical RMSD		0.0677	0.0815	0.1399	0.0664	0.1277	0.1126	0.1553
SMA	SMA L3 RMSD	0.2624	0.2476	0.0950	-	0.0899	0.0715	0.1277
	data #pts	39	39	29	0	34	32	12
	Corr. ρ_W	0.15	0.47	-0.04	-	-0.02	0.37	0.30
SMP	SMP L3 RMSD	0.3363	0.1579	0.1519	0.1101	0.0618	0.0445	0.0745
	data #pts	43	42	37	41	38	41	17
	Corr. ρ_W	0.79	0.85	-	0.81	0.70	0.80	-0.05

Table 5.3. Summary of results comparing soil moisture retrieved from SMAP HH backscatter vs. in-situ measurements at 7 selected ground stations. The SMAP L3 version 3 active soil moisture product (radar-only) and passive soil moisture product (radiometer-only) are also included for comparison. See main text for discussion.

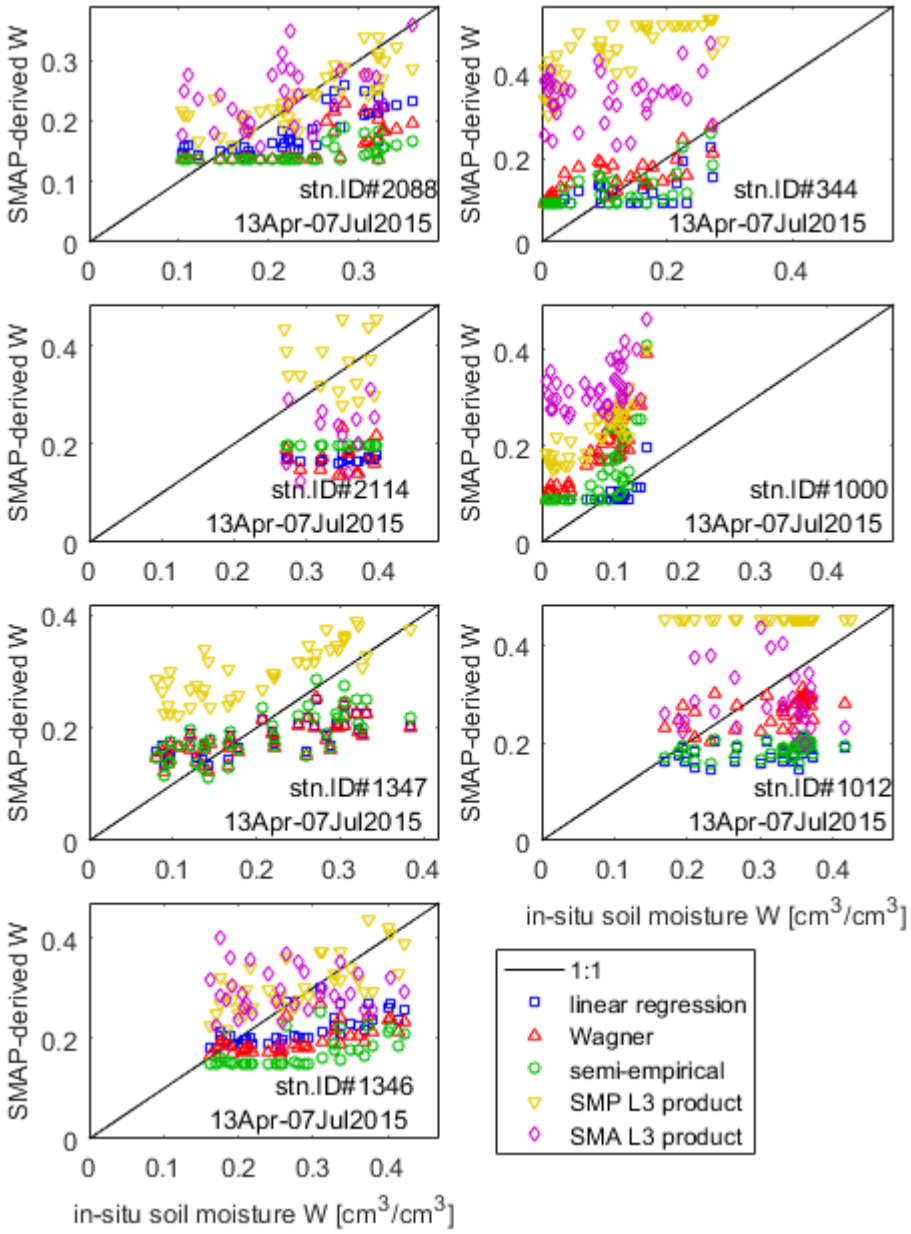

Figure 5.5. Scatterplots of retrieved volumetric soil moisture values derived from SMAP HH backscatter (vertical axes) against corresponding in-situ values (horizontal axes), for each of the 7 selected ground stations, and for each of the three algorithms described in Section 5.2, as well as the SMAP products. Blue squares are the results from linear regression, red upright triangles from Wagner's approach, green circles from the semi-empirical method, yellow inverted triangles from SMAP's passive soil moisture product (radiometer-only), and purple diamonds from SMAP's active soil moisture product (radar only). See main text for discussion.

Overall amongst the three algorithms, again linear regression gave the lowest RMSD values on the SMAP test set, while the Wagner and semi-empirical algorithms were comparable in RMSD. RMSD was unsurprisingly high at stations #1012 and #2114 where correlation is poor on both training and test sets. SMA and SMP had exceptionally high RMSD values on station #344, but had lower RMSDs on some of the other stations. Inspection of Figure 5.5 shows that there was correlation between SMP and the in-situ values, however SMP sometimes suffered from a large offset bias from the in-situ values. There were several instances of horizontal/near-horizontal-line retrievals in Figure 5.5 for some of the algorithms; these were either due to floor or ceiling values, or low values of correlation coefficient ρ from Aquarius causing near-zero slopes in the retrieval equation.

With just seven stations, it is difficult and unwise to read too much into the evaluations. We next perform evaluation over all 38 stations shown in Figure 5.2, but this time using a subset of the SMAP data for training our three algorithms (linear regression, Wagner, semi-empirical), since we lack training data from Aquarius for most of these stations. At each station, 5 pairs of data points (pairs of in-situ soil moisture values and corresponding SMAP HH backscatter values) were chosen at random for training, out of all the data from April 13, 2015 to July 7, 2015. The remainder (about 20-40 data points) of the data points forms our test set. For linear regression, these 5 points were used to find the line of best fit and estimation of slope and intercept. For the semi-empirical method, here we take the correlation coefficient to be ρ=0.60 for all 38 stations, which is the median value of the Aquarius-computed correlation coefficients at the seven stations (last row of Table 5.2, or third row of Table 5.3). For Wagner's approach, the minimum and maximum observed HH values, out of those 5, were used as the 10^{th} and 90^{th} percentiles of the observed HH values in the algorithm. Here, Wagner's approach is expected to fail disastrously since the maximum and minimum of only 5 data points from April-July are unlikely to correspond to the field capacity W_{wet} and wilting point W_{dry} . The results of the evaluation are displayed in the boxplots of Figure 5.6 and Figure 5.7.

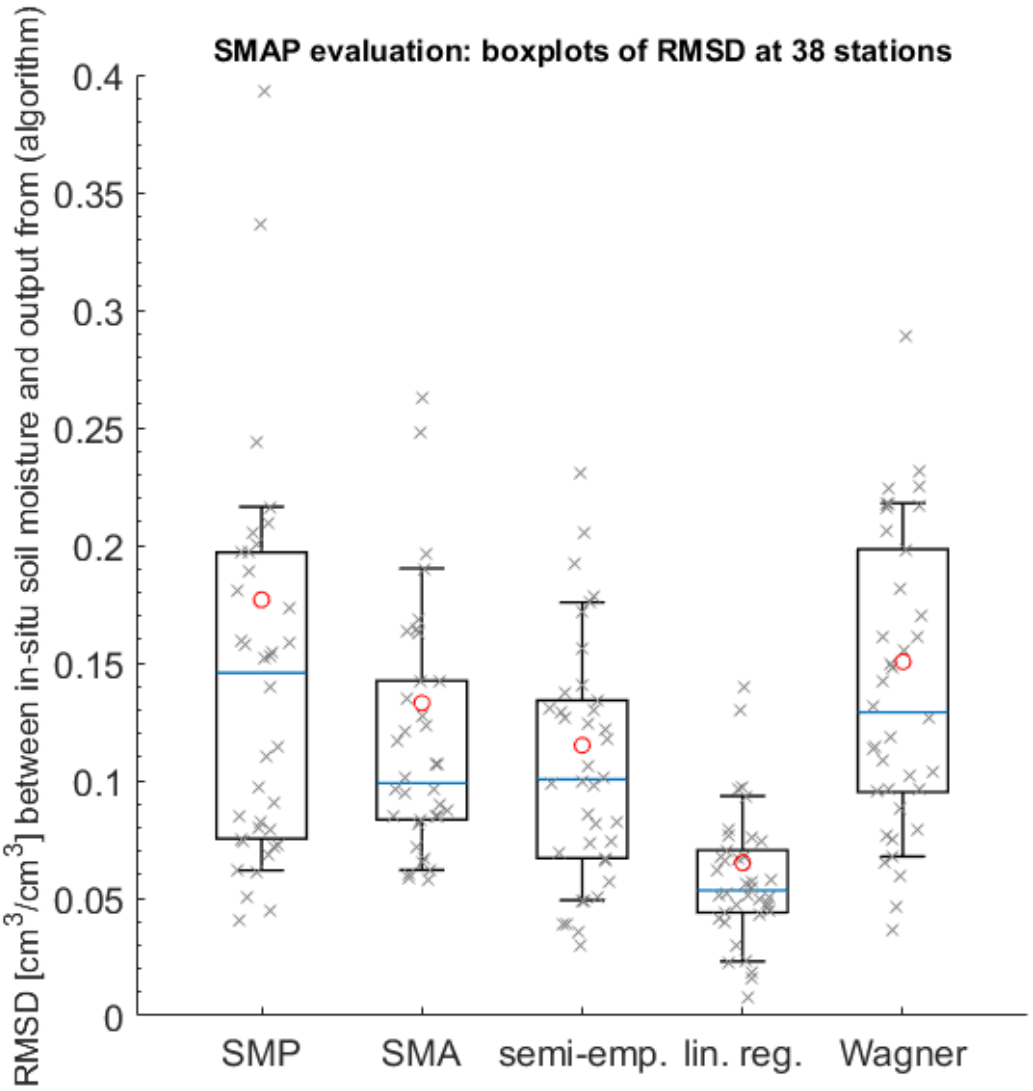

Figure 5.6. Boxplots of the root-mean-squared-difference (RMSD) between the in-situ soil moisture and the retrieved values output from the various algorithms, showing the distribution over all 38 stations. The RMSD from each of the 38 stations is scattered (marked with crosses), the boxes show the inter-quartile range, the horizontal blue lines the median, the whiskers the 10th and 90th percentiles. The red circles show the overall RMSD computed by pooling all data points from all stations.

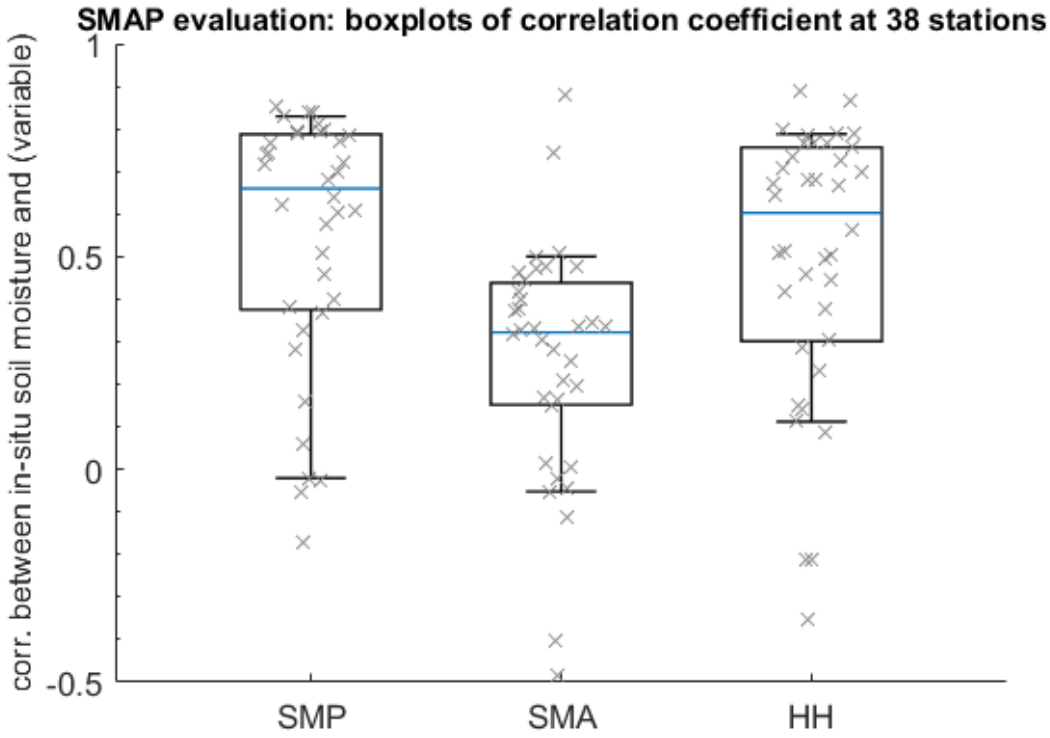

Figure 5.7. Boxplots of the Pearson correlation coefficient between the in-situ soil moisture and the retrieved soil moisture values in the SMP and SMA products, showing the distribution over the 38 ground stations. The correlation coefficient between the in-situ values and the SMAP HH backscatter is also plotted. The respective correlation coefficients from each of the 38 stations is scattered (marked with crosses), the boxes show the inter-quartile range, the horizontal blue lines the median, the whiskers the 10^{th} and 90^{th} percentiles.

Figure 5.6 contains boxplots of the root-mean-squared-difference (RMSD) between the in-situ volumetric soil moisture measured at the ground stations and the retrieved values output from the various algorithms, showing the distribution over all 38 stations. Figure 5.7 contains boxplots of the Pearson correlation coefficient between the in-situ soil moisture and the retrieved soil moisture values in the SMP and SMA products, as well as the correlation coefficient between the in-situ values and the SMAP HH backscatter. From Figure 5.6, we see that, in terms of this RMSD metric, linear regression gives the lowest RMSD, followed by the semi-empirical algorithm, followed by SMA, Wagner's approach, and SMP has the highest RMSD. The failure of Wagner's algorithm was expected due to the small training size. For the small training size, linear regression gave a surprisingly low RMSD, while the semi-empirical algorithm was by design robust against a small training size. It is noteworthy

that our rudimentary semi-empirical algorithm, which does not use any in-situ soil moisture values for training, slightly outperformed the baseline SMA product in RMSD. This is not to say that the semi-empirical algorithm in this form has a good RMSD performance; rather, all the algorithms, with the exception of linear regression, have median RMSDs ≥10%, which should be considered a poor performance.

On the other hand, caution must be taken in interpreting these SMAP and Aquarius evaluations to mean that linear regression is overwhelmingly the most superior algorithm. Note that the in-situ measurements are at a single point on the ground, which may not necessarily be representative of the entire 30-40km SMAP radar footprint (approximately the size of a 36km EASE2 grid pixel); there may be significant variation in the soil moisture at different points within the same footprint. COSMOS offered a much larger ground area (at the scale of hundreds of metres [127]) than a point measurement, but still significantly smaller than the SMAP footprint. Famiglietti et al. [133] analyzed the spatial variability of soil moisture using data from various field experiments in Oklahoma and Iowa to be about $0.07 cm^3/cm^3$ standard deviation at the 50km scale, but this standard deviation itself had scatter and variability; the coefficient of variation (ratio of standard deviation to mean) was a decreasing function of soil moisture. An analysis of the in-situ soil moisture data from SMAP core-validation sites [134] verifies the findings by Famiglietti et al. [133]. One of the purposes of the SMAP core-validation sites is to have accurate and validated soil moisture ground-truth measurements made over multiple sites within a 30-40km footprint so as to mitigate this variability. However, the SMAP product was not expected to be valid in areas of dense vegetation, thus these sites are mostly over cropland or grassland areas with none in the forest land cover classes. As such, we have to rely mostly on the available point-measurement data. There may also be significant land cover heterogeneity within each 30-40km footprint. Since our linear regression algorithm made use of in-situ soil moisture data for training and evaluation was performed on the same station, the linear regression results may thus be "over-trained" on just the immediate vicinity of the stations themselves instead of the soil moisture over the entire footprint. It might be better to think of the RMSD of the linear regression results as a "best case" scenario, or performance upper limit, in the absence

of spatial heterogeneity within a footprint, or if that could somehow be taken care of. Indeed, as a rough estimate, if we suppose that spatial heterogeneity contributes $0.07 cm^3/cm^3$ of variation, adding this to the linear regression RMSD values yields a median of about $0.09 cm^3/cm^3$, which is now a much smaller advantage over the semi-empirical algorithm, an advantage that we might presume related to the utilization of some soil moisture training information vs. no soil moisture training information for the semi-empirical algorithm.

Further remarks can be made concerning Figure 5.7. Empirically, we observe that there is some correlation, though not a tight one, between in-situ soil moisture and the L-band HH backscatter. In fact the passive soil moisture product SMP exhibited a slightly better correlation with the in-situ values, implying that the radiometer data contains useful information for soil moisture retrieval. The poor RMSD of SMP in Figure 5.6 suggests that instead it is merely the algorithm for the passive soil moisture product (which was not designed for areas with vegetation water content $>5 kg/m^2$) that is not suitable over forested areas and led to significant errors. Likewise, the much poorer correlation of SMA with in-situ soil moisture despite correlation between HH and the in-situ values suggests that the algorithm for the baseline SMA active soil moisture product had not managed to properly exploit the radar data in areas of high vegetation, where it was not designed for. Evidently different algorithms should be used for soil moisture retrieval under forests; perhaps some of our approaches introduced in this chapter might have better exploited the radar data.

Separately, based on the correlation coefficients we could make the interpretation that the square of the correlation coefficient is the fraction of the variance in HH backscatter explained by soil moisture variation. A significant fraction (very roughly half) of variance in HH backscatter is thus not due to soil moisture. Compare this to the model-based analysis of Chapter 3 where it seemed that, though secondary factors were important and far from negligible, the primary source of temporal variation was thought to be likely due to soil moisture. Indeed, based on the model-based approaches of Chapter 2 and Chapter 3, one might raise the very compelling suggestion that simply using the HH radar backscatter in our proposed algorithms does not make full use of the polarization information; the HH radar backscatter is “contaminated” by changes due to other factors e.g. changes in vegetation.

Based on the model, one might attempt to "cancel out" the vegetation changes by using corrections of the form HH $-\gamma$HV, or similar forms in logarithmic (dB) scale. Figure 5.8 does a quick exploration of this idea on the data, as a function of γ. However as can be seen from the figure, such an approach might be marginally beneficial at best. Using logarithmic-transformed variables yields a similar conclusion (not displayed). It is unclear if this is due to inadequacies in our modelling, the limited ability of one station to represent the soil moisture of the entire radar footprint, or other unaccounted-for factors driving HH backscatter changes.

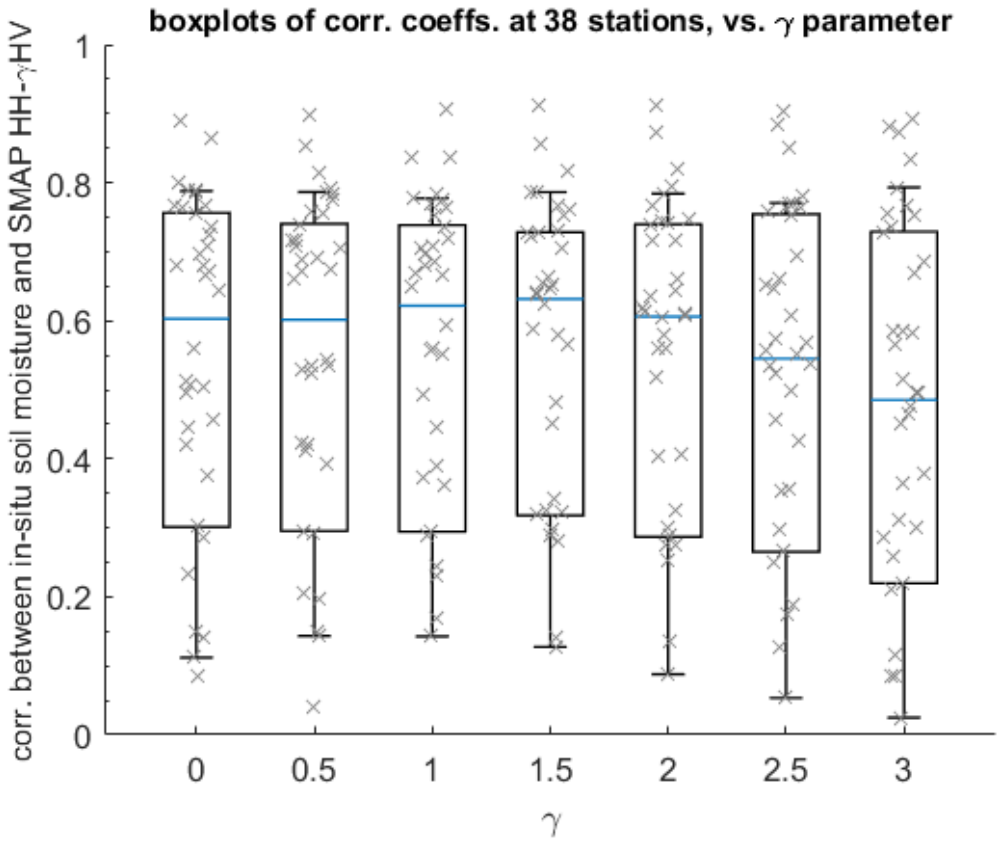

Figure 5.8. Boxplots of the Pearson correlation coefficient between the in-situ soil moisture and the quantity HH $-\gamma$HV computed from the SMAP radar data over the 38 ground stations, for various values of γ.

5.4 Conclusion

We proposed and evaluated several soil-moisture retrieval algorithms for forested areas using Aquarius and SMAP L-band radar backscatter and in-situ volumetric soil moisture

measurements at 38 ground stations in the United States. In terms of the RMSD between the ground measurements and the retrieval algorithm outputs, our linear regression algorithm (overall RMSD 6.5%) and semi-empirical algorithm (overall RMSD 11.5%) performed better than baseline algorithms from SMAP that used either passive radiometer (overall RMSD 18%) or active radar (overall RMSD 13%) measurements only, though these SMAP soil moisture products were not intended to be valid in forested areas. Our semi-empirical algorithm, which is based on our forward model from earlier chapters, did not make use of any in-situ soil moisture values for training. The linear regression algorithm relied on in-situ soil moisture values for training and may represent an estimate of the expected limit of soil moisture retrieval RMSD performance over forested areas using just L-band multi-polarization scatterometry with no additional phase information. Correlations between the radar backscatter and the in-situ soil moisture were positive but not tight, showing a significant fraction of the temporal variance in radar backscatter not directly due to soil moisture changes; despite suggestions by our forward model in earlier chapters, it remained difficult to properly separate these different factors using L-band multi-polarization scatterometry alone.

BIBLIOGRAPHY

[1] G. E. Kindermann, I. McCallum, S. Fritz and M. Obersteiner, "A global forest growing stock, biomass and carbon map based on FAO statistics," *Silva Fennica*, vol. 42, no. 3, p. 244, 2008. DOI: 10.14214/sf.244.

[2] National Research Council, Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond. Washington, DC: The National Academies Press, 2007. DOI: 10.17226/11820.

[3] Y. Zhang et al., "Multi-decadal trends in global terrestrial evapotranspiration and its components," *Scientific Reports*, vol. 6, p. 19124, 2016. DOI: 10.1038/srep19124.

[4] Entekhabi et al., SMAP Handbook, National Aeronautics and Space Administration, 2014. url: https://smap.jpl.nasa.gov/system/internal_resources /details/original/ 178_SMAP_Handbook_FINAL_1_JULY_2014_Web.pdf

[5] R. Sessa and H. Dolman (Eds.), Terrestrial essential climate variables for climate change assessment, mitigation and adaptation. Rome: FAO GTOS-52. 2008.

[6] T. Le Toan, S. Quegan, M. Davidson, H. Balzter, P. Paillou, K. Papathanassiou, S. Plummer, F. Rocca, S. Saatchi, H. Shugart and L. Ulander, "The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle," *Remote Sensing of Environment*, vol. 115, no. 11, pp. 2850-2860, 2011. DOI: 10.1016/j.rse.2011.03.020.

[7] S. S. Saatchi et al, "Benchmark map of forest carbon stocks in tropical regions across three continents," *Proceedings of the National Academy of Sciences*, vol. 108, no. 24, pp. 9899-9904, 2011. DOI: 10.1073/pnas.1019576108.

[8] M. Moghaddam, S. Saatchi and R. H. Cuena, "Estimating subcanopy soil moisture with radar," *Journal of Geophysical Research*, vol. 105, no. D11, pp. 14899-14911, 2000. DOI: 10.1029/2000JD900058.

[9] A. Tabatabaeenejad, M. Burgin and M. Moghaddam, "Potential of L-band Radar for Retrieval of Canopy and Subcanopy Parameters of Boreal Forests," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 6, pp. 2150-2160, 2012. DOI: 10.1109/TGRS.2011.2173349.

[10] S. R. Cloude, Polarisation: Applications in Remote Sensing, Oxford: Oxford University Press, 2009.

[11] J. J. van Zyl and Y. Kim, Synthetic Aperture Radar Polarimetry, New Jersey: Wiley, 2011.

[12] J. J. van Zyl, "Application of Cloude's target decomposition theorem to polarimetric imaging radar data," in *Proc. SPIE 1748, Radar Polarimetry*, 1993.

[13] A. Freeman and S. L. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 36, no. 3, pp. 963-973, 1998. DOI: 10.1109/36.673687.

[14] S. L. Durden, J. J. van Zyl and H. A. Zebker, "Modeling and Observation of the Radar Polarization Signature of Forested Areas," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 27, no. 3, pp. 290-301, 1989. DOI: 10.1109/36.17670.

[15] N. S. Chauhan and R. S. Lang, "Microwave Modelling of Orchard Canopy," in *International Geoscience and Remote Sensing Symposium, 'Remote Sensing: Moving Toward the 21st Century,'* Edinburgh, UK, 1988.

[16] N. S. Chauhan, R. H. Lang and K. J. Ranson, "Radar modeling of a boreal forest," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 29, no. 4, pp. 627-638, 1991. DOI: 10.1109/36.135825.

[17] F. T. Ulaby, K. Sarabandi, K. McDonald, M. Whitt and M. C. Dobson, "Michigan microwave canopy scattering model," *International Journal of Remote Sensing,* vol. 11, no. 7, pp. 1223-1253, 1990. DOI: 10.1080/01431169008955090.

[18] Y. Wang, J. L. Day, F. W. Davis and j. M. Melack, "Modeling L-band radar backscatter of Alaskan boreal forest," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 31, no. 6, pp. 1146-1154, 1993. DOI: 10.1109/36.317448.

[19] M. A. Karam, A. K. Fung, R. H. Lang and N. S. Chauhan, "A microwave scattering model for layered vegetation," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 30, no. 4, pp. 767-784, 1992. DOI: 10.1109/36.158872.

[20] G. Sun and K. J. Ranson, "Three-dimensional radar backscatter model of forest canopies," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 33, no. 2, pp. 372-382, 1995. DOI: 10.1109/36.377937.

[21] M. Burgin, D. Clewley, R. M. Lucas and M. Moghaddam, "A Generalized Radar Backscattering Model Based on Wave Theory for Multilayer Multispecies Vegetation," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 49, no. 12, pp. 4832-4845, 2011. DOI: 10.1109/TGRS.2011.2172949.

[22] S.-B. Kim, M. Moghaddam, L. Tsang, M. Burgin, X. Xu and E. G. Njoku, "Models of L-band Radar Backscattering Coefficients Over Global Terrain for Soil Moisture Retrieval," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 52, no. 2, pp. 1381-1396, 2014. DOI: 10.1109/TGRS.2013.2250980.

[23] F. T. Ulaby and P. P. Batlivala, "Diurnal Variations of Radar Backscatter from a Vegetation Canopy," *IEEE Transactions on Antennas and Propagation,* Vols. AP-24, no. 1, January 1976. DOI: 10.1109/TAP.1976.1141298.

[24] B. Brisco, R. J. Brown, J. A. Koehler, G. J. Sofko and M. J. McKibben, "The diurnal pattern of microwave backscattering by wheat," *Remote Sensing of Environment,* vol. 34, no. 1, pp. 37-47, 1990. DOI: 10.1016/0034-4257(90)90082-W.

[25] J. Way, J. Paris, M. C. Dobson, K. McDonald, F. T. Ulaby, J. A. Weber, S. L. Ustin, V. C. Vanderbilt and E. S. Kasischke, "Diurnal Changes in Trees as

Observed by Optical and Microwave Sensors: The EOS Synergism Study," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 29, no. 6, pp. 807-821, 1991. DOI: 10.1109/TGRS.1991.1019466.

[26] J. A. Weber and S. L. Ustin, "Diurnal water relations of walnut trees: Implications for remote sensing," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 29, no. 6, pp. 864-874, 1991. DOI: 10.1109/TGRS.1991.1019470.

[27] S. C. Steele-Dunne, J. Friesen and N. van de Giesen, "Using Diurnal Variation in Backscatter to Detect Vegetation Water Stress," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 50, no. 7, pp. 2618-2629, 2012. DOI: 10.1109/TGRS.2012.2194156.

[28] F. T. Ulaby and D. G. Long, Microwave Radar and Radiometric Remote Sensing, Ann Arbor: The University of Michigan Press, 2014.

[29] F. T. Ulaby, R. K. Moore and A. K. Fung, Radar remote sensing and surface scattering and emission theory. Vol. II of Microwave remote sensing: Active and passive, Norwood, MA: Artech House, 1986.

[30] C. F. Bohren and D. R. Huffman, Absorption and Scattering of Light by Small Particles, New York: Wiley, 1998.

[31] J. J. van Zyl and F. T. Ulaby, Chapter 2 in Radar Polarimetry for Geoscience Applications, F. T. Ulaby and C. Elachi, Eds., Norwood, MA: Artech House, 1990.

[32] H. C. van de Hulst, Light scattering by small particles, New York: Dover, 1981.

[33] A. Ishimaru, Wave Propagation and Scattering in Random Media, Academic Press, 1978.

[34] R. G. Newton, "Optical theorem and beyond," *American Journal of Physics,* vol. 44, no. 7, pp. 639-642, 1976. DOI: 10.1119/1.10324.

[35] T. R. Loveland and A. S. Belward, "The International Geosphere Biosphere Programme Data and Information System global land cover data set (DISCover)," *Acta Astronautica,* vol. 41, no. 4-10, pp. 681-689, 1997. DOI: 10.1016/S0094-5765(98)00050-2.

[36] R. H. Lang and J. S. Sighu, "Electromagnetic Backscattering from a Layer of Vegetation: A Discrete Approach," *IEEE Transactions on Geoscience and Remote Sensing,* Vols. GE-21, no. 1, pp. 62-71, 1983. DOI: 10.1109/TGRS.1983.350531.

[37] L. Tsang, J. A. Kong and R. T. Shin, Theory of Microwave Remote Sensing, New York: Wiley, 1985.

[38] H. Huang, L. Tsang, E. G. Njoku, A. Colliander, T. Liao and K. Ding, "Propagation and Scattering by a Layer of Randomly Distributed Dielectric Cylinders Using Monte Carlo Simulations of 3D Maxwell Equations With Applications in Microwave Interactions With Vegetation," *IEEE Access,* vol. 5, pp. 11985-12003, 2017. DOI: 10.1109/ACCESS.2017.2714620.

[39] M. Arii, J. J. van Zyl and Y. Kim, "A General Characterization for Polarimetric Scattering From Vegetation Canopies," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 48, no. 9, pp. 3349-3357, 2010. DOI: 10.1109/TGRS.2010.2046331.

[40] F. T. Ulaby, R. K. Moore and A. K. Fung, From theory to applications. Vol. III of Microwave remote sensing: Active and passive, Norwood, MA: Artech House, 1986.

[41] A. K. Fung and K. S. Chen, Microwave Scattering and Emission Models for Users, Boston: Artech House, 2010.

[42] L. Tsang, K. Ding, S. Huang and X. Xu, "Electromagnetic Computation in Scattering of Electromagnetic Waves by Random Rough Surface and Dense Media in Microwave Remote Sensing of Land Surfaces," *Proceedings of the IEEE,* vol. 101, no. 2, pp. 255-279, 2013. DOI: 10.1109/JPROC.2012.2214011.

[43] V. Mironov, Y. Kerr, J. Wigneron, L. Kosolapova and F. Demontoux, "Temperature- and Texture-Dependent Dielectric Model for Moist Soils at 1.4 GHz," *IEEE Geoscience and Remote Sensing Letters,* vol. 10, no. 3, pp. 419-423, 2013. DOI: 10.1109/LGRS.2012.2207878.

[44] P. Beckmann and A. Spizzichino, The scattering of electromagnetic waves from rough surfaces, New York: Pergamon/MacMillan, 1963.

[45] F. T. Ulaby and M. A. El-rayes, "Microwave Dielectric Spectrum of Vegetation - Part II: Dual-Dispersion Model," *IEEE Transactions on Geoscience and Remote Sensing,* Vols. GE-25, no. 5, pp. 550-557, 1987. DOI: 10.1109/TGRS.1987.289833.

[46] S. L. Durden, J. D. Klein and H. A. Zebker, "Polarimetric radar measurements of a forested area near Mt Shasta," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 29, no. 3, pp. 444-450, 1991. DOI: 10.1109/36.79435.

[47] T. Le Toan, A. Beaudoin, J. Riom and D. Guyon, "Relating forest biomass to SAR data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 30, no. 2, pp. 403-411, 1992. DOI: 10.1109/36.134089.

[48] K. P. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 39, no. 11, pp. 2352-2363, 2001. DOI: 10.1109/36.964971.

[49] Y.-C. Lin and K. Sarabandi, "A Monte Carlo coherent scattering model for forest canopies using fractal-generated trees," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 37, no. 1, pp. 440-451, 1999. DOI: 10.1109/36.739083.

[50] T. Castel, A. Beaudoin, N. Floury, T. Le Toan, Y. Caraglio and J. F. Barczi, "Deriving forest canopy parameters for backscatter models using the AMAP architectural plant model," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 39, no. 3, pp. 571-583, 2001. DOI: 10.1109/36.911115.

[51] A. Beaudoin, T. Le Toan, S. Goze, E. Nezry, A. Lopes, E. Mougin, C. C. Hsu, H. C. Han, J. A. Kong and R. T. Shin, "Retrieval of forest biomass from SAR data,"

International Journal of Remote Sensing, vol. 15, no. 14, pp. 2777-2796, 1994. DOI: 10.1080/01431169408954284.

[52] E. S. Kasischke, N. L. Christensen and E. M. Haney, "Modeling of geometric properties of loblolly pine tree and stand characteristics for use in radar backscatter studies," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 32, no. 4, pp. 800-822, 1994. DOI: 10.1109/36.298009.

[53] G. B. West, J. H. Brown and B. J. Enquist, "A general model for the structure and allometry of plant vascular systems," *Nature,* vol. 400, pp. 664-667, 1999. DOI: 10.1038/23251.

[54] J. P. Richter ed., The Notebooks of Leonardo da Vinci, Vol. I, New York: Dover, 1883. Reprinted 1970.

[55] K. Shinozaki, K. Yoda, K. Hozumi and T. Kira, "A Quantitative Analysis of Plant Form - The Pipe Model Theory : I. Basic Analyses," *Japanese Journal of Ecology,* vol. 14, no. 3, pp. 97-105, 1964. DOI: 10.18960/seitai.14.3_97.

[56] K. Shinozaki, K. Yoda, K. Hozumi and T. Kira, "A Quantitative Analysis of Plant Form - The Pipe Model Theory : II. Further evidence of the theory and its application in forest ecology," *Japanese Journal of Ecology,* vol. 14, no. 4, pp. 133-139, 1964. DOI: 10.18960/seitai.14.4_133.

[57] L.-C. Jiang, R. Zhang and F.-R. Li, "Modeling Branch Length and Branch Angle with Linear Mixed Effects for Dahurian Larch," *Scientia Silvae Sinicae,* vol. 48, no. 5, pp. 53-60, 2012.

[58] T. A. McMahon and R. E. Kronauer, "Tree structures: Deducing the principle of mechanical design," *Journal of Theoretical Biology,* vol. 59, no. 2, pp. 443-466, 1976. DOI: 10.1016/0022-5193(76)90182-X.

[59] G. B. West, B. J. Enquist and J. H. Brown, "A general quantitative theory of forest structure and dynamics," *Proceedings of the National Academy of Sciences,* vol. 106, no. 17, pp. 7040-7045, 20098. DOI: 10.1073/pnas.0812294106.

[60] B. J. Enquist, G. B. West and J. H. Brown, "Extensions and evaluations of a general quantitative theory of forest structure and dynamics," *Proceedings of the National Academy of Sciences,* vol. 106, no. 17, pp. 7046-7051, 2009. DOI: 10.1073/pnas.0812303106.

[61] D. M. Le Vine, G. S. E. Lagerloef, F. R. Colomb, S. H. Yueh and F. A. Pellerano, "Aquarius: An Instrument to Monitor Sea Surface Salinity From Space," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 45, no. 7, pp. 2040-2050, 2007. DOI: 10.1109/TGRS.2007.898092.

[62] A. G. Fore et al., "Aquarius Scatterometer Calibration," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 8, no. 12, pp. 5424-5432, 2015. DOI: 10.1109/JSTARS.2015.2493449.

[63] G. Lagerloef, F. R. Colomb, D. Le Vine, F. Wentz, S. Yueh, C. Ruf, J. Lilly, J. Gunn, Y. Chao, A. Decharon, G. Feldman and C. Swift, "The Aquarius/SAC-D Mission: Designed to Meet the Salinity Remote-Sensing Challenge," *Oceonography,* vol. 21, no. 1, pp. 68-81, 2008.

[64] M. C. Dobson, F. T. Ulaby and L. E. Pierce, "Land-cover classification and estimation of terrain attributes using synthetic aperture radar," *Remote Sensing of Environment,* vol. 51, no. 1, pp. 199-214, 1995. DOI: 10.1016/0034-4257(94)00075-X.

[65] J. M. Kellndorfer, L. E. Pierce, M. C. Dobson and F. T. Ulaby, "Towards consistent regional-to-global-scale vegetation characterization using orbital SAR systems," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 36, no. 5, pp. 1396-1411, 1998. DOI: 10.1109/36.718844.

[66] D. E. Alsdorf, J. M. Melack, T. Dunne, L. A. K. Mertes, L. L. Hess and L. C. Smith, "Interferometric radar measurements of water level changes on the Amazon flood plain," *Nature,* vol. 404, pp. 174-177, 2000. DOI: 10.1038/35004560.

[67] L. L. Hess, J. M. Melack, S. Filoso and Y. Wang, "Delineation of inundated area and vegetation along the Amazon floodplain with the SIR-C synthetic aperture radar," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 33, no. 4, pp. 896-904, 1995. DOI: 10.1109/36.406675.

[68] S. H. Yueh, A. Fore, A. Freedman, M. J. Chaubell, W. Tang and G. Neumann, Aquarius Scatterometer Algorithm Theoretical Basis Document, Technical Document AQ-014-PS-0019, JPL/CalTech, 2012. ftp://podaac-ftp.jpl.nasa.gov/OceanWinds/aquarius/docs/v2/AQ-014-PS-0019_AquariusATBD_scatterometer.pdf.

[69] Aquarius/SAC-D. Aquarius Dataset Version 4.0. PO.DAAC, CA, USA. Dataset accessed 2017-04-25. ftp://podaac-ftp.jpl.nasa.gov/allData/aquarius/L2/V4

[70] Aquarius User Guide. Aquarius Dataset Version 4.0. JPL Document# JPL D-70012, Aquarius Document# AQ-010-UG-0008. 2015

[71] M. J. Brodzik, B. Billingsley, T. Haran, B. Raup and M. H. Savoie, "EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets," *International Journal of Geo-Information,* vol. 1, no. 1, pp. 32-45, 2012. DOI: 10.3390/ijgi1010032.

[72] G. Fischer, F. Nachtergaele, S. Prieler, H. T. van Velthuizen, L. Verelst and D. Wiberg, Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy, 2008. url: http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/

[73] J. Way et al., "The effect of changing environmental conditions on microwave signatures of forest ecosystems: preliminary results of the March 1988 Alaskan aircraft SAR experiment," *International Journal of Remote Sensing,* vol. 11, no. 7, pp. 1119-1144, 1990. DOI: 10.1080/01431169008955084 .

[74] X. Xu, C. Derksen, S. H. Yueh, R. S. Dunbar and A. Colliander, "Freeze/Thaw Detection and Validation Using Aquarius' L-Band Backscattering Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 9, no. 4, pp. 1370-1381, 2016. DOI: 10.1109/JSTARS.2016.2519347.

[75] M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley and X. Huang, "MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets," *Remote Sensing of Environment,* vol. 114, no. 1, pp. 168-182, 2010. DOI: 10.1016/j.rse.2009.08.016.

[76] S. Channan, K. Collins and W. R. Emanuel, Global mosaics of the standard MODIS land cover type data, University of Maryland and the Pacific Northwest National Laboratory, College Park, Maryland, USA, 2014. url: http://glcf.umd.edu/data/lc/

[77] K. Saleh, A. Porte, D. Guyon, P. Ferrazzoli and J.-P. Wigneron, "A forest geometric description of a maritime pine forest suitable for discrete microwave models," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 43, no. 9, pp. 2024-2035, 2005. DOI: 10.1109/TGRS.2005.853568.

[78] J. R. Wang and T. J. Schmugge, "An Empirical Model for the Complex Dielectric Permittivity of Soils as a Function of Water Content," *IEEE Transactions on Geoscience and Remote Sensing,* Vols. GE-18, no. 4, pp. 288-295, 1980. DOI: 10.1109/TGRS.1980.350304.

[79] M. T. Hallikainen, F. T. Ulaby, M. C. Dobson, M. A. El-rayes and L. Wu, "Microwave Dielectric Behavior of Wet Soil-Part 1: Empirical Models and Experimental Observations," *IEEE Transactions on Geoscience and Remote Sensing,* Vols. GE-23, no. 1, pp. 25-34, 1985. DOI: 10.1109/TGRS.1985.289497.

[80] Y. Oh, K. Sarabandi and F. T. Ulaby, "An empirical model and an inversion technique for radar scattering from bare soil surfaces," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 30, no. 2, pp. 370-381, 1992. DOI: 10.1109/36.134086.

[81] R. Nelson, K. J. Ranson, G. Sun, D. S. Kimes, V. Kharuk and P. Montesano, "Estimating Siberian timer volume using MODIS and ICEsat/GLAS," *Remote Sensing of Environment,* vol. 113, no. 3, pp. 691-701, 2009. DOI: 10.1016/j.rse.2008.11.010.

[82] A. E. Zanne, G. Lopez-Gonzalez, D. A. Coomes, J. Ilic, S. Jansen, S. L. Lewis, R. B. Miller, N. G. Swenson, M. C. Wiemann and J. Chave, Global wood density database, 2009. DOI: 10.5061/dryad.234/1.

[83] H. Margolis, G. Sun, P. M. Montesano and R. F. Nelson, "NACP LiDAR-based Biomass Estimates, Boreal Forest Biome, North America, 2005-2006", ORNL DAAC, Oak Ridge, Tennessee, USA, 2015. DOI: 10.3334/ORNLDAAC/1273.DOI: 10.3334/ORNLDAAC/1273.

[84] D. Shepashenko, A. Shvidenko and S. Nilsson, "Phytomass (live biomass) and carbon of Siberian forests," *Biomass and Bioenergy,* vol. 14, no. 1, pp. 21-31, 1998. DOI: 10.1016/S0961-9534(97)10006-X.

[85] J. Friesen, S. C. Steele-Dunne and N. van de Giesen, "Diurnal Differences in Global ERS Scatterometer Backscatter Observations of the Land Surface," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 50, no. 7, pp. 2592-2602, 2012. DOI: 10.1109/TGRS.2012.2193889.

[86] USDA Natural Resources Conservation Service (2017). SNOwpack TELemetry Network (SNOTEL). NRCS. https://data.nal.usda.gov/dataset/snowpack-telemetry-network-snotel.

[87] W. A. Dorigo, W. Wagner, R. Hohensinn, S. Hahn, C. Paulik, A. Xaver, A. Gruber, M. Drusch, S. Mecklenburg, P. van Oevelen, A. Robock and T. Jackson, "The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements," *Hydrology and Earth System Sciences,* vol. 15, pp. 1675-1698, 2011. DOI: 10.5194/hess-15-1675-2011.

[88] W. A. Dorigo, A. Xaver, M. Vreugdenhil, A. Gruber, A. Hegyjova, A. D. Sanchis-Dufau, D. Zamojski, C. Cordes, W. Wagner and M. Drusch, "Global Automated Quality Control of In situ Soil Moisture data from the International Soil Moisture Network," *Vadose Zone Journal*, vol. 12, no. 3, 2013.

[89] M. Segura and M. Kanninen, "Allometric Models for Tree Volume and Total Aboveground Biomass in a Tropical Humid Forest in Costa Rica," *Biotropica,* vol. 37, no. 1, pp. 2-8, 2005. DOI: 10.1111/j.1744-7429.2005.02027.x.

[90] S. Brown and A. E. Lugo, "Aboveground biomass estimates for tropical moist forests of the Brazilian Amazon," *Interciencia,* vol. 17, no. 1, pp. 8-18, 1992.

[91] E. M. Nogueira, P. M. Fearnside, B. W. Nelson, R. I. Barbosa and E. W. H. Keizer, "Estimates of forest biomass in the Brazilian Amazon: New allometric equations and adjustments to biomass from wood-volume inventories," *Forest Ecology and Management,* vol. 256, no. 11, pp. 1853-1867, 2008. DOI: 10.1016/j.foreco.2008.07.022.

[92] Y. Wang, L. Hess, S. Filoso and J. Melack, "Understanding the Radar Backscattering from Flooded and Nonflooded Amazonian Forests: Results from Canopy Backscattering Modeling," *Remote Sensing of Environment,* vol. 54, pp. 324-332, 1995. DOI: 10.1016/0034-4257(95)00140-9.

[93] T. v. Emmerik, S. Steele-Dunne, A. Paget, R. S. Oliveira, P. R. L. Bittencourt, F. d. V. Barros and N. v. d. Giesen, "Water stress detection in the Amazon using radar," *Geophysical Research Letters,* vol. 44, no. 13, pp. 6841-6849, 2017. DOI: 10.1002/2017GL073747.

[94] M. Satake and H. Hanado, "Diurnal change of Amazon rain forest sigma0 observed by Ku-band spaceborne radar," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 42, no. 6, pp. 1127-1134, 2004. DOI: 10.1109/TGRS.2004.825589.

[95] A. G. Konings, Y. Yu, L. Xu, Y. Yang, D. S. Schimel and S. S. Saatchi, "Active microwave observations of diurnal and seasonal variations of canopy water content across the humid African tropical forests," *Geophysical Research Letters,* vol. 44, no. 5, pp. 2290-2299, 2017. DOI: 10.1002/2016GL072388.

[96] D. E. Prado, "What is the Gran Chaco vegetation in South America? I. A review," *Candollea,* vol. 48, no. 1, pp. 145-172, 1993.

[97] M. A. El-rayes and F. T. Ulaby, "Microwave Dielectric Spectrum of Vegetation-Part I: Experimental Observations," *IEEE Transactions on Geoscience and*

Remote Sensing, Vols. GE-25, no. 5, pp. 541-549, 1987. DOI: 10.1109/TGRS.1987.289832.

[98] A. Konings, M. Piles, K. Rötzer, K. A. McColl, S. Chan and D. Entekhabi, "Vegetation optical depth and scattering albedo retrieval using time series of dual-polarized L-band radiometer observations," *Remote Sensing of Environment*, vol. 172, pp. 178-189, 2016. DOI: 10.1016/j.rse.2015.11.009.

[99] K. Rötzer, C. Montzka, D. Entekhabi, A. G. Konings, K. A. McColl, M. Piles and H. Vereecken, "Relationship Between Vegetation Microwave Optical Depth and Cross-Polarized Backscatter from Multiyear Aquarius Observations," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 10, pp. 4493-4503, 2017. DOI: 10.1109/JSTARS.2017.2716638.

[100] Y. H. Kerr, P. Waldteufel, P. Richaume, J. P. Wigneron, P. Ferrazzoli, A. Mahmoodi, A. Al Bitar, F. Cabot, C. Gruhier, S. E. Juglea, D. Leroux, A. Mialon and S. Delwart, "The SMOS Soil Moisture Retrieval Algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1384-1403, 2012. DOI: 10.1109/TGRS.2012.2184548.

[101] R. Rahmoune, P. Ferrazzoli, Y. H. Kerr and P. Richaume, "SMOS Level 2 Retrieval Algorithm over Forests: Description and Generation of Global Maps," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 3, pp. 1430-1439, June 2013. DOI: 10.1109/JSTARS.2013.2256339.

[102] D. Entekhabi et al., "The Soil Moisture Active Passive (SMAP) Mission," *Proceedings of the IEEE*, vol. 98, no. 5, pp. 704-716, 2010. DOI: 10.1109/JPROC.2010.2043918.

[103] SMAP data 2015 (NASA). Dataset: SMAP SMAP_L1B_S0_LoRes_V3. Retrieved from ASF DAAC 13 June 2018. DOI: 10.5067/NEWTOCOKVZHF.

[104] B. Weiss and M. Maatyan, SMAP Level 1B_S0_LoRes Product Specification Document, JPL D-72544, 2016. url: https://media.asf.alaska.edu/uploads/smap%20pdf/d-72544-b_smap_l1b_s0_lores_product_specification_document.pdf

[105] R. West, SMAP L1B_S0, L1C_S0 Algorithm Theoretical Basis Document, 2014. url: https://smap.jpl.nasa.gov/system/internal_resources/details/original/280_L1C_S0_RevA_web.pdf

[106] X. Xu, R. S. Dunbar, C. Derksen, A. Colliander, Y. Kim and J. S. Kimball, SMAP L3 Radiometer Global and Northern Hemisphere Daily 36 km EASE-Grid Freeze/Thaw State, Version 2, Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center, 2018. Accessed Feb 12, 2019. DOI: 10.5067/YN94K53QM061.

[107] H. Stephen, "Microwave Remote Sensing of Saharan Ergs and Amazon Vegetation," 2006. PhD Dissertation. url: https://scholarsarchive.byu.edu/etd/495.

[108] H. Stephen and D. G. Long, "Microwave Backscatter Modeling of Erg Surfaces in the Sahara Desert," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 43, no. 2, pp. 238-247, 2005. DOI: 10.1109/TGRS.2004.840646.

[109] Z. Bartalis, K. Scipal and W. Wagner, "Azimuthal anisotrophy of scatterometer measurements over land," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, no. 8, pp. 2083-2092, 2006. DOI: 10.1109/TGRS.2006.872084.

[110] J. P. Sparks, G. S. Campbell and R. A. Black, "Water content, hydraulic conductivity, and ice formation in winter stems of Pinus contorta: a TDR case study," *Oecologia,* vol. 127, pp. 468-475, 2001. DOI: 10.1007/s004420000587.

[111] K. McDonald, M. C. Dobson and F. T. Ulaby, "Using Mimics To Model L-band Multiangle and Multitemporal Backscatter From A Walnut Orchard," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 28, no. 4, pp. 477-491, 1990. DOI: 10.1109/TGRS.1990.572925.

[112] K. C. McDonald, R. Zimmermann and J. S. Kimball, "Diurnal and Spatial Variation of Xylem Dielectric Constant in Norway Spruce (Picea abies [L.] Karst.) as Related to Microclimate, Xylem Sap Flow, and Xylem Chemistry," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 40, no. 9, pp. 2063-2082, 2002. DOI: 10.1109/TGRS.2002.803737.

[113] A. R. Monteith and L. M. Ulander, "Temporal Survey of P- and L-band Polarimetric Backscatter in Boreal Forests," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 11, no. 10, pp. 3564-3577, 2018. DOI: 10.1109/JSTARS.2018.2814825.

[114] A. C. Paget, D. G. Long and N. M. Madsen, "RapidScat Diurnal Cycles Over Land," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 54, no. 6, pp. 3336-3344, 2016. DOI: 10.1109/TGRS.2016.2515022.

[115] George Huffman, GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V05, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), 2017. Accessed: 8 February 2019. DOI: 10.5067/GPM/IMERG/3B-HH/05.

[116] N. Wagner, K. Emmerich, F. Bonitz and K. Kupfer, "Experimental Investigations on the Frequency- and Temperature-Dependent Dielectric Material Properties of Soil.," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 49, no. 7, pp. 2518-2530, 2011. DOI: 10.1109/TGRS.2011.2108303.

[117] R. Rahmoune, P. Ferrazzoli, Y. K. Singh, Y. H. Kerr, P. Richaume and A. Al Bitar, "SMOS Retrieval Results Over Forests: Comparisons With Independent Measurements," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 7, no. 9, pp. 3858-3866, 2014. DOI: 10.1109/JSTARS.2014.2321027.

[118] M. S. Burgin, L. Mandrake, G. B. Doran, B. D. Bue and J. J. van Zyl, "Soil Moisture Estimation by Linear Regression from Smap Polarimetric Radar Data with Aquarius Derived Coefficients," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018.

[119] M. S. Burgin and J. J. van Zyl, "Analysis of Polarimetric Radar Data and Soil Moisture From Aquarius: Towards a Regression-Based Soil Moisture Estimation Algorithm," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 9, no. 8, pp. 3497-3504, 2016. DOI: 10.1109/JSTARS.2016.2526899.

[120] Y. Kim and J. J. van Zyl, "A Time-Series Approach to Estimate Soil Moisture Using Polarimetric Radar Data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 47, no. 8, pp. 2519-2527, 2009. DOI: 10.1109/TGRS.2009.2014944.

[121] W. Wagner, G. Lemoine and H. Rott, "A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data," *Remote Sensing of Environment,* vol. 70, no. 2, pp. 191-207, 1999. DOI: https://doi.org/10.1016/S0034-4257(99)00036-X.

[122] Y. Zhang, M. G. Schaap and Y. Zha, "A High-Resolution Global Map of Soil Hydraulic Properties Produced by a Hierarchical Parameterization of a Physically-Based Water Retention Model," *Water Resources Research,* vol. 54, no. 12, pp. 9774-9790, 2018. DOI: 10.7910/DVN/UI5LCE.

[123] J. Peng, P. Mohammed, J. Chaubell, S. Chan, S. Kim, N. Das, S. Dunbar, R. Bindlish and X. Xu, Soil Moisture Active Passive (SMAP) L1-L3 Ancillary Static Data, Version 1. Soil Attributed Dataset. Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center, 2019. Accessed April 20, 2019. doi: 10.5067/HB8BPJ13TDQJ.

[124] N. Das, SMAP Ancillary Data Report: Soil Attributes. JPL D-53058. https://smap.jpl.nasa.gov/system/internal_resources/details/original/286_044_soil_attrib.pdf.

[125] G. L. Schaefer, M. H. Cosh and T. J. Jackson, "The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN)," *Journal of Atmospheric and Oceanic Technology,* vol. 24, pp. 2073-2077, 2007. DOI: 10.1175/2007JTECHA930.1.

[126] J. E. Bell et al, "U.S. Climate Reference Network soil moisture and temperature observations," *Journal of Hydrometeorology,* vol. 14, pp. 977-988, 2013. DOI: 10.1175/JHM-D-12-0146.1.

[127] M. Zreda, W. J. Shuttleworth, X. Zeng, C. Zweck, D. Desilets, T. Franz and R. Rosolem, "COSMOS: the COsmic-ray Soil Moisture Observing System.," *Hydrology and Earth System Sciences,* vol. 16, pp. 4079-4099, 2012. DOI: 10.5194/hess-16-4079-2012.

[128] M. Zreda, D. Desilets, T. P. Ferré and R. L. Scott, "Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons," *Geophysical Research Letters,* vol. 35, no. 21, p. L21402, 2008. DOI: 10.1029/2008GL035655.

[129] S.-b. Kim, J. van Zyl, S. Dunbar, J. Johnson, M. Moghaddam and L. Tsang, SMAP Algorithm Theoretical Basis Document, Level 2 & 3 Soil Moisture (Active) Data Products, Revision B, 2015. https://nsidc.org/sites/nsidc.org/files/technical-references/L2%263_SM_A_RevB_web151031.pdf.

[130] P. O'Neill, S. Chan, E. Njoku, T. Jackson and R. Bindlish, SMAP Algorithm Theoretical Basis Document, Level 2 & 3 Soil Moisture (Passive) Data Products, Revision B, 2015. JPL D-66480.

[131] S. Kim, J. van Zyl, R. S. Dunbar, E. G. Njoku, J. T. Johnson, M. Moghaddam and L. Tsang, 2016. SMAP L3 Radar Global Daily 3 km EASE-Grid Soil Moisture, Version 3. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. Accessed Jan 4, 2018.DOI: 10.5067/IGQNPB6183ZX.

[132] P. E. O'Neill, S. Chan, E. G. Njoku, T. Jackson and R. Bindlish, 2016. SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 3. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. Accessed August 8, 2016.DOI: 10.5067/7MINGFDCZTES.

[133] J. S. Famiglietti, D. Ryu, A. A. Berg, M. Rodell and T. J. Jackson, "Field observations of soil moisture variability across scales," *Water Resources Research,* vol. 44, p. W01423, 2008. DOI: 10.1029/2006WR005804.

[134] A. Colliander, H. Al Jassar, W. Dorigo, J. Martinez-Fernandez, C. Montzka, M. Seyfried et al. 2017. SMAP/In Situ Core Validation Site Land Surface Parameters Match-Up Data, ver. 1. Boulder, Colorado USA, NASA NSIDC DAAC, Accessed Apr 1, 2019. DOI: 10.5067/DXAVIXLY18KM.

[135] J. R. Wait, "Scattering of a Plane Wave from a Circular Dielectric Cylinder at Oblique Incidence," *Canadian Journal of Physics,* vol. 33, no. 5, pp. 189-195, 1955. DOI: 10.1139/p55-024.

[136] G. T. Ruck, D. E. Barrick, W. D. Stuart and C. K. Krichbaum, Radar Cross Section Handbook Volume I, Chapter 4, New York-London: Plenum Press, 1970.

[137] R. T. Wang and H. C. van de Hulst, "Application of the exact solution for scattering by an infinite cylinder to the estimation of scattering by a finite cylinder," *Applied Optics,* vol. 34, no. 15, pp. 2811-2821, 1995. DOI: 10.1364/AO.34.002811.

[138] L. Tsang, J. A. Kong and K. H. Ding, Scattering of Electromagnetic Waves: Theories and Applications, Wiley Interscience, 2000.

[139] T. Liao, S. Kim, S. Tan, L. Tsang, C. Su and T. J. Jackson, "Multiple Scattering Effects With Cyclical Correction in Active Remote Sensing of Vegetated Surface Using Vector Radiative Transfer Theory," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 9, no. 4, pp. 1414-1429, 2016. DOI: 10.1109/JSTARS.2015.2505638.

[140] P. Ferrazzoli, L. Guerriero and D. Solimini, "Numerical model of microwave backscattering and emission from terrain covered with vegetation," *Applied Computational Electromagnetics Society Journal,* vol. 6, no. 1, pp. 175-191, 1991.

[141] S. Twomey, H. Jacobowitz and H. B. Howell, "Matrix Methods for Multiple-Scattering Problems," *Journal of the Atmospheric Sciences,* vol. 23, pp. 289-296, 1966. DOI: 10.1175/1520-0469(1966)023<0289:MMFMSP>2.0.CO;2.

[142] H. C. van de Hulst, Multiple light scattering. Tables, formulas and applications. Volume 2, New York: Academic Press, 1980.

[143] L. Tsang and A. Ishimaru, "Backscattering enhancement of random discrete scatterers," *Journal of the Optical Society of America A*, vol. 1, no. 8, pp. 836-839, 1984. DOI: 10.1364/JOSAA.1.000836.

[144] C. Dahon, L. Ferro-Famil, C. Titin-Schnaider and E. Pottier, "Computing the double-bounce reflection coherent effect in an incoherent electromagnetic scattering model," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 2, pp. 241-245, 2006. DOI: 10.1109/LGRS.2005.863397.

[145] J.-P. Wigneron et al., "Modelling the passive microwave signature from land surfaces: A review of recent results and application to the L-band SMOS & SMAP soil moisture retrieval algorithms," *Remote Sensing of Environment*, vol. 192, pp. 238-262, 2017. DOI: 10.1016/j.rse.2017.01.024.

[146] H. C. van de Hulst, Light Scattering by Small Particles, New York: Dover, 1981.

[147] F. T. Ulaby, R. K. Moore and A. K. Fung, Microwave remote sensing fundamentals and radiometry. Vol. I of Microwave remote sensing: Active and passive, Norwood, MA: Artech House, 1986.

[148] Integrated Hydrologic Database of the Secretary of Water Resources, Argentina. Accessed 21 Feb 2019. http://bdhi.hidricosargentina.gob.ar.

[149] Tropical Rainfall Measuring Mission (TRMM) (2011), TRMM (TMPA) Rainfall Estimate L3 3 hour 0.25 degree x 0.25 degree V7, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: July 2, 2017DOI: 10.5067/TRMM/TMPA/3H/7.

Appendix A

BISTATIC SCATTERING FROM A DIELECTRIC CYLINDER

The expression for the far-field bistatic scattering from a dielectric cylinder is given in this section (backscatter being a special case of bistatic scattering). There is a known exact analytical solution only for infinitely long cylinders, first derived by Wait [135]; here we follow more closely the notation and expressions by Bohren and Huffman [30] for the case of a vertically oriented, infinitely long dielectric cylinder in vacuum. For long (compared to the wavelength) but finite cylinders, we apply the approximation made by Ruck and Barrick [136] and van Zyl and Kim [11], by multiplying by a sinc function to the solution for infinite cylinders.

Let $k = 2\pi/\lambda$ be the wavenumber, ε the dielectric constant of the cylinder, r the radius of the cylinder, and L the cylinder length. Let θ_i be the angle of the incident wave from the vertical, and ϕ_i and ϕ_s the incident and scattered azimuth angles, following the FSA/BSA conventions as in Chapter 2. Note that for infinitely long cylinders, scattering only takes place on a forward scattering cone making the same angle with the vertical cylinder axis, i.e. $\theta_s = \pi - \theta_i$. Also, by the azimuthal symmetry, the dependence on ϕ_i and ϕ_s occurs only through the relative angle $\phi_s - \phi_i$ between them.

Let

$$\xi = kr \sin\theta_i$$

$$\eta = kr\sqrt{\varepsilon - \cos^2\theta_i}$$

$$A_n = i\xi[\xi J_n'(\eta)J_n(\xi) - \eta J_n(\eta)J_n'(\xi)]$$

$$B_n = \xi[\varepsilon\xi J_n'(\eta)J_n(\xi) - \eta J_n(\eta)J_n'(\xi)]$$

$$C_n = n\eta \cos\theta_i J_n(\eta)J_n(\xi)[(\xi/\eta)^2 - 1]$$

$$D_n = n\eta \cos\theta_i J_n(\eta)H_n^{(1)}(\xi)[(\xi/\eta)^2 - 1]$$

$$V_n = \xi\left[\varepsilon\xi J_n'(\eta)H_n^{(1)}(\xi) - \eta J_n(\eta)H_n^{(1)'}(\xi)\right]$$

$$W_n = i\xi\left[\eta J_n(\eta)H_n^{(1)'}(\xi) - \xi J_n'(\eta)H_n^{(1)}(\xi)\right]$$

where J_n is the Bessel function of the first kind, order , J_n' is its derivative, $H_n^{(1)}$ is the Hankel function, and $H_n^{(1)'}$ its derivative. Further compute

$$a_{n\mathrm{I}} = \frac{C_nV_n - B_nD_n}{W_nV_n + iD_n^2}\,, \qquad b_{n\mathrm{I}} = \frac{W_nB_n + iD_nC_n}{W_nV_n + iD_n^2}$$

$$a_{n\mathrm{II}} = -\frac{A_nV_n - iC_nD_n}{W_nV_n + iD_n^2}\,, \qquad b_{n\mathrm{II}} = -i\frac{C_nW_n + A_nD_n}{W_nV_n + iD_n^2} = -a_{n\mathrm{I}}$$

$$T_1 = b_{0\mathrm{I}} + 2\sum_{n=1}^{\infty}(-1)^n\cos(n[\phi_s - \phi_i - \pi])\,b_{n\mathrm{I}}$$

$$T_1 = a_{0\mathrm{II}} + 2\sum_{n=1}^{\infty}(-1)^n\cos(n[\phi_s - \phi_i - \pi])\,a_{n\mathrm{I}}$$

$$T_3 = 2i\sum_{n=1}^{\infty}(-1)^n\sin(n[\phi_s - \phi_i - \pi])\,a_{n\mathrm{I}}$$

$$T_4 = -T_3$$

We keep only the first 20 terms in the infinite series for computation, as a reasonable trade-off between computational accuracy and speed.

The expression for our approximation for the (dimensionless) scattering matrix of a finite cylinder is then

$$[S] = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} = \frac{ikL}{\pi}\frac{\sin\theta_s}{\sin\theta_i}\,\mathrm{sinc}\left(\frac{kL}{2}(\cos\theta_i + \cos\theta_s)\right)\begin{bmatrix} -T_2 & T_3 \\ -T_4 & T_1 \end{bmatrix} \qquad \text{(A. 1)}$$

with

$$\mathrm{sinc}(x) = \frac{\sin x}{x}\,.$$

Here equation (A. 1) is written for the BSA convention. For the FSA convention, the signs of S_{hh} and S_{hv} should be flipped (see Section 0). For finite cylinders, we need not impose

$\theta_s = \pi - \theta_i$, but the sinc function keeps the scattering small if θ_s departs from this significantly and the cylinder length is large compared to the wavelength. The approximation is not expected to be good if the incident direction is close to the cylinder axis. In our computer implementation, we set the scattering matrix to zero if we are within 5 degrees of the cylinder axis. There are a few further caveats to the approximation of equation (A. 1). Realize that in this approximation, there is apparently no reflection of the induced current from the ends of the cylinder and thus no resonance length. Also note that reciprocity may not be respected exactly by equation (A. 1) when $\theta_i \neq \pi/2 \neq \theta_s$, in the sense that $S_{hv}(\theta_i, \phi_i, \theta_s, \phi_s) \neq S_{vh}(\theta_s, \phi_s + \pi, \theta_i, \phi_i + \pi)$. The physical optics approximation of multiplying by the sinc function has only been rigorously shown for the case $\theta_i = \pi/2$ [30] [137]. We still use approximation (A. 1) in spite of these issues because of the following reasons. Firstly, because we intend to use finite cylinders to model conceptual sections of branches and other similar linear structures in vegetation, the neglect of the end reflections may be reasonable. Secondly, because the majority of direct single-scattering backscatter from vegetation will come from cylinders oriented near $\theta_i = \pi/2$, the approximation is valid where it matters most. Any potential issues with using approximation (A. 1) are thus more likely to come from using it for calculations involving multiple scattering, where all permutations of bistatic angles may be involved.

The above expression was for a vertically oriented cylinder. For a cylinder with axis arbitrarily oriented in the direction (θ_c, ϕ_c), first consider the local rotated frame in which the cylinder is vertically oriented, and let (θ_{ic}, ϕ_{ic}) and (θ_{sc}, ϕ_{sc}) be the local incident and scattered directions as in the BSA convention of Figure 1.2. Using the scattering matrix $[S(\theta_{ic}, \phi_{ic}, \theta_{sc}, \phi_{sc})]$ in this local rotated frame, we will then apply coordinate rotations to obtain the scattering matrix in the global frame [11].

Let

$$\hat{\mathbf{c}} = \sin\theta_c \cos\phi_c \, \hat{\mathbf{x}} + \sin\theta_c \sin\phi_c \, \hat{\mathbf{y}} + \cos\theta_c \, \hat{\mathbf{z}}$$

$$\hat{\mathbf{h}}_{ic} = \frac{\hat{\mathbf{c}} \times \hat{\mathbf{k}}_i}{|\hat{\mathbf{c}} \times \hat{\mathbf{k}}_i|} , \qquad \hat{\mathbf{v}}_{ic} = \hat{\mathbf{h}}_{ic} \times \hat{\mathbf{k}}_i , \qquad \hat{\mathbf{h}}_{sc} = \frac{\hat{\mathbf{c}} \times \hat{\mathbf{k}}_s}{|\hat{\mathbf{c}} \times \hat{\mathbf{k}}_s|} , \qquad \hat{\mathbf{v}}_{sc} = \hat{\mathbf{h}}_{sc} \times \hat{\mathbf{k}}_i .$$

The bistatic scattering matrix (in the global frame) from a dielectric cylinder is then

$$[\mathbf{S}(\theta_i, \phi_i, \theta_s, \phi_s, \theta_c, \phi_c)] =$$

$$\begin{bmatrix} \hat{\mathbf{h}}_s \cdot \hat{\mathbf{h}}_{sc} & \hat{\mathbf{h}}_s \cdot \hat{\mathbf{v}}_{sc} \\ \hat{\mathbf{v}}_s \cdot \hat{\mathbf{h}}_{sc} & \hat{\mathbf{v}}_s \cdot \hat{\mathbf{v}}_{sc} \end{bmatrix} [\mathbf{S}(\theta_{ic}, \phi_{ic}, \theta_{sc}, \phi_{sc})] \begin{bmatrix} \hat{\mathbf{h}}_{ic} \cdot \hat{\mathbf{h}}_i & \hat{\mathbf{h}}_{ic} \cdot \hat{\mathbf{v}}_i \\ \hat{\mathbf{v}}_{ic} \cdot \hat{\mathbf{h}}_i & \hat{\mathbf{v}}_{ic} \cdot \hat{\mathbf{v}}_i \end{bmatrix}. \tag{A. 2}$$

Note that (θ_{ic}, ϕ_{ic}) can easily be found from $\theta_c, \phi_c, \theta_i, \phi_i$ using

$$\hat{\mathbf{k}}_i = \begin{bmatrix} \sin\theta_i \cos\phi_i \\ \sin\theta_i \sin\phi_i \\ -\cos\theta_i \end{bmatrix}$$

$$\begin{bmatrix} \sin\theta_{ic} \cos\phi_{ic} \\ \sin\theta_{ic} \sin\phi_{ic} \\ -\cos\theta_{ic} \end{bmatrix} = \begin{bmatrix} \cos\theta_c \cos\phi_c & \cos\theta_c \sin\phi_c & -\sin\theta_c \\ -\sin\phi_c & \cos\phi_c & 0 \\ \sin\theta_c \cos\phi_c & \sin\theta_c \sin\phi_c & \cos\theta_c \end{bmatrix} \hat{\mathbf{k}}_i.$$

Appendix B

MULTIPLE SCATTERING CORRECTION FACTOR

In this Appendix, the estimation of the multiple-scattering correction factors $\mathcal{F}_{\mathrm{HH}}(\tau_{cn}(\theta_{i0}),\theta_{i0},\varepsilon_v)$, $\mathcal{F}_{\mathrm{VV}}(\tau_{cn}(\theta_{i0}),\theta_{i0},\varepsilon_v)$ and $\mathcal{F}_{\mathrm{HV}}(\tau_{cn}(\theta_{i0}),\theta_{i0},\varepsilon_v)$ to the canopy vegetation backscatter term (Section 2.2) is described. Within this Appendix, the subscript i for the incidence angle θ_i is temporarily changed to $i0$ to avoid notational confusion (for reasons that would become apparent). The overall approach is as follows. For the co-polarized corrections $\mathcal{F}_{\mathrm{HH}}$ and $\mathcal{F}_{\mathrm{VV}}$, the method of radiative transfer is used to estimate the backscatter without making the single-scattering assumption. A further correction is made to account for coherent backscatter enhancement not modelled by the radiative transfer equations. The cross-polarization correction $\mathcal{F}_{\mathrm{HV}}$ is estimated from the co-polarization corrections via Monte-Carlo estimates of the ratio between cross-polarization and co-polarization returns for double scattering.

The radiative transfer equation for our scattering problem can be written as [40]

$$\frac{d\mathbf{I}}{ds} = -[\mathbf{K}]\mathbf{I} + \int_{4\pi} \langle \mathbf{M} \rangle \mathbf{I}\, d\Omega\,. \tag{B.1}$$

Here s is the distance along the propagation direction being considered, and $[\mathbf{K}]$ is the extinction coefficient matrix, often denoted $\boldsymbol{\kappa}_e$ in the literature, but here we change the notation slightly to avoid excessive subscripts and confusion with the field extinction. $\mathbf{I}$ is the modified Stokes vector satisfying

$$\mathbf{I}\, d\Omega = \begin{bmatrix} I_v \\ I_h \\ U \\ V \end{bmatrix} d\Omega = \frac{1}{\eta} \begin{bmatrix} \langle |E_v|^2 \rangle \\ \langle |E_h|^2 \rangle \\ 2\mathrm{Re}\langle E_v E_h^* \rangle \\ 2\mathrm{Im}\langle E_v E_h^* \rangle \end{bmatrix} \tag{B.2}$$

with units of power per unit solid angle per unit area (area perpendicular to the propagation direction), where η is the charateristic impedance of the medium (≈377Ω for free space). $\langle \mathbf{M} \rangle$

is the average Stokes matrix density, with dimensions of area per unit volume; the entries of $\langle \mathbf{M} \rangle$ for the form of the radiative transfer equation that shall be used here will be written out explicitly later.

For our horizontal slab geometry, with θ the angle from the vertical, let $\mathbf{G}$ be the power per solid angle per unit horizontal area, instead of power per unit perpendicular area. Then $\mathbf{G} = \mathbf{I}|\cos\theta|$, and the radiative transfer equation becomes

$$\frac{d\mathbf{G}}{dz} = -\frac{1}{|\cos\theta|}[\mathbf{K}]\mathbf{G} + \int_{4\pi} \langle \mathbf{M} \rangle \mathbf{G} \frac{1}{|\cos\theta|}\, d\Omega\,. \tag{B. 3}$$

There are various ways to solve the radiative transfer equation [40, 138-141]. A numerical approach to solving the radiative transfer equation may be to solve a linear system as described below. Discretize the 4π sphere into $2J$ directions, J upwards and J downwards, and use subscripts u or d for "upwards" or "downwards" respectively. (In the computer implementation, $J = 162$ was used, with angular intervals of 10° in elevation and 20° in azimuth). Subscripts s, i, or j , each running from 1 to J, will be reserved as indices for these directions, with subscript i typically used for an "incident" direction, and subscript s typically used for a "scattered" direction; when there is a need to subscript both "scattered" and "incident" simultaneously, a semicolon will be used to separate them, e.g. $(s; i)$ or $(us; di)$. Associate solid angle $\Delta\Omega_{uj}$ or solid angle $\Delta\Omega_{dj}$ with each direction, such that

$$2\pi = \sum_{j=1}^{J} \Delta\Omega_{uj} \quad , \qquad 2\pi = \sum_{j=1}^{J} \Delta\Omega_{dj} \tag{B. 4}$$

Let $z^{(l)} = \left(l - \frac{1}{2}\right)\Delta z\,, \Delta z = (Z_2 - Z_1)/l_{max}$ where $(Z_2 - Z_1)$ is the total height of the vegetation layer, and l_{max} is the total number of thin sublayers that whole layer is being divided into for computation. l is an integer indexing the sublayers and Δz is the thickness of each sublayer. The highest sublayer is thus associated with index $l = l_{max}$, and the lowest sublayer with index $l = 1$. Our discretized radiative transfer equation then reads, for each $s = 1, \ldots, J$ and for each l (except for the upper and lower boundaries)

$$\frac{d}{dz_+}\mathbf{G}_{us}^{(l)} = -\frac{1}{|\cos\theta_{us}|}[\mathbf{K}]_{us}\mathbf{G}_{us}^{(l)} + \sum_{i=1}^{J}\left(\langle\mathbf{M}_{us;ui}\rangle\frac{\mathbf{G}_{ui}^{(l)}}{|\cos\theta_{ui}|} + \langle\mathbf{M}_{us;di}\rangle\frac{\mathbf{G}_{di}^{(l)}}{|\cos\theta_{di}|}\right) \quad \text{(B. 5a)}$$

$$\frac{d}{dz_-}\mathbf{G}_{ds}^{(l)} = -\frac{1}{|\cos\theta_{ds}|}[\mathbf{K}]_{ds}\mathbf{G}_{us}^{(l)} + \sum_{i=1}^{J}\left(\langle\mathbf{M}_{ds;ui}\rangle\frac{\mathbf{G}_{ui}^{(l)}}{|\cos\theta_{ui}|} + \langle\mathbf{M}_{ds;di}\rangle\frac{\mathbf{G}_{di}^{(l)}}{|\cos\theta_{di}|}\right) \quad \text{(B. 5b)}$$

In these equations, $\mathbf{G}_{uj}^{(l)}$ is a 4x1 vector of Stokes parameters to be interpreted as the power (in watts) through a unit horizontal plane area (1m^2) through solid angle $\Delta\Omega_{uj} \approx \sin\theta_{uj}\,\Delta\theta\Delta\phi$ centred at the direction (θ_{uj}, ϕ_{uj}), at position $z = \left(l - \frac{1}{2}\right)\Delta z$ (the l -th sublayer), and likewise for $\mathbf{G}_{dj}^{(l)}$ and $\Delta\Omega_{uj}$ (replace subscript u for "upwards" by d for "downwards"). The extinction coefficient matrix in the direction (θ_{us}, ϕ_{us}) is

$$[\mathbf{K}]_{us} = n_{cn}\begin{bmatrix} 2\langle\kappa_{v,cn}(\theta_{us})\rangle & 0 & 0 & 0 \\ 0 & 2\langle\kappa_{h,cn}(\theta_{us})\rangle & 0 & 0 \\ 0 & 0 & \langle\kappa_{v,cn}(\theta_{us}) + \kappa_{h,cn}(\theta_{us})\rangle & 0 \\ 0 & 0 & 0 & \langle\kappa_{v,cn}(\theta_{us}) + \kappa_{h,cn}(\theta_{us})\rangle \end{bmatrix}_{us} \quad \text{(B. 6)}$$

and likewise for $[\mathbf{K}]_{ds}$ (replace u by d). Azimuthal symmetry in the cylinder orientation distribution is assumed, removing the dependence of $[\mathbf{K}]_{us}$ on ϕ_{us} and $[\mathbf{K}]_{ds}$ on ϕ_{ds} . Recall that n_{cn} is the number density of cylinders in the canopy layer, while$\langle\kappa_{h,cn}\rangle$ and $\langle\kappa_{v,cn}\rangle$ are per-cylinder extinction cross-sections for the field, to be distinguished from the extinction coefficient matrices $[\mathbf{K}]$ for the power or intensity. $\langle\kappa_{v,cn}(\theta_{us})\rangle$ and $\langle\kappa_{h,cn}(\theta_{us})\rangle$ are found from the optical theorem (equation 1.23); the angular brackets denote averaging over the orientation and size distribution of the scatterers (i.e. cylinders).

$$\langle\mathbf{M}_{us;di}\rangle = \frac{n_{cn}}{k^2}\Delta\Omega_{us} \quad \times$$

$$\left\langle\begin{bmatrix} S_{vv}S_{vv}^* & S_{vh}S_{vh}^* & \mathrm{Re}(S_{vv}S_{vh}^*) & -\mathrm{Im}(S_{vv}S_{vh}^*) \\ S_{hv}S_{hv}^* & S_{hh}S_{hh}^* & \mathrm{Re}(S_{hv}S_{hh}^*) & -\mathrm{Im}(S_{hv}S_{hh}^*) \\ 2\mathrm{Re}(S_{vv}S_{hv}^*) & 2\mathrm{Re}(S_{vh}S_{hh}^*) & \mathrm{Re}(S_{vv}S_{hh}^* + S_{vh}S_{hv}^*) & -\mathrm{Im}(S_{vv}S_{hh}^* - S_{vh}S_{hv}^*) \\ 2\mathrm{Im}(S_{vv}S_{hv}^*) & 2\mathrm{Im}(S_{vh}S_{hh}^*) & \mathrm{Im}(S_{vv}S_{hh}^* + S_{vh}S_{hv}^*) & \mathrm{Re}(S_{vv}S_{hh}^* - S_{vh}S_{hv}^*) \end{bmatrix}\right\rangle \quad \text{(B. 7)}$$

is the average Stokes matrix density (entries listed in [40]) for incident direction (θ_{di}, ϕ_{di}) to scattered direction (θ_{us}, ϕ_{us}), and $\langle \mathbf{M}_{ds;ui} \rangle$ is for incident direction (θ_{ui}, ϕ_{ui}) to scattered direction (θ_{ds}, ϕ_{ds}), and likewise for $\langle \mathbf{M}_{ds;di} \rangle$ and $\langle \mathbf{M}_{us;ui} \rangle$. $k = 2\pi/\lambda$ is the wavenumber. The angular brackets for the Stokes matrix denote averaging over orientation of the cylinder orientation and size distribution, normalized to one cylinder. To clarify, we thus need to compute the dimensionless bistatic scattering matrix (Appendix A), in FSA coordinates,

$$\begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}$$

between incident direction (θ_{di}, ϕ_{di}) and scattered direction (θ_{us}, ϕ_{us}) for a single cylinder, and average the second order quantities $S_{vv}S_{vv}^*$, $S_{vh}S_{vh}^*$, etc. over the cylinder size and orientation distribution, to get the matrix elements $\langle S_{vv}S_{vv}^* \rangle$, $\langle S_{vh}S_{vh}^* \rangle$, etc. (Care must be taken that in the FSA coordinate definition, θ_i is the supplement of what one might expect.) The following distribution (Table B.1) of cylinder sizes was used:

Radius [cm]	0.10	0.14	0.20	0.29	0.42	0.59	0.85	1.21	1.72	2.46	3.51
Length [cm]	21.5	27.3	34.6	43.9	55.6	70.6	89.4	113.4	143.7	182.2	231.0
number dist. (%)	50.9	25.0	12.3	6.02	2.96	1.45	0.712	0.349	0.172	0.084	0.041
volume dist. (%)	2.13	2.70	3.42	4.33	5.49	6.96	8.83	11.19	14.18	17.98	22.80

Table B.1. Distribution of cylinder sizes for radiative transfer computation.

The cylinder lengths were chosen to follow the $L(r) \propto r^{2/3}$ relationship (equation 2.44). The cylinder number distribution was chosen to satisfy the $n_{cn}p(r)\,\Delta \ln r \propto r^{-2}$ relationship discussed in equation (2.39) of Section 2.7 (since log-uniform bins are used for the radius distribution). n_{cn} was varied to give the desired optical thickness τ_{cn}. The radius distribution was chosen for approximate consistency with the 1mm-3cm distribution discussed in Section 2.7. Strictly speaking, a distribution up to slightly larger radii might have been more appropriate. The reason is because while 1mm-3cm covered the main

resonance region with the strongest extinction and single-scattering backscatter per unit volume, for multiple-scattering, the single-scattering albedo (which considers the whole 4π sphere instead of just the backscatter direction) is small towards the Rayleigh regime, but continues to be significantly large at radii larger than the resonance size. However, it becomes unclear how far beyond 3cm radius the number density distribution will be valid; e.g. in the boreal forests, 3cm radius already approaches the typical maximum size of tree branches.

The orientation distribution was chosen to be uniform over the 4π sphere. (The multiple-scattering correction factors would turn out to be quite similar if we had chosen a cosine-squared distribution instead, so for simplicity we used a uniform distribution to compute the correction factors, and apply them even for non-uniform orientation distributions.) Note also that the number density n_{cn} and total layer height $(Z_2 - Z_1)$ always occur together so only their product $n_{cn}\,(Z_2 - Z_1)$ (number of cylinders per unit ground area) matters in this computation. Also for our case of a uniformly random cylinder orientation distribution, by symmetry $\langle \kappa_{h,cn}(\theta) \rangle = \langle \kappa_{v,cn}(\theta) \rangle = \langle \kappa_{cn} \rangle$ is independent of θ.

Directly computing and averaging the bistatic scattering matrices for each $\langle \mathbf{M}_{us;di} \rangle$ is computationally intensive. For randomly oriented cylinders, the average Stokes matrix should be a function only of the angle in between the incident and scattered directions, up to a coordinate transformation. This symmetry can be used to speed up computations. Recall that in FSA coordinates,

$$\hat{\mathbf{k}}_i = \sin\theta_i \cos\phi_i\, \hat{\mathbf{x}} + \sin\theta_i \sin\phi_i\, \hat{\mathbf{y}} - \cos\theta_i\, \hat{\mathbf{z}}$$

$$\hat{\mathbf{k}}_s = \sin\theta_s \cos\phi_s\, \hat{\mathbf{x}} + \sin\theta_s \sin\phi_s\, \hat{\mathbf{y}} + \cos\theta_s\, \hat{\mathbf{z}}$$

$$\hat{\mathbf{h}}_i = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_i}{\left|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_i\right|}\,, \qquad \hat{\mathbf{v}}_i = \hat{\mathbf{h}}_i \times \hat{\mathbf{k}}_i\,, \qquad \hat{\mathbf{h}}_s = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_s}{\left|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_s\right|}\,, \qquad \hat{\mathbf{v}}_s = \hat{\mathbf{h}}_s \times \hat{\mathbf{k}}_i$$

Let the direction vector perpendicular to the scattering plane be

$$\widehat{\boldsymbol{w}} = \frac{\hat{\mathbf{k}}_i \times \hat{\mathbf{k}}_s}{\left|\hat{\mathbf{k}}_i \times \hat{\mathbf{k}}_s\right|} \tag{B.8}$$

and define

$$\hat{\mathbf{h}}_{iw} = \hat{\boldsymbol{w}} \times \hat{\mathbf{k}}_i \,, \qquad \hat{\mathbf{v}}_{iw} = \hat{\mathbf{h}}_{iw} \times \hat{\mathbf{k}}_i = -\,\hat{\boldsymbol{w}}, \tag{B.9a}$$

$$\hat{\mathbf{h}}_{sw} = \hat{\boldsymbol{w}} \times \hat{\mathbf{k}}_s \,, \qquad \hat{\mathbf{v}}_{sw} = \hat{\mathbf{h}}_{sw} \times \hat{\mathbf{k}}_s = -\,\hat{\boldsymbol{w}}. \tag{B.9b}$$

Let

$$[\mathbf{S}]_w = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_w \tag{B.10}$$

be the scattering matrix in the rotated frame where $\hat{\boldsymbol{w}}$ is vertical. The incident and scattered directions lie in the horizontal plane in this frame, and the 2x2 scattering matrix $[\mathbf{S}]_w = [\mathbf{S}(\Delta\phi)]_w$ depends only on the angle between them,

$$\Delta\phi = \operatorname{acos}\left(\hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_s\right). \tag{B.11}$$

For the cases of forward scattering $\Delta\phi = 0$ or back scattering $\Delta\phi = \pi$ (or approximate cases), pick $\hat{\boldsymbol{w}} = -\hat{\mathbf{v}}_i$. The 2x2 scattering matrix in the original global frame is then given by the coordinate transformation

$$[\mathbf{S}(\theta_i, \phi_i, \theta_s, \phi_s)] = [\mathbf{U}]_{sw}[\mathbf{S}(\Delta\phi)]_w[\mathbf{U}]_{wi} \tag{B.12}$$

where the rotation matrices

$$[\mathbf{U}]_{sw} = \begin{bmatrix} \hat{\mathbf{h}}_s \cdot \hat{\mathbf{h}}_{sw} & \hat{\mathbf{h}}_s \cdot \hat{\mathbf{v}}_{sw} \\ \hat{\mathbf{v}}_s \cdot \hat{\mathbf{h}}_{sw} & \hat{\mathbf{v}}_s \cdot \hat{\mathbf{v}}_{sw} \end{bmatrix}, \qquad [\mathbf{U}]_{wi} = \begin{bmatrix} \hat{\mathbf{h}}_{iw} \cdot \hat{\mathbf{h}}_i & \hat{\mathbf{h}}_{iw} \cdot \hat{\mathbf{v}}_i \\ \hat{\mathbf{v}}_{iw} \cdot \hat{\mathbf{h}}_i & \hat{\mathbf{v}}_{iw} \cdot \hat{\mathbf{v}}_i \end{bmatrix}. \tag{B.13}$$

Writing out the 2x2 scattering matrix as a vector, equation (B.13) becomes

$$\begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix} = [\mathbf{U}]_{sw} \otimes [\mathbf{U}]_{wi}^T \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix}_w \tag{B.14}$$

where

$$[\mathbf{U}]_{sw}\otimes[\mathbf{U}]_{wi}^{T} = \begin{bmatrix} \hat{\mathbf{h}}_s\cdot\hat{\mathbf{h}}_{sw}\begin{bmatrix}\hat{\mathbf{h}}_{iw}\cdot\hat{\mathbf{h}}_i & \hat{\mathbf{v}}_{iw}\cdot\hat{\mathbf{h}}_i\\ \hat{\mathbf{h}}_{iw}\cdot\hat{\mathbf{v}}_i & \hat{\mathbf{v}}_{iw}\cdot\hat{\mathbf{v}}_i\end{bmatrix} & \hat{\mathbf{h}}_s\cdot\hat{\mathbf{v}}_{sw}\begin{bmatrix}\hat{\mathbf{h}}_{iw}\cdot\hat{\mathbf{h}}_i & \hat{\mathbf{v}}_{iw}\cdot\hat{\mathbf{h}}_i\\ \hat{\mathbf{h}}_{iw}\cdot\hat{\mathbf{v}}_i & \hat{\mathbf{v}}_{iw}\cdot\hat{\mathbf{v}}_i\end{bmatrix}\\ \hat{\mathbf{v}}_s\cdot\hat{\mathbf{h}}_{sw}\begin{bmatrix}\hat{\mathbf{h}}_{iw}\cdot\hat{\mathbf{h}}_i & \hat{\mathbf{v}}_{iw}\cdot\hat{\mathbf{h}}_i\\ \hat{\mathbf{h}}_{iw}\cdot\hat{\mathbf{v}}_i & \hat{\mathbf{v}}_{iw}\cdot\hat{\mathbf{v}}_i\end{bmatrix} & \hat{\mathbf{v}}_s\cdot\hat{\mathbf{v}}_{sw}\begin{bmatrix}\hat{\mathbf{h}}_{iw}\cdot\hat{\mathbf{h}}_i & \hat{\mathbf{v}}_{iw}\cdot\hat{\mathbf{h}}_i\\ \hat{\mathbf{h}}_{iw}\cdot\hat{\mathbf{v}}_i & \hat{\mathbf{v}}_{iw}\cdot\hat{\mathbf{v}}_i\end{bmatrix}\end{bmatrix}. \quad \text{(B. 15)}$$

The 4x4 scatterer covariance matrices in the global frame and the rotated frame are thus related by

$$[\mathbf{C}(\theta_i,\phi_i,\theta_s,\phi_s)] = ([\mathbf{U}]_{sw}\otimes[\mathbf{U}]_{wi}^{T})[\mathbf{C}(\Delta\phi)]_w([\mathbf{U}]_{sw}\otimes[\mathbf{U}]_{wi}^{T})^{\dagger} \quad \text{(B. 16)}$$

where

$$[\mathbf{C}] = \begin{bmatrix} S_{hh}S_{hh}^{*} & S_{hh}S_{hv}^{*} & S_{hh}S_{vh}^{*} & S_{hh}S_{vv}^{*}\\ S_{hv}S_{hh}^{*} & S_{hv}S_{hv}^{*} & S_{hv}S_{vh}^{*} & S_{hv}S_{vv}^{*}\\ S_{vh}S_{hh}^{*} & S_{vh}S_{hv}^{*} & S_{vh}S_{vh}^{*} & S_{vh}S_{vv}^{*}\\ S_{vv}S_{hh}^{*} & S_{vv}S_{hv}^{*} & S_{vv}S_{vh}^{*} & S_{vv}S_{vv}^{*}\end{bmatrix}$$

is the 4x4 scatterer covariance matrix and the dagger symbol † denotes matrix transpose conjugation (to be distinguished from the superscript ^{T} denoting matrix transpose). In this way, we can precompute $\langle[\mathbf{C}(\Delta\phi)]_w\rangle$ over a grid of $\Delta\phi$ values (a 2° grid spacing was used), and obtain $\langle[\mathbf{C}(\theta_i,\phi_i,\theta_s,\phi_s)]\rangle$ for any incident and scattered direction pair by applying the transformation. The angular brackets denote averaging over the cylinder orientation and size distribution. Finally, the elements of the average Stokes matrix

$$\left\langle\begin{bmatrix} S_{vv}S_{vv}^{*} & S_{vh}S_{vh}^{*} & \mathrm{Re}(S_{vv}S_{vh}^{*}) & -\mathrm{Im}(S_{vv}S_{vh}^{*})\\ S_{hv}S_{hv}^{*} & S_{hh}S_{hh}^{*} & \mathrm{Re}(S_{hv}S_{hh}^{*}) & -\mathrm{Im}(S_{hv}S_{hh}^{*})\\ 2\mathrm{Re}(S_{vv}S_{hv}^{*}) & 2\mathrm{Re}(S_{vh}S_{hh}^{*}) & \mathrm{Re}(S_{vv}S_{hh}^{*}+S_{vh}S_{hv}^{*}) & -\mathrm{Im}(S_{vv}S_{hh}^{*}-S_{vh}S_{hv}^{*})\\ 2\mathrm{Im}(S_{vv}S_{hv}^{*}) & 2\mathrm{Im}(S_{vh}S_{hh}^{*}) & \mathrm{Im}(S_{vv}S_{hh}^{*}+S_{vh}S_{hv}^{*}) & \mathrm{Re}(S_{vv}S_{hh}^{*}-S_{vh}S_{hv}^{*})\end{bmatrix}\right\rangle$$

can be obtained from the elements of the average covariance matrix.

Having computed the Stokes matrices $\langle\mathbf{M}_{us;di}\rangle$, $\langle\mathbf{M}_{ds;ui}\rangle$, etc., returning to the radiative transfer equation, the differential equations (B.5a) and (B.5b) can be solved numerically by the (backward) Euler method:

$$\mathbf{G}_{uj}^{(l)} = \mathbf{G}_{uj}^{(l-1)} + \Delta z \frac{d}{dz_+} \mathbf{G}_{uj}^{(l)} \quad , \qquad \forall l = 2, \dots, l_{max}\,, j = 1, \dots, J \tag{B.17a}$$

$$\mathbf{G}_{dj}^{(l)} = \mathbf{G}_{dj}^{(l+1)} + \Delta z \frac{d}{dz_-} \mathbf{G}_{dj}^{(l)} \quad , \qquad \forall l = 1, \dots, l_{max} - 1\,, j = 1, \dots, J. \tag{B.17b}$$

With appropriate boundary conditions on the top and bottom sublayers, these equations can be collected into a linear system whose solution would be our desired solution to the radiative transfer equation. In computations, $l_{max} \approx 100\tau_{cn}(\theta_{i0}) \cos\theta_{i0}$ was used (with $l_{max} = 20$ if $\tau_{cn}(\theta_{i0}) \cos\theta_{i0} < 0.2$), such that each sublayer is indeed thin.

For our radar scattering calculations, there are two additional special directions to be taken care of. Firstly, the incident radar beam is treated as a plane wave, i.e. its beamwidth is a delta function in angular space, so it has to be taken into account separately. The subscript $i0$ shall be used to denote quantities associated with this incident beam, e.g. the incident direction is (θ_{i0}, ϕ_{i0}) , with ϕ_{i0} set to 0 for convenience. (Again, care must be taken that in the FSA coordinate definition, θ_i is the supplement of what one might expect.) Let $\mathbf{G}_{i0}^{(l)}$ be a 4x1 vector of Stokes parameters to be interpreted as the power (in watts), in this direction, through a unit horizontal plane area (1m^2) at the l -th sublayer.

Secondly, the final output direction of interest (i.e. the backscatter direction) may not be coincident with the centre of one the discretized directions, so it is best to also handle it separately. The subscript $s0$ shall be used to denote terms associated with final scattering direction of interest for evaluation, i.e. the direction (θ_{s0}, ϕ_{s0}). For the particular case of backscatter, $(\theta_{s0}, \phi_{s0}) = (\theta_{i0}, \pi - \phi_{i0})$. Let $\mathbf{G}_{s0}^{(l)}$ be a 4x1 vector of Stokes parameters to be interpreted as the power (in watts) through a unit horizontal plane area (1m^2) through a unit solid angle $\Delta\Omega_{s0} = 1\ \text{steradian}$ centred at the direction (θ_{s0}, ϕ_{s0}), at the l -th sublayer.

Thus the linear system we actually solve is of the form

$$\mathbf{x} = [\mathbf{B}]\mathbf{x} + \mathbf{b} \tag{B.18}$$

$$\mathbf{x}=\begin{bmatrix}\mathbf{x}^{(1)}\\ \vdots\\ \mathbf{x}^{(l)}\\ \vdots\\ \mathbf{x}^{(l_{max})}\end{bmatrix},[\mathbf{B}]=\begin{bmatrix}\mathbf{B}^{(1,1)} & \cdots & \mathbf{B}^{(1,l_{max})}\\ \vdots & \ddots & \vdots\\ \mathbf{B}^{(l_{max},1)} & \cdots & \mathbf{B}^{(l_{max},l_{max})}\end{bmatrix},\mathbf{b}=\begin{bmatrix}\mathbf{b}^{(1)}\\ \vdots\\ \mathbf{b}^{(l)}\\ \vdots\\ \mathbf{b}^{(l_{max})}\end{bmatrix}$$

where $\mathbf{x}$ is a concatenation of l_{max} $\mathbf{x}^{(l)}$ vectors, each of which is $8J+8$ long:

$$\mathbf{x}^{(l)}=\begin{bmatrix}\mathbf{G}_{s0}^{(l)}\\ \mathbf{G}_{u1}^{(l)}\\ \vdots\\ \mathbf{G}_{uJ}^{(l)}\\ \mathbf{G}_{d1}^{(l)}\\ \vdots\\ \mathbf{G}_{dJ}^{(l)}\\ \mathbf{G}_{i0}^{(l)}\end{bmatrix} \tag{B.19}$$

and $[\mathbf{B}]$ has $l_{max}\times l_{max}$ blocks, each block a $(8J+8)\times(8J+8)$ matrix. Using $[\mathbf{0}]_{(8J+8)\times(8J+8)}$ to denote the $(8J+8)\times(8J+8)$ zero matrix, $[\mathbf{0}]_{(4J+4)\times(4J+4)}$ to denote $(4J+4)\times(4J+4)$ zero matrix, $[\mathbf{I}]_{(4J+4)\times(4J+4)}$ to denote the $(4J+4)\times(4J+4)$ identity matrix, the (l,m)-th block of $[\mathbf{B}]$ is

$$\mathbf{B}^{(l,m)}=\Delta z[\mathbf{A}]\,,\qquad \forall l=m,\qquad l=1,\dots l_{max} \tag{B.20}$$

$$\mathbf{B}^{(l,m)}=\begin{bmatrix}[\mathbf{I}]_{(4J+4)\times(4J+4)} & [\mathbf{0}]_{(4J+4)\times(4J+4)}\\ [\mathbf{0}]_{(4J+4)\times(4J+4)} & [\mathbf{0}]_{(4J+4)\times(4J+4)}\end{bmatrix},\qquad \forall l=m+1,\qquad l=2,\dots l_{max}$$

$$\mathbf{B}^{(l,m)}=\begin{bmatrix}[\mathbf{0}]_{(4J+4)\times(4J+4)} & [\mathbf{0}]_{(4J+4)\times(4J+4)}\\ [\mathbf{0}]_{(4J+4)\times(4J+4)} & [\mathbf{I}]_{(4J+4)\times(4J+4)}\end{bmatrix},\qquad \forall l=m-1,\qquad m=2,\dots l_{max}$$

$$\mathbf{B}^{(l,m)}=[\mathbf{0}]_{(8J+8)\times(8J+8)}\,,\qquad otherwise$$

where $[\mathbf{A}]$ is the following $(8J+8)\times(8J+8)$ matrix:

$$[\mathbf{A}] = \begin{bmatrix} [\mathbf{A}_{s0;s0}] & [\mathbf{A}_{s0;u}] & [\mathbf{A}_{s0;d}] & [\mathbf{A}_{s0;i0}] \\ [\mathbf{A}_{u;s0}] & [\mathbf{A}_{u;u}] & [\mathbf{A}_{u;d}] & [\mathbf{A}_{u;i0}] \\ [\mathbf{A}_{d;s0}] & [\mathbf{A}_{d;u}] & [\mathbf{A}_{d;d}] & [\mathbf{A}_{d;i0}] \\ [\mathbf{A}_{i0;s0}] & [\mathbf{A}_{i0;u}] & [\mathbf{A}_{i0;d}] & [\mathbf{A}_{i0;i0}] \end{bmatrix}. \tag{B.21}$$

The matrices within $[\mathbf{A}]$ are defined as follows:

$$[\mathbf{A}_{s0;s0}] = \frac{1}{|\cos\theta_{s0}|}\left(\langle \mathbf{M}_{s0;s0}\rangle - [\mathbf{K}]_{s0}\right), \qquad \text{a } 4\times 4 \text{ matrix}$$

$$[\mathbf{A}_{s0;i0}] = \frac{1}{|\cos\theta_{i0}|}\langle \mathbf{M}_{s0;i0}\rangle, \qquad \text{a } 4\times 4 \text{ matrix}$$

$$[\mathbf{A}_{i0;i0}] = -\frac{1}{|\cos\theta_{i0}|}[\mathbf{K}]_{i0}, \qquad \text{a } 4\times 4 \text{ matrix}$$

$$[\mathbf{A}_{i0;s0}] = [\mathbf{0}]_{4\times 4}, \qquad \text{a } 4\times 4 \text{ zero matrix}$$

$$[\mathbf{A}_{s0;u}] = \left[[\mathbf{A}_{s0;u1}][\mathbf{A}_{s0;u2}]\dots[\mathbf{A}_{s0;uJ}]\right], \qquad \text{a } 4\times 4J \text{ matrix which is a row concantenation of } J \text{ 4x4 submatrices, with } [\mathbf{A}_{s0;ui}] = \frac{1}{|\cos\theta_{ui}|}\langle \mathbf{M}_{s0;ui}\rangle \quad \forall i = 1,\dots,J$$

$$[\mathbf{A}_{s0,d}] = \left[[\mathbf{A}_{s0;d1}][\mathbf{A}_{s0;d2}]\dots[\mathbf{A}_{s0;dJ}]\right], \qquad \text{a } 4\times 4J \text{ matrix which is a row concantenation of } J \text{ } 4\times 4 \text{ submatrices, with } [\mathbf{A}_{s0;di}] = \frac{1}{|\cos\theta_{di}|}\langle \mathbf{M}_{s0;di}\rangle \quad \forall i = 1,\dots,J$$

$$[\mathbf{A}_{u;i0}] = \begin{bmatrix} [\mathbf{A}_{u1;i0}] \\ \vdots \\ [\mathbf{A}_{uJ;i0}] \end{bmatrix}, \qquad \text{a } 4\text{J}\times 4 \text{ matrix which is a column concantenation of } J \text{ } 4\times 4 \text{ submatrices, with } [\mathbf{A}_{us;i0}] = \frac{1}{|\cos\theta_{i0}|}\langle \mathbf{M}_{us;i0}\rangle \quad \forall s = 1,\dots,J$$

$$[\mathbf{A}_{d;i0}] = \begin{bmatrix} [\mathbf{A}_{d1;i0}] \\ \vdots \\ [\mathbf{A}_{dJ;i0}] \end{bmatrix}, \qquad \text{a } 4\text{J}\times 4 \text{ matrix which is a column concantenation of } J \text{ } 4\times 4 \text{ submatrices, with } [\mathbf{A}_{ds;i0}] = \frac{1}{|\cos\theta_{i0}|}\langle \mathbf{M}_{ds;i0}\rangle \quad \forall s = 1,\dots,J$$

$$[\mathbf{A}_{u;s0}] = [\mathbf{0}]_{4J\times 4}, \qquad \text{a } 4\text{J}\times 4 \text{ zero matrix}$$

$$[\mathbf{A}_{d;s0}] = [\mathbf{0}]_{4J\times 4}, \qquad \text{a } 4\text{J}\times 4 \text{ zero matrix}$$

$$\left[\mathbf{A}_{i0;u}\right] = [\mathbf{0}]_{4\times 4J}\,, \qquad \text{a } 4 \times 4\text{J zero matrix}$$

$$\left[\mathbf{A}_{i0;d}\right] = [\mathbf{0}]_{4\times 4J}\,, \qquad \text{a } 4 \times 4\text{J zero matrix}$$

$$\left[\mathbf{A}_{u;d}\right] = \begin{bmatrix} \left[\mathbf{A}_{u1;d1}\right] & \cdots & \left[\mathbf{A}_{u1;dJ}\right] \\ \vdots & \ddots & \vdots \\ \left[\mathbf{A}_{uJ;d1}\right] & \cdots & \left[\mathbf{A}_{uJ;dJ}\right] \end{bmatrix}, \text{comprising } J \times J \text{ blocks of } 4 \times 4 \text{ matrices, with}$$

$$\left[\mathbf{A}_{us;di}\right] = \frac{1}{|\cos\theta_{di}|}\langle \mathbf{M}_{us;di}\rangle \quad \forall s, i = 1, \dots, J$$

$$\left[\mathbf{A}_{d;u}\right] = \begin{bmatrix} \left[\mathbf{A}_{d1;u1}\right] & \cdots & \left[\mathbf{A}_{d1;uJ}\right] \\ \vdots & \ddots & \vdots \\ \left[\mathbf{A}_{dJ;u1}\right] & \cdots & \left[\mathbf{A}_{dJ;uJ}\right] \end{bmatrix}, \text{comprising } J \times J \text{ blocks of } 4 \times 4 \text{ matrices, with}$$

$$\left[\mathbf{A}_{ds;ui}\right] = \frac{1}{|\cos\theta_{ui}|}\langle \mathbf{M}_{ds;ui}\rangle \quad \forall s, i = 1, \dots, J$$

$$\left[\mathbf{A}_{u;u}\right] = \begin{bmatrix} \left[\mathbf{A}_{u1;u1}\right] & \cdots & \left[\mathbf{A}_{u1;uJ}\right] \\ \vdots & \ddots & \vdots \\ \left[\mathbf{A}_{uJ;u1}\right] & \cdots & \left[\mathbf{A}_{uJ;uJ}\right] \end{bmatrix}, \text{comprising } J \times J \text{ blocks of } 4 \times 4 \text{ matrices, with}$$

$$\left[\mathbf{A}_{us;ui}\right] = \frac{1}{|\cos\theta_{ui}|}\langle \mathbf{M}_{us;ui}\rangle \quad \forall s, i = 1, \dots, J\,, s \neq i$$

$$\left[\mathbf{A}_{us;ui}\right] = \frac{1}{|\cos\theta_{ui}|}\left(\langle \mathbf{M}_{us;ui}\rangle - [\mathbf{K}]_{us}\right) \quad \forall s, i = 1, \dots, J\,, s = i$$

$$\left[\mathbf{A}_{d;d}\right] = \begin{bmatrix} \left[\mathbf{A}_{d1;d1}\right] & \cdots & \left[\mathbf{A}_{d1;dJ}\right] \\ \vdots & \ddots & \vdots \\ \left[\mathbf{A}_{dJ;d1}\right] & \cdots & \left[\mathbf{A}_{dJ;dJ}\right] \end{bmatrix}, \text{comprising } J \times J \text{ blocks of } 4 \times 4 \text{ matrices, with}$$

$$\left[\mathbf{A}_{ds;di}\right] = \frac{1}{|\cos\theta_{di}|}\langle \mathbf{M}_{ds;di}\rangle \quad \forall s, i = 1, \dots, J\,, s \neq i$$

$$\left[\mathbf{A}_{ds;di}\right] = \frac{1}{|\cos\theta_{di}|}\left(\langle \mathbf{M}_{ds;di}\rangle - [\mathbf{K}]_{ds}\right) \quad \forall s, i = 1, \dots, J\,, s = i.$$

It remains to specify $\mathbf{b}$, the column concatenation of l_{max} $\mathbf{b}^{(l)}$ vectors, each of which is $8J + 8$ long. Most entries are zero:

$$\mathbf{b}^{(l)} = [\mathbf{0}]_{(8J+8)\times 1}\,, \qquad \text{a } (8\text{J} + 8) \times 1 \text{ zero vector}\,, \forall l = 2, \dots, l_{max} - 1 \qquad \text{(B. 22a)}$$

while $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(l_{max})}$ specify the boundary conditions at the bottom and top respectively. At the top, we neglect internal reflection, but provide a 1W/m^2 incident plane wave from direction (θ_{i0}, ϕ_{i0})

$$\mathbf{G}_{i0}^{(l_{max}+1)} = \begin{bmatrix} 0 \\ |\cos\theta_{i0}| \\ 0 \\ 0 \end{bmatrix} \text{ for h} - \text{polarization,}$$

$$\mathbf{G}_{i0}^{(l_{max}+1)} = \begin{bmatrix} |\cos\theta_{i0}| \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ for v} - \text{polarization}$$

so we set the boundary condition at the top to be

$$\mathbf{b}^{(l_{max})} = \begin{bmatrix} [\mathbf{0}]_{(8J+4)\times 1} \\ \mathbf{G}_{i0}^{(l_{max}+1)} \end{bmatrix} \qquad \text{(B22. b)}$$

and at the bottom boundary, set to be

$$\mathbf{b}^{(1)} = [\mathbf{0}]_{(8J+8)\times 1}\ ,\ \text{a } (8\text{J} + 8) \times 1 \text{ zero vector.} \qquad \text{(B22. c)}$$

This simplified boundary condition corresponds to no reflections off the ground nor incident radiation upwelling from the ground. This may seem unphysical, but remember that the purpose of this radiative transfer calculation is just to compute the multiple-scattering correction to the canopy backscattering term only. Reflections involving the ground had been taken into account as a separate term. If we chose to include reflections off the ground within the radiative transfer computation as is often done in the literature, some rearrangement of terms and entries of $[\mathbf{B}]$ would need to be modified correspondingly. Ground physical parameters would enter and we would need to solve the radiative transfer equation for each set of ground parameters, increasing the complexity and computational effort required to explore the full parameter space, and yet still not having taken into account coherent effects that would require further correction. The simplification adopted here requires only vegetation parameters for the radiative transfer calculation and isolates contributions from different scattering components to the total radar backscatter.

With the linear system fully specified, solving equation (B.18) for $\mathbf{x}$, and recalling that

$$\mathbf{x}^{(l_{max})} = \begin{bmatrix} \mathbf{G}_{s0}^{(l_{max})} \\ \mathbf{G}_{u1}^{(l_{max})} \\ \vdots \\ \mathbf{G}_{uJ}^{(l_{max})} \\ \mathbf{G}_{d1}^{(l_{max})} \\ \vdots \\ \mathbf{G}_{dJ}^{(l_{max})} \\ \mathbf{G}_{i0}^{(l_{max})} \end{bmatrix}$$

our desired radar backscatter is contained within the solution $\mathbf{G}_{s0}^{(l_{max})}$, which is a 4x1 vector of Stokes parameters to be interpreted as the power (in watts) through a unit horizontal plane area (1m^2) through a unit solid angle $\Delta\Omega_{s0} = 1$ steradian centred in the backscatter direction (θ_{s0}, ϕ_{s0}), that was due to a 1W/m^2 incident plane wave. From Section 1.3, the normalized radar cross-section is 4π of the scattered power per steradian per unit ground area per unit illumination intensity

$$\sigma_{vh}^0 = 4\pi \frac{1}{\mathrm{A}} \frac{dP_v^s}{d\Omega} \frac{1}{\mathcal{S}_h^i} \tag{B.23}$$

here written for the example of incident h-polarization and scattered v-polarization. Thus we have, for the case of incident h-polarization (setting $\mathbf{G}_{i0}^{(l_{max}+1)} = [0\ |\cos\theta_{i0}|\ 0\ 0]^T$), normalized backscatter radar cross-sections

$$\mathrm{VH}_{cn,rt} = 4\pi\ [1\ 0\ 0\ 0] \cdot\ \mathbf{G}_{s0}^{(l_{max})} \tag{B.24a}$$

$$\mathrm{HH}_{cn,rt} = 4\pi\ [0\ 1\ 0\ 0] \cdot\ \mathbf{G}_{s0}^{(l_{max})} \tag{B.24b}$$

and for the case of incident v-polarization (setting $\mathbf{G}_{i0}^{(l_{max}+1)} = [|\cos\theta_{i0}|\ 0\ 0\ 0]^T$), normalized backscatter radar cross-sections

$$\mathrm{VV}_{cn,rt} = 4\pi\ [1\ 0\ 0\ 0] \cdot\ \mathbf{G}_{s0}^{(l_{max})} \tag{B.24c}$$

$$\mathrm{HV}_{cn,rt} = 4\pi\ [0\ 1\ 0\ 0] \cdot\ \mathbf{G}_{s0}^{(l_{max})} \tag{B.24d}$$

We expect HV and VH to be the same and indeed the computed values are in close agreement (this is an important check because the approximation (A.1) that we used for bistatic scattering from cylinders did not necessarily respect reciprocity for all cases); thus the $\mathrm{HV}_{cn,rt}$ and $\mathrm{VH}_{cn,rt}$ results are averaged and treated together as a single cross-polarized case. $\mathrm{HH}_{cn,rt}$ and $\mathrm{VV}_{cn,rt}$ are almost the same, with a very slight excess in $\mathrm{VV}_{cn,rt}$ compared to $\mathrm{HH}_{cn,rt}$ at larger optical thickness. An explanation of this difference is provided by van de Hulst [142]: it is due to double-scatterings involving pairs of scatterers near the surface such that the direction vector between them is parallel to $\pm\hat{\mathbf{h}}_i$. These pathways provide the greater backscatter return for VV compared to HH; this difference is more pronounced at large incidence angles. For our case, it is typically no more than 0.2dB.

The results from this radiative transfer computation are compared in Figure B.1 to the results from the single-scattering calculation of equations (2.7a-c), at radar wavelength $= 0.24\mathrm{m}$, incidence angle $\theta_i = 38.49°$, and cylinder relative permittivity $\varepsilon_v = 35.94 + 11.09i$.

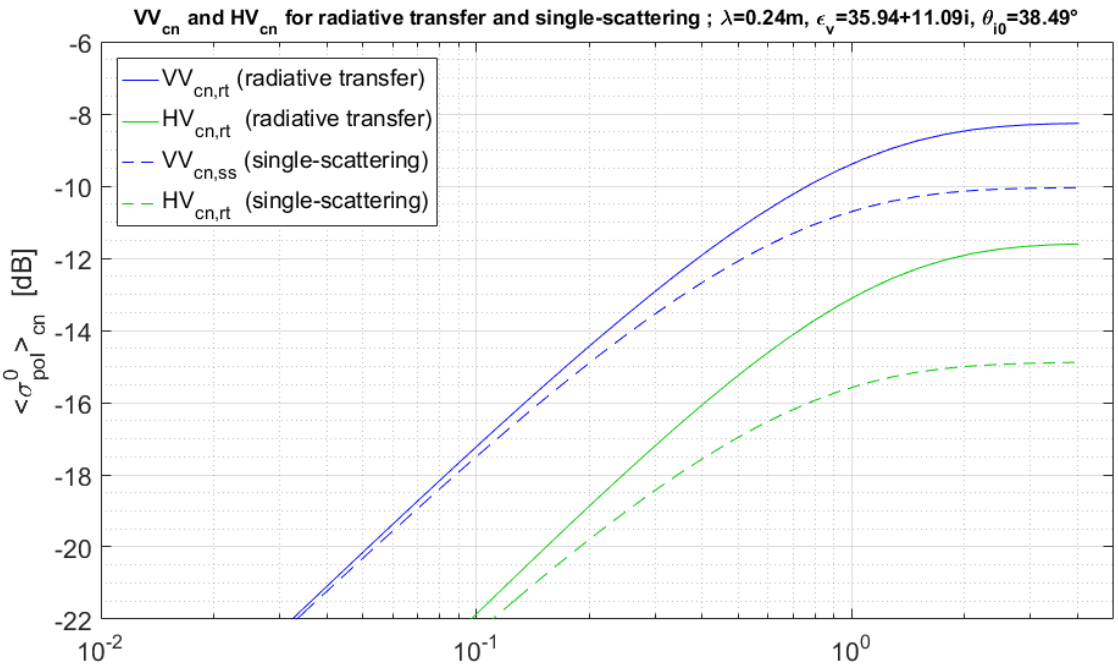

Figure B.2. Normalized backscatter radar cross-sections $\mathrm{VV}_{cn,rt}$ ($\approx \mathrm{HH}_{cn,rt}$), $\mathrm{VV}_{cn,ss}$ ($= \mathrm{HH}_{cn,ss}$), $\mathrm{HV}_{cn,rt}$ ($= \mathrm{VH}_{cn,rt}$), $\mathrm{HV}_{cn,ss}$ ($= \mathrm{VH}_{cn,ss}$), as a function of the optical thickness $\tau_{cn}(\theta_{i0}) = 2n_{cn}\langle\kappa_{cn}\rangle(Z_2 - Z_1)/\cos\theta_{i0}$, for the radiative transfer and single-scattering methods.

This radiative transfer solution however is not the final result. In the backscatter direction, there is an enhancement of co-polarized backscatter due to coherent effects [142, 143, 138]. The backscatter enhancement is essentially due to the constructive interference of reciprocal (for the co-polarized case) counter-propagating paths with exactly equal path length. This radiative transfer solution does not contain such a coherent effect. To account for this backscatter enhancement, the estimated additional co-polarized backscatter due to multiple-scattering (i.e. the difference between the radiative transfer solution and the single-scattering solution) is multiplied by 2. For the cross-polarized backscatter, however, the counter-propagating paths are not reciprocal. Thus the backscatter enhancement factor may not be 2 for cross-polarization; in fact, de-enhancement may even occur [144]. In view of this issue, we do not use the $\mathrm{HV}_{cn,rt}$ solution from radiative transfer, but instead independently estimate the ratio of the cross-polarized to co-polarized multiple-scattering contribution, so as to leverage on the greater confidence in the co-polarized quantities. A separate set of Monte-Carlo simulations (with the afore-mentioned distribution) was performed to estimate this cross-to-co-polarization ratio, denoted by ϱ, for double-scattering. This value should be fairly independent of incident angle due to the uniformly random orientation distribution of cylinders, however it would vary with cylinder relative permittivity ε_v. Higher-order scatterings were not computed owing to the significantly greater effort; from Liao et al. [139], we expect most of the multiple scattering contributions to be from double-scattering. Such an approach is not expected to be valid when the single-scattering albedo (ratio of total scattering to extinction, of individual particles) is very high, but in our case we estimate single-scattering albedos not to exceed 0.75.

The results from the Monte-Carlo simulations for several values of cylinder relative permittivity ε_v are tabulated in Table B.2.

Cylinder relative permittivity ε_v	$5.15 + 1.41i$	$17.10 + 5.83i$	$35.94 + 11.09i$	$62.75 + 18.22i$
Estimated double-scattering cross-to-co-pol ratio, $\varrho(\varepsilon_v)$	0.27	0.38	0.42	0.43

Table B.2. Monte-Carlo estimate of cross-pol to co-pol ratio for double-scattering.

The final estimated multiple-scattering correction factors are then:

$$\mathrm{HH}_{cn,ss}\mathcal{F}_{\mathrm{HH}} = \mathrm{HH}_{cn,ss} + 2\left(\mathrm{HH}_{cn,rt} - \mathrm{HH}_{cn,ss}\right)$$

$$\mathcal{F}_{\mathrm{HH}}(\tau_{cn}(\theta_{i0}), \theta_i, \varepsilon_v) = 1 + 2\left(\frac{\mathrm{HH}_{cn,rt}(\tau_{cn}, \theta_{i0}, \varepsilon_v)}{\mathrm{HH}_{cn,ss}(\tau_{cn}, \theta_{i0}, \varepsilon_v)} - 1\right) \tag{B.25}$$

$$\mathrm{VV}_{cn,ss}\mathcal{F}_{\mathrm{VV}} = \mathrm{VV}_{cn,ss} + 2\left(\mathrm{VV}_{cn,rt} - \mathrm{VV}_{cn,ss}\right)$$

$$\mathcal{F}_{\mathrm{VV}}(\tau_{cn}(\theta_{i0}), \theta_i, \varepsilon_v) = 1 + 2\left(\frac{\mathrm{VV}_{cn,rt}(\tau_{cn}, \theta_{i0}, \varepsilon_v)}{\mathrm{VV}_{cn,ss}(\tau_{cn}, \theta_{i0}, \varepsilon_v)} - 1\right) \tag{B.26}$$

$$\mathrm{HV}_{cn,ss}\mathcal{F}_{\mathrm{HV}} = \mathrm{HV}_{cn,ss} + 2\varrho\left(\frac{\mathrm{HH}_{cn,rt} + \mathrm{VV}_{cn,rt}}{2} - \mathrm{VV}_{cn,ss}\right)$$

$$\mathcal{F}_{\mathrm{HV}}(\tau_{cn}(\theta_{i0}), \theta_{i0}, \varepsilon_v) = 1 + \varrho(\varepsilon_v) \times \left(\frac{\mathrm{HH}_{cn,rt}(\tau_{cn}, \theta_{i0}, \varepsilon_v) + \mathrm{VV}_{cn,rt}(\tau_{cn}, \theta_{i0}, \varepsilon_v) - 2\mathrm{VV}_{cn,ss}(\tau_{cn}, \theta_{i0}, \varepsilon_v)}{\mathrm{HV}_{cn,ss}(\tau_{cn}, \theta_{i0}, \varepsilon_v)}\right). \tag{B.27}$$

These estimated multiple-scattering correction factors $\mathcal{F}_{\mathrm{HH}}(\tau_{cn}(\theta_{i0}), \theta_{i0}, \varepsilon_v)$, $\mathcal{F}_{\mathrm{VV}}(\tau_{cn}(\theta_{i0}), \theta_{i0}, \varepsilon_v)$ and $\mathcal{F}_{\mathrm{HV}}(\tau_{cn}(\theta_{i0}), \theta_{i0}, \varepsilon_v)$ are plotted in Figure 2.2, Figure 2.3, and Figure 2.4 as a function of optical thickness $\tau_{cn}(\theta_{i0}) = n_{cn}\left(\langle\kappa_{h,cn}(\theta_{i0})\rangle + \langle\kappa_{v,cn}(\theta_{i0})\rangle\right)(Z_2 - Z_1)/\cos\theta_{i0} = 2n_{cn}\langle\kappa_{cn}\rangle(Z_2 - Z_1)/\cos\theta_{i0}$ for several values of θ_{i0} and ε_v. Linear interpolation is used to obtain values of $\mathcal{F}_{\mathrm{HH}}$, $\mathcal{F}_{\mathrm{HV}}$, $\mathcal{F}_{\mathrm{VV}}$ at intermediate values of $\tau_{cn}(\theta_{i0})$, θ_{i0} and ε_v. Note that in Figure 2.2, Figure 2.3, and Figure 2.4, the notation for the radar incidence angle is reverted to θ_i instead of θ_{i0}.

Appendix C

MIRONOV DIELECTRIC MODEL OF MOIST SOILS

Dielectric models are necessary to relate the relative permittivity of moist soils to other physical properties including the water content. There are several popular models used for microwave remote sensing; see Wigneron et al. [145] for a review. Here we adopt the model by Mironov et al. [43] tailored for practical use at 1.4GHz, with validated formulae for the soil relative permittivity as a function of soil volumetric moisture, soil temperature, and clay fraction. In this model, the real and imaginary parts of the dielectric relative permittivity ε_g of a moist soil are given by

$$\text{Real}(\varepsilon_g) = n_s^2 - \kappa_s^2 , \qquad \text{Imag}(\varepsilon_g) = 2n_s\kappa_s$$

$$n_s = \begin{cases} n_d + (n_b - 1)W, & W \le W_t \\ n_d + (n_b - 1)W_t + (n_u - 1)(W - W_t), & W \ge W_t \end{cases}$$

$$\kappa_s = \begin{cases} \kappa_d + \kappa_b W, & W \le W_t \\ \kappa_d + \kappa_b W_t + (\kappa_u - 1)(W - W_t), & W \ge W_t \end{cases}$$

where, in this Appendix only, n and κ are the real and imaginary parts of the complex refractive index, not to be confused with notations elsewhere in this thesis. The subscripts s, d, b, u are for moist soil, dry soil, bound soil water, and free soil water, respectively. W is the volumetric moisture content of the soil in cm^3/cm^3, and W_t is the maximum bound water fraction (for a given type of soil). Expressions for these parameters at 1.4GHz were provided by Mironov et al. [43] as fitted polynomial functions of clay content C in percent (valid in the range 0-70), and temperature T in degrees Celsius (valid in the range 10-40):

$$W_t = 0.0286 + 0.00307C$$

$$n_d = 1.634 - 0.00539C + 2.75 \times 10^{-5} C^2$$

$$\kappa_d = 0.0395 - 4.038 \times 10^{-4} C$$

$$n_b = (8.86 + 0.00321T) + (-0.0644 + 7.96 \times 10^{-4}T)C + (2.97 \times 10^{-4} - 9.6 \times 10^{-6}T)C^2$$

$$\kappa_b = (0.738 - 0.000903T + 8.57 \times 10^{-5}T^2) + (-0.00215 + 1.47 \times 10^{-4}T)C + (7.36 \times 10^{-5} - 1.03 \times 10^{-6}T + 1.05 \times 10^{-8}T^2)C^2$$

$$n_u = (10.3 - 0.0173T) + (6.5 \times 10^{-4} + 8.82 \times 10^{-5}T)C + (-6.34 \times 10^{-6} - 6.32 \times 10^{-7}T)C^2$$

$$\kappa_u = (0.738 - 0.017T) + (1.78 \times 10^{-4}T^2) + (0.0161 + 7.25 \times 10^{-4}T)C + (-1.46 \times 10^{-4} - 6.03 \times 10^{-6}T - 7.87 \times 10^{-9}T^2)C^2$$

This model allows us to compute the soil relative permittivity ε_g given the soil volumetric moisture content W, temperature T and clay content C. To retrieve the soil moisture W, we also need the inverse relationship. Since the dependence on temperature is quite weak over most of the range of interest, for convenience and ease of inversion, we shall use only $T = 15°C$. We shall also neglect the imaginary part of ε_g as before (see Section 3.3). We then compute ε_g over a grid of values of W and C, and use this as a look-up table for the inverse relationship to obtain W from ε_g and C .