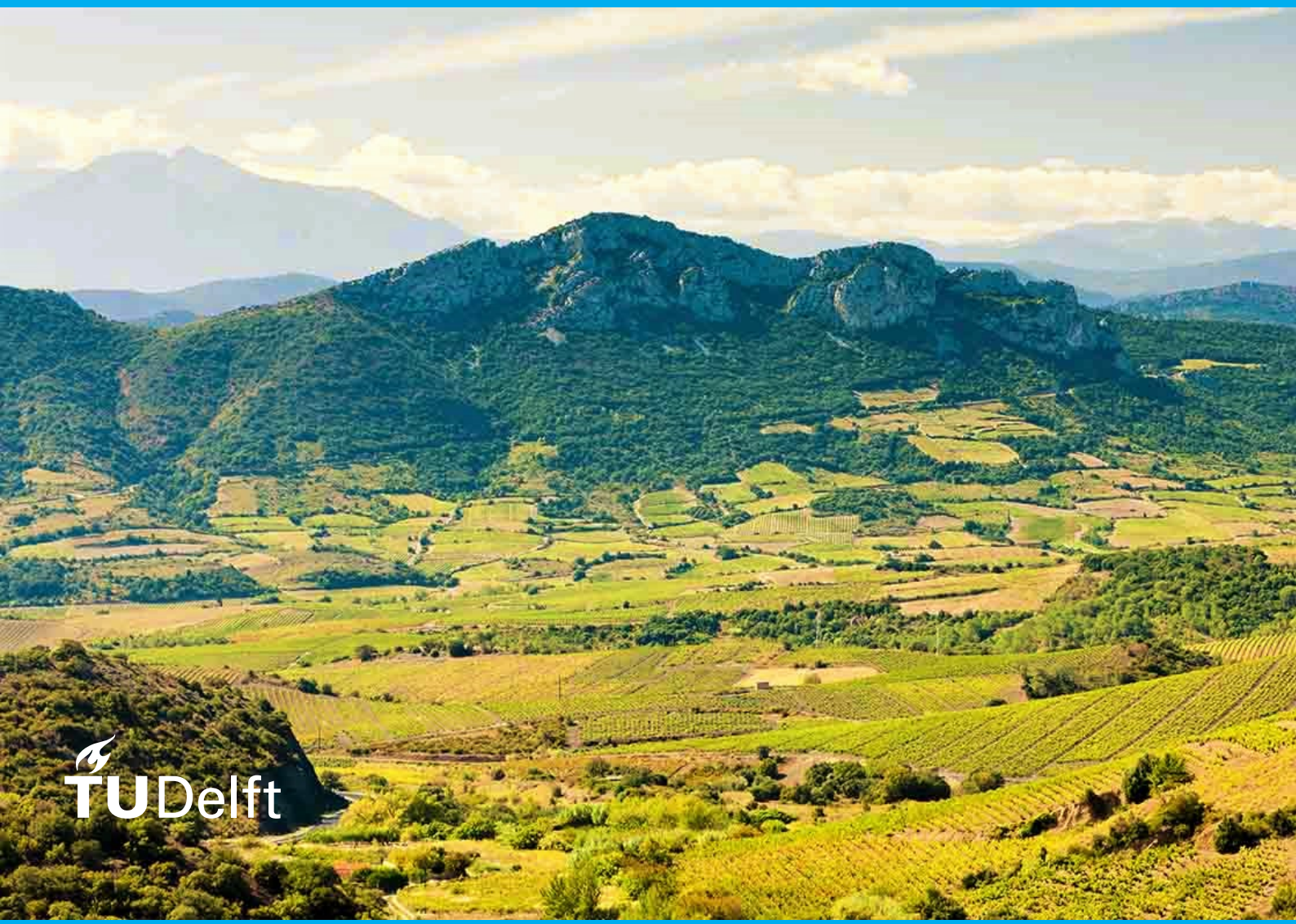
Sensitivity Assessment of Sentinel-1 SAR Closure Phase to Vegetation and Soil Moisture Dynamics

A Case Study for Regions in Southern France

Noah Hadler

MSc Thesis



TU Delft

Sensitivity Assessment of Sentinel-1 SAR Closure Phase to Vegetation and Soil Moisture Dynamics

A Case Study for Regions in Southern France

by

Noah Hadler

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday January 27, 2022 at 10:00 AM.

Student number:	4927656	
Project duration:	October 2020 – January 2022	
Thesis committee:	Prof. Dr. ir. S.C. Steele-Dunne,	TU Delft, chair, daily supervisor
	Dr. ir. P. López-Dekker,	TU Delft, supervisor
	Dr. S.L.M. Lhermitte,	TU Delft, external assessor
	Dr. ir. E. Abraham,	TU Delft, external assessor

TU Delft

Abstract

From three coherent SAR images it is possible to estimate three interferograms. Combined in a circular way, the sum of the three interferometric phases is called the closure phase which necessarily adds up to zero on a single pixel level. However, if the interferograms are spatially averaged, phase consistency is not guaranteed. In most of the interferometric studies, those mismatches were assumed to be caused by decorrelation noise alone, and were either not considered or deemed negligible, eluding further investigations of its origin. However, recent publications have confirmed that inconsistent phase closures are systematic and not the exception, pointing to an underlying geophysical cause. Comparisons of the spatial signatures of phase closures with land cover maps suggest a spatial and temporal correlation that is related to the characteristics of different land cover types. Since interferometric measurements are sensitive to variations of the dielectric constant, those similarities have been attributed to dynamics in vegetation and soil moisture. A closure phase significance test developed at the Geoscience and Remote Sensing department at TU Delft aimed to increase the signal-to-noise ratio of this geophysical signal component by providing a significance ratio for phase closures. However, the sensitivity of (significant) phase closures to dynamics in vegetation and soil over different land cover types has not been assessed yet. Here we show that with enough averaging of the interferometric phase, the spatial and temporal characteristics of closure phase can be used to distinguish between different land cover types. We found that the degree of spatial averaging has a significant impact on both the phase closure values and its spatial and temporal consistency. The magnitudes of significant phase closures generally increased over low-vegetated land covers, suggesting that closure phases are most sensitive to soil moisture dynamics, whereas vegetation cover was associated with decreasing phase closure magnitudes and spatial inconsistency. Besides spatial averaging, significant differences were observed between closure phases from different polarizations. Furthermore, we found that amplitude backscatter and closure phase are spatially and temporally correlated, pointing to similar influencing mechanisms. Our results demonstrate the importance of applying a closure phase significance test and describe the effect of spatial averaging on the characteristics of phase closures with respect to different land cover types. We anticipate this study to provide useful steps towards using the closure phase for soil and vegetation monitoring in the future. For example, the findings could be used to further exploit potential synergies with amplitude backscatter for soil moisture retrieval from closure phase or develop more sophisticated methods for land cover mapping using InSAR. If not used for applications linked to land cover, vegetation or soil, being able to better predict the effect of those parameters on the interferometric phase and coherence, eventually enables to separate their contribution from other signals, such as deformation estimates. Additional research is needed to relate significant phase closures to moisture changes in vegetation.

On the cover: photo of a landscape in southern France

Preface

The present master's thesis marks the end of my time as an university student. (Still can't believe it.) I would like to thank several people. Without their help and support this work would not have been possible.

First and foremost, I would like to thank Susan and Paco for guiding me through this process and for always providing me with valuable feedback and good discussions in our numerous Thursday meetings. In particular with the rather open topic in the beginning, and everyone being isolated and working remotely, it was reassuring to know those weekly meetings exist and that there is a time to share my progress and exchange ideas. I really appreciate all the time and effort you invested! Thank you Stef for your valuable feedback and ideas during my mid-term review. Your input was crucial and inspired me to rethink one of my research objectives. Thank you Edo for being part of my thesis committee and for taking the time to read this thesis for my defense. I would also like to thank Yan for all the discussions that we had about closure phases but also for the many random conversations over Slack in the past year. It was always fun to chat! Thank you to Gert, Malte and Tina for your support with *RIPPL* and *TPT*, the processing setup on the HPC cluster, and the assistance with the spatial averaging approach applied in this thesis.

Looking back at seven years of studying in Aachen, Baton Rouge and Delft, I am very happy that I have chosen an academic path that allowed me to explore the Earth with all its processes in so many different ways and with excellent teachers along the way. The time in Aachen and my study abroad at LSU have sparked my interest in satellite Earth observation which eventually motivated me to apply for *Geoscience & Remote Sensing* in Delft. It has been a challenging and formative adventure with lots of ups and downs, but always accompanied by amazing people that I got to meet and can now call friends. I am very grateful for everything and aware that not everyone is provided with such great opportunities and chances.

A very big thank you also to all my friends who kept me sane and provided me with all the support I could ask for during all of this. You are amazing! You probably not going to read that ever so I hope I am doing a good enough job of communicating that frequently.

And last, but certainly not least I would like to thank my family for always supporting me, encouraging me to make my own decisions, and for letting me choose my own path early on. You always trusted me along my journey and I cannot thank you enough for that. Ich habe euch lieb!

Noah Hadler
Mönchengladbach, January 2022

Contents

List of Figures

List of Tables

1

Introduction

1.1. Vegetation and Soil Monitoring from Space

Under changing climate conditions and population growth, we are facing all kinds of challenges, many of which involve a more intelligent use of water resources and accurate predictions of environmental changes. Global warming, and the resulting increase in the recurrence of natural hazards associated with aridity, demand constant monitoring of landscape values such as soil moisture (SM) and vegetation water content (VWC). This is essential to counter water and food scarcity in the future, and to mitigate potential wildfire hazards (Karimi et al., 2013; Yebra et al., 2013). Earth observation satellites provide the tools for large-scale measurements of soil and vegetation properties which are important references for landscape monitoring, agricultural applications and drought forecasting (Nikaein et al., 2021; Steele-Dunne et al., 2019; Vreugdenhil et al., 2018).

Nowadays, there are several different satellite-based data sources relevant for assessing the moisture content in soil and vegetation including observations from radar remote sensing (Bousbih et al., 2017; Entekhabi et al., 2010; Wagner et al., 2013). The ability of radar to penetrate vegetation canopy up to a certain depth, its sensitivity to vegetation structure and soil moisture changes (in the top few centimeters) have made it a widely used technique for vegetation and soil monitoring. Unlike optical remote sensing, radar can operate independent from solar illumination of the Earth surface and penetrates through clouds and haze. On average, $\sim$ 55% of the Earth's land surface is covered by clouds (King et al., 2013) which makes optical instruments not a good candidate for monitoring dynamic targets such as vegetation and soil. For decision making in applications such as agricultural monitoring, image acquisitions of a region every few days are necessary, e.g. to monitor water stress in crops (Steele-Dunne et al., 2017). Advising on irrigation over an agricultural field with the right amount of water at the right time can therefore not rely on constant cloud-free conditions.

Most of the studies in the past have demonstrated vegetation and soil monitoring from radar systems including JAXAs Advanced Microwave Scanning Radiometer 2 (AMSR2), EUMETSATs Metop Advanced SCATterometers (ASCAT), ESAs Soil Moisture Ocean Salinity (SMOS) and NASAs Soil Moisture Active Passive (SMAP) (Bartalis et al., 2007; Cho et al., 2015; Entekhabi et al., 2010; Kerr et al., 2001). Although the revisit time of those microwave sensors is as high as 1-2 days, the spatial resolution is relatively low, often on the order of several kilometers. Within the scope of the Copernicus

program, the Sentinel-1A and B satellites provide radar data from C-Band Synthetic Aperture Radar (SAR) at a temporal resolution of 1.5-4 days over Europe and an unprecedented spatial resolution of 20 m. Launched in 2014 and 2016 respectively, Sentinel-1 has since been used to measure dynamics in vegetation and soil moisture (El Hajj et al., 2017; Khabbazan et al., 2019; Torbick et al., 2017).

Yet, most approaches related to vegetation and soil monitoring have focused on using only the amplitude information in SAR imagery. Radar backscatter is influenced by several target parameters, including surface roughness, vegetation cover and topography as well as soil moisture changes (Attema & Ulaby, 1978; Ulaby et al., 1978). The sensitivity of microwaves to dynamics in vegetation and soil moisture is attributed to differences of the dielectric constant which is orders of magnitude smaller for dry matter than for water. However, products from SAR backscatter fail to fully exploit the high resolution of SAR imagery and require compensation for impeding influences from vegetation growth and surface roughness (Brocca et al., 2011) which are associated with the necessity for extensive spatial averaging (Thoma et al., 2008).

There is clear evidence that vegetation and soil moisture dynamics also affect the SAR interferometric phase and have a significant influence on accurate estimations of ground displacements using InSAR (Zwieback et al., 2017). Yet, temporal changes of soil or vegetation moisture and the effects on the SAR interferometric phase have only recently been investigated and largely represent an unexplored field of research.

1.2. Differential Interferometric SAR (DInSAR) Closure Phase

DInSAR is a remote sensing technique used to measure and monitor deformations of the Earth's surface (Zwieback et al., 2017). The deformations are derived from the phase differences of radar images acquired at different times. If three or more images are available, one is able to compute the displacement between the first and third acquisition and derive an equal estimate from the sum of the two intermediate phases (De Zan et al., 2013). For those piston deformations it is expected that the wrapped sum of the three interferometric phases - the closure phase - is equal to zero in the absence of noise and influences due to elevation and atmospheric effects (Ferretti et al., 2011). However, this is not the case in acquired data. Phase excesses or deficits are common and also referred to as *inconsistency* or *lack of triangularity* (De Zan et al., 2015). Considering the single pixel level, closure phases from the triple product of three interferograms always add up to zero which one can easily verify mathematically:

$$\text{Single pixel level}: \quad \Phi_{i,j,k} = \phi_{i,j} + \phi_{j,k} - \phi_{i,k} = 0, \tag{1.1}$$

$$\text{Spatially averaged}: \quad \bar{\Phi}_{i,j,k} = \bar{\phi}_{i,j} + \bar{\phi}_{j,k} - \bar{\phi}_{i,k} \neq 0, \tag{1.2}$$

where $\Phi_{i,j,k}$ denotes the closure phase and $\phi_{i,j}$, $\phi_{j,k}$, $\phi_{i,k}$ the phases of three interferograms. The bar over the respective variable denotes the spatial average. It has to be stressed that spatial averaging is imperative to obtain non-zero phase closures (De Zan et al., 2015). In SAR terminology, spatial averaging of adjacent pixels is achieved through multilooking, which trades geometric resolution for phase noise reduction (Goldstein & Werner, 1998; Lee et al., 1998). In this study, multilooking is based on a per field basis, i.e. the interferograms are averaged over an homogenous land cover field which takes advantage of the spatial context of the respective area. The fact that an averaged phase closure $\bar{\Phi}_{i,j,k}$ is not necessarily equal to zero can be due to decorrelation noise alone (Zwieback et al.,

2016) or linked to multiple scattering effects, i.e. volume scattering and variations in the propagative properties of the semi-transparent medium which mainly depend on the moisture level (De Zan et al., 2015). As shown in recent publications, inconsistent phase closures are systematic and not the exception, pointing to an underlying geophysical cause, such as vegetation and soil moisture dynamics. A comparison of the spatial signatures of phase closure with land cover maps suggests a spatial and temporal correlation that is related to the characteristics of different land cover types. According to Zwieback et al. (2017), interferometric closure phase contains valuable information for vegetation and soil monitoring, especially at lower frequencies and short temporal baselines.

1.3. A Closure Phase Significance Test

In order to relate inconsistent closure phases to geophysical causes, it is necessary find a quantification of the phase noise component in closure phases. A numerical model developed in-house at the Geoscience and Remote Sensing department (GRS) at TU Delft (P. López-Dekker, personal communication) aims to reduce the decorrelation induced phase noise by predicting the expected standard deviation of an estimated phase closure, denoted as $\hat{\sigma}_{\hat{\Phi}}$, from three coherence magnitudes and the effective number of looks (ENL). The model is based on the null hypothesis that closure phases are caused by random phase noise as a result of the finite interferometric coherence. In order to derive estimates for $\hat{\sigma}_{\hat{\Phi}}$, the model performs several Monte Carlo simulations of an expected phase closure. From the resulting estimates of $\hat{\sigma}_{\hat{\Phi}}$, a closure phase significance ratio can be computed as the fraction of the observed phase closure and the respective estimate of $\hat{\sigma}_{\hat{\Phi}}$:

$$\text{Significance Ratio} = \frac{\bar{\Phi}^{\text{noise}}_{i,j,k} + \bar{\Phi}^{\text{geophysical}}_{i,j,k}}{\hat{\sigma}_{\hat{\Phi}}}. \tag{1.3}$$

The noise component of the observed phase closure $\bar{\Phi}_{i,j,k}$ is assumed to be normally distributed whereas the “geophysical” phase closure is not. A sufficient level of spatial averaging plays a crucial role in reducing the phase noise and isolating the geophysical component of the phase closure which leads to higher significance ratios.

The significance test is implemented in the Python application *TPT* ([T]riplet [P]rocessing [T]oolbox) (M. Manne, personal communication), which enables to process phase closures, estimate $\hat{\sigma}_{\hat{\Phi}}$, and export the closure phase significance ratio. The estimation of $\hat{\sigma}_{\hat{\Phi}}$ is based on the respective degree of multilooking as an input parameter, and three coherence values, which refer to the given closure phase. The three interferometric phases in the phase triplet each have a respective coherence value. For a given multilook factor, $\hat{\sigma}_{\hat{\Phi}}$ is estimated for all possible coherence value combinations and stored as a look-up table for a nearest neighbour search of the expected closure phase referring to three given coherence values. A refined setup of this coherence interpolator, and determining an appropriate number of model simulations, has not been implemented in *TPT* yet, in particular with regard to phase closures that are averaged over fields with similar land cover and different numbers of looks.

1.4. Research Objectives

The motivation behind this project was very curiosity driven in the beginning. The general goal was to further understand the factors that are influencing the closure phase over different land cover types. The initial approach was based on an exploratory characterisation of closure phase from Sentinel-1, standardized through a newly developed significance test. The potential of using closure phase for soil

and vegetation monitoring has been recognized since 2013 (De Zan et al., 2013) - the latest studies have focused on retrieving soil moisture over different land cover types (De Zan & Gomba, 2018; Gomba et al., 2021). **This study aimed to provide essential steps towards using the closure phase for vegetation and soil monitoring in the future**. The main objectives of this study were:

I. Present necessary steps to derive a closure phase that is spatially averaged with respect to land cover.

II. Provide guidelines to determine the parameter-settings to the closure phase significance test based on statistical properties.

III. Explore the impact of spatial averaging (multilooking) on the values and consistency of the significance ratio.

IV. Describe the spatial and temporal characteristics of the significance ratio for different land cover types.

In addition, differences between closure phases derived from VV or VH polarization and potential synergies and similarities with other observables are highlighted. The research objectives are addressed through the following *research questions*:

1. What processing steps are necessary and what parameter-based decisions are essential before performing a closure phase significance test?

2. What is the effect of spatial averaging on the significance ratio of closure phase and how does the consistency between individual time series of significance ratio change as a function of spatial averaging?

3. How consistent is the significance ratio spatially and temporally as a function of different land cover types?

4. Is the significance ratio consistent with spatial characteristics observed for SAR amplitude?

1.5. Research Approach

This study presents a characterisation of closure phase over different land cover types and was carried out for the region Occitanie in southern France, where fine resolution land cover maps for the Sentinel-1 era were available. SAR imagery for coherence magnitudes, amplitude backscatter and interferometric phases were downloaded and pre-processed to 20 m by 20 m pixels using a conventional boxcar filter for spatial multilooking. Only consecutive acquisitions were used to derive the complex interferograms assuming they are less affected by temporal decorrelation. Groups of pixels were further aggregated based on homogeneous land cover polygons given by the land cover maps. This enabled to average as many looks as possible by exploiting the spatial context of the scene. Phase closures were computed from the phases of the averaged complex interferograms. The subsequent closure phase significance test required adaptations in *TPT* for the field averaging approach, the setup of the coherence interpolator and a measure to determine a minimum number of model simulations. Determining these parameters was based on the statistical relationship between the degree of multilooking per field and the standard deviation in coherence magnitude estimation as outlined by Touzi et al. (1999).

The model is run for the respective number of looks per field. The resulting estimates for the expected closure phase standard deviation, $\hat{\sigma}_{\hat{\Phi}}$, are used to standardize the observed phase closures,

i.e. computing the significance ratio for closure phases. A sensitivity assessment is performed to determine the effect of multilooking on the magnitudes of significance ratio over different land cover types. The effect of spatial averaging on the similarity between individual closure phase time series was assessed using Pearson's *R* correlation coefficient. The spatial consistency of significance ratio was analysed using Pearson's *R* for time series at different spatial distances to one another. Lastly, temporal characteristics and spatial signatures were visualized and discussed in regard to different land cover types. In addition, differences between closure phases derived from VV or VH polarization and potential synergies and similarities with other observables are discussed.

1.6. Thesis Outline

This thesis commences with a review of recent studies about closure phase and its potential for soil and vegetation monitoring in chapter 2. This chapter also introduces the use of radar remote sensing for vegetation and soil monitoring more generally, and highlights the characteristics of Sentinel-1 C-Band and different polarizations (VV-pol, VH-pol). This chapter is continued with a detailed description of the closure phase significance model with its assumptions and parameter-inputs. The used data and the methodology applied to derive phase closures, the necessary model-adaptations, and the comprehensive sensitivity assessment is discussed in chapter 3. Chapter 4 presents the results in different sections, each primarily addressing one of the research objectives. A discussion of the results, limitations of this work and recommendations are provided in chapter 5. A conclusion of this thesis and an outlook for future work are summarized in chapter 6.

2

Background and Model

This chapter commences with a brief introduction to Sentinel-1 C-Band SAR and a characterisation of co- and cross-polarisation (VV-pol and VH-pol) for soil and vegetation monitoring. The chapter further introduces the term *coherence* and provides background on *closure phase*. At last, the numerical closure phase significance test is described in detail.

2.1. Dual Polarimetric Sentinel-1 C-Band SAR

The ability of a microwave radar to reach the ground's surface, primarily depends on its wavelength, polarization and the density of the above ground biomass (Ulaby et al., 1978). For shorter wavelength, such as Sentinel-1 C-Band, the radar's ability to penetrate layers of vegetation is lower than compared to longer wavelength (Ulaby et al., 1978). A comparison for the penetration capability of different bands is illustrated in Fig. 2.1, showing that the capability of the microwave to reach the ground surface increases as the wavelength increases. Over bare soils, an analysis carried out by Sekertekin et al. (2020) could show a good sensitivity of Sentinel-1 C-Band to soil moisture. On the contrary, for vegetation covered areas, the sensitivity of Sentinel-1 backscatter to soil moisture was lower than the backscatter obtained over bare soils. El Hajj et al. (2019) investigated the potential of radar at C-Band and L-Band to penetrate the canopy cover. They found that penetration is limited for developed vegetation cover (NDVI > 0.7) which suggests that backscatter values from Sentinel-1 are more representative for the characteristics of the vegetation canopy (Oon et al., 2019). The penetration depth is reduced by the moisture of the target as an inverse function of water content (Lillesand et al., 2015). Those findings indicate that Sentinel-1 C-Band can provide a great utility for applications that are aimed at monitoring of the above ground biomass while soil moisture assessments seem to be limited to surfaces with low vegetation cover.

Besides wavelength, the choice of polarization plays a crucial role when it comes to assessing either soil or vegetation properties. Sentinel-1 SAR in Interferometric Wide Swath (IW) mode supports operation in VV and VH polarization over land which initiated studies that aimed at testing the sensitivity of both polarizations and ratios thereof to dynamics in natural vegetation, crops and soil. Several studies could demonstrate that C-Band SAR in VV polarization has a better potential for soil moisture estimation than VH (Holah et al., 2005; Y.-y. Li et al., 2014; Zhang et al., 2008), thus co-polarized backscatter (VV) is often used to measure soil moisture variabilities (Autret et al., 1989; Ulaby et al., 1986). Dostálová

et al. (2018) assessed the annual seasonality in Sentinel-1 backscatter over forested areas and could demonstrate that VH polarization is sensitive to seasonal changes of tree foliage, suggesting that the leaves scatter the signal forward rather than scattering it back (Ahern et al., 1993). Many studies indicate that cross-polarized backscatter VH is sensitive to vegetated areas, as a consequence of its high sensitivity to volume scattering. As volume scattering mainly depends on the characteristics of the vegetation structure and its dielectric properties, cross-polarized backscatter can provide useful information about biomass dynamics, in particular vegetation water content (VWC) and phenology. (Bousbih et al., 2018; Bousbih et al., 2017; El Hajj et al., 2019; Gao et al., 2017; J. Li & Wang, 2018; McNairn & Brisco, 2004; Sekertekin et al., 2020; Vreugdenhil et al., 2018).

A consensus between the majority of studies exists on the significance of C-band backscatter for environmental monitoring, highlighting the sensitivity of co- and cross-polarization to soil and vegetation properties: While VV is more sensitive to soil conditions, VH could demonstrate a sensitivity to the influence of vegetation and could be used to discriminate between the effects of vegetation. Thus, in order to improve soil and vegetation monitoring, the complementary use of both cross- and co-polarized backscatter can provide more information in a SAR dataset. Yet, a broad sensitivity assessment for C-Band SAR interferometric phase derived from different polarizations has not been carried out.

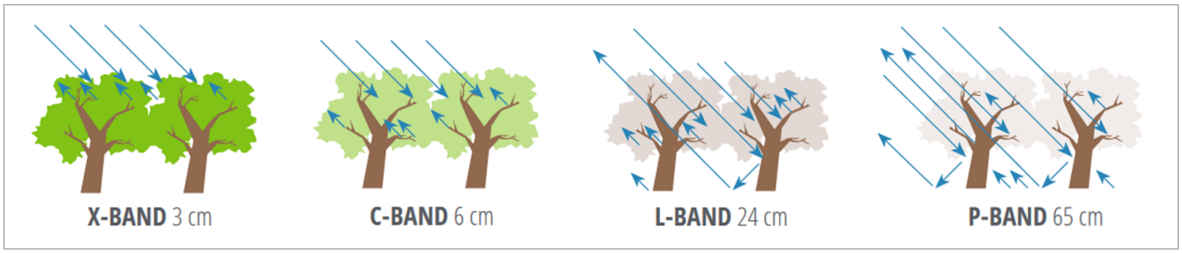


Figure 2.1: Sensitivity of SAR microwaves to forest structure and penetration into the canopy at different wavelengths used for airborne or spaceborne remote sensing observations of the land surface. The graphic shows the limited capability of C-Band SAR to penetrate vegetation canopy (Flores-Anderson et al., 2019).

2.2. Background on Closure Phase

A first relation between SAR interferometric phase and soil moisture was discovered in 1989 from centimeter scale movements that were associated with watering of the target area (Gabriel et al., 1989). Following this study, interferometric signals linked to changes in soil moisture have been reported also for other SAR data with the majority of explanations indicating motions that are linked to soil shrinking and swelling. Morrison et al. (2011) and Hensley et al. (2011) questioned this explanation by indicating that the motion of the soil surface is smaller than the corresponding C-band phase shift and that the phase change is too large to be solely attributed to surface deformation respectively. In 2013, De Zan et al. (2013) presented a model based on plane waves and Born scattering that is able to relate variations in soil moisture to L-band InSAR observables. This study could show a good correlation between phase triplet mismatches and soil moisture variations over bare soils. In 2015, De Zan et al. (2015) outlined explanations for phase inconsistencies based on models which challenge trivial interpretations for non-zero closure phases such as the presence of statistical noise as the only contribution. The main contributors to decorrelation (and hence phase triplet mismatches) are volumetric scattering (Zebker & Villasenor, 1992), in the presence of normal baselines, and variations in the properties of the semi-transparent media, mainly attributed to moisture variations (De Zan et al., 2015). Phase effects such as delays in atmospheric propagation and target motion are irrelevant for phase consistency, mean-

ing that the closure phase is expected to be zero (Zwieback et al., 2017). However, changes in the relative phase combined with changes in the relative intensity of two or more scatterers can increase phase inconsistency causing effects. This refers back to the model presented in De Zan et al. (2013) where variations in the dielectric permittivity, associated with moisture levels, would lead to variations in both amplitudes and phases for scatterers at different depths into the soil. These findings are derived from L-band aircraft imagery over agricultural fields, C-band imagery (ERS-1) SAR for the wider region of Rome and TerraSAR-X images at X-band for the “Afar triangle” in East Africa. Other observations from TropiScat experiments at P-band first indicate that the phase inconsistencies are related to water content variations in the sapwood of trees, thus changing their dielectric permittivity at different heights. De Zan et al. (2015) point out that the phase inconsistencies are, most likely, not actual range displacements but rather could be explained by water accumulation at mm-scale in the sapwood.

One major challenge that is associated with retrieving soil moisture from closure phase is the issue of ambiguity, i.e. that multiple soil moisture time series can correspond to the same closure phase. Zwieback et al. (2017) investigated the feasibility of separating soil moisture estimates, derived from closure phases and coherence magnitude, from surface displacements. In conclusion, Zwieback et al. (2017) have identified two main problems:

1) Neither the phase triplets nor the coherence magnitude contain enough information to estimate soil moisture due to ambiguity in the moisture order.

2) Soil moisture estimation is impaired by noise-sensitivity, model calibration errors and additional influences like changes in the surface roughness.

In 2018, De Zan and Gomba (2018) presented a model to solve for the occurring ambiguity problem associated with the moisture order and to deal with the spatio-temporal variability of soil moisture in a selection of experiments. The closure phase signs allow sorting acquisitions for increasing or decreasing moisture (De Zan & Gomba, 2018), however the signs alone can not solve for the correct moisture order, and rather provide constraints bounded by cyclic permutations of different moisture levels (Zwieback et al., 2017). The presented model was tested for ALOS-2/PALSAR-2 data over the area of Kumamoto in Japan. According to De Zan and Gomba (2018), an unexpected finding was that the moisture inversion yields a very clean closure phase signal over forested areas and a stable retrieval. Initially, the model was developed for bare soils, however the vegetation water content itself seems to contribute to the closure phases at L-band. The stronger signals over forested areas could be an indication of differences in moisture dynamics between soils and vegetation with lower variances for moisture in soils. Considering the similar dielectric behaviour of soils and vegetation, the two media can share the same interferometric modeling approach according to De Zan and Gomba (2018). From the fact that the retrieval showed promising results over forested areas, De Zan and Gomba (2018) assume that closure phases are sensitive to tree moisture variations.

2.3. DInSAR Coherence

The term coherence is defined as the normalized cross-correlation between two coregistered SAR images and indicates how much a scene has changed between two SAR acquisitions (Marino et al., 2012). Differences in phase (and amplitude) between scatterers for two acquisition times determine the coherence magnitude (Touzi et al., 1999). Among several factors that are contributing to decorrelation, interferometric coherence is affected by dielectric volume scattering (Zwieback et al., 2015) and can therefore be used as an indicator for vegetation dynamics and soil moisture changes (Palmisano et al., 2019). Since increasing or decreasing moisture changes cause the same coherence loss, coherence

magnitudes alone can not be used to retrieve soil or vegetation moisture unambiguously (De Zan et al., 2013).

Many studies have analysed coherence with respect to land cover classification, e.g. Jacob et al., 2020; Khalil et al., 2018; Nikaein et al., 2021; Santoro et al., 2007; Wegmüller et al., 2015, by exploiting the fact that 1) vegetated areas, i.e. temporally inconstant volumes, are affected by volume and temporal decorrelation and 2) for different cover types, vegetation emerges at different seasonal times. Temporal decorrelation is related to wind induced motion of the vegetation, precipitation events between two acquisitions or temperature changes, perhaps associated with freezing (Santoro et al., 2007). Those influences play a fundamental role even for short time intervals. Loss of coherence due to volume decorrelation was observed for increasing vegetation growth which suggests the possibility to use coherence for forest/non-forest mapping, to distinguish between different phenology stages (Ouaadi et al., 2020), harvest detection (Khabbazan et al., 2019) or to identify the beginning of a crop growth cycle (Nikolaeva et al., 2015; Santoro et al., 2007; Wegmüller et al., 2015). Volume decorrelation was observed even for small canopy thickness with the effect being different depending on the transmissivity of the vegetated layer, whereas bare soils generally show higher coherence magnitudes in comparison (Santoro et al., 2007).

Coherence magnitude could demonstrate a sensitivity to vegetation and soil moisture dynamics and can be exploited as a measure for phase closures that are not caused by decorrelation noise alone (Zwieback et al., 2016). Loss of coherence is generally a good indicator of whether some event has taken place that changes the dielectric properties or geometry of a scene (Zwieback et al., 2015). In the context of this thesis, coherence will be a useful observable to detect rapid vegetation growth or harvest events and can provide information about changing soil moisture, precipitation and freezing over low vegetated regions. The coherence magnitudes associated with the triplet interferograms will provide the essential input to the closure phase significance test.

2.4. Numerical Significance Test for Phase Closures

2.4.1. Aim of the Model

Phase triplets derived from three spatially averaged interferograms yield phase inconsistencies that are either solely induced by noise or originate from geophysical causes such as soil moisture variabilities. In general, the more two moisture values deviate from one another, the smaller the coherence (De Zan et al., 2015) and the larger the level of phase noise. The numerical model presented here is based on the null hypothesis that phase inconsistencies are solely caused by decorrelation, where the noise component is zero-mean Gaussian distributed. The model attempts to separate a possible geophysical phase closure from phase noise by estimating the standard deviation of expected phase closures, denoted as $\hat{\sigma}_{\hat{\Phi}}$, from simulated coherence values and the effective number of looks (ENL). The estimation of $\hat{\sigma}_{\hat{\Phi}}$ as a description of the phase noise component will be explained in detail in the model description in the following section. From the resulting estimates, a closure phase significance ratio is computed as the ratio between the observed phase closures $\bar{\Phi}_{ijk}$ and $\hat{\sigma}_{\hat{\Phi}}$:

$$\Psi_{ijk} = \frac{\bar{\Phi}_{ijk}}{\hat{\sigma}_{\hat{\Phi}}} \tag{2.1}$$

The greek letter Ψ for the significance ratio is used for a clearer separation from other observables and to facilitate the readability of this report. The significance test allows to make a comparison between fields with different coherences and different numbers of looks. The degree of multilooking of the

interferometric phase will be important in order to exploit the geophysical origins of phase closures (De Zan et al., 2015).

2.4.2. Description of the Model

The presented numerical model is based on the assumption that closure phases are solely produced by random noise associated with the finite interferometric coherence. The results are estimates for standard deviation of the expected phase closure $\hat{\sigma}_{\hat{\Phi}}$ associated with three coherence magnitudes.

In a first step, a complex random vector is defined as $y \sim \mathcal{CN}(0, C_3)$ representing the pixel in three SLC acquisition, and is characterised by a circular Gaussian distribution with zero-mean. The second-order statistics of the simulated vector are defined through the covariance matrix

$$C_3 = E\{yy^H\} = \Delta\Gamma\Delta^T, \tag{2.2}$$

where H is the Hermitian transposition, Δ the diagonal matrix with the square roots of the intensities, and T denotes the conjugate transpose. The coherence matrix Γ is a Hermitian matrix, i.e. a complex square matrix that is equal to its own conjugate transpose (Samiei-Esfahany & Hanssen, 2017), and is defined as

$$\Gamma_3 = \begin{bmatrix} 1 & \gamma_{12}^* & \gamma_{13}^* \\ \gamma_{12} & 1 & \gamma_{23}^* \\ \gamma_{13} & \gamma_{23} & 1 \end{bmatrix}. \tag{2.3}$$

Its elements can also be expressed in polar form given by

$$\gamma_{ij} = |\gamma_{ij}| e^{j\phi_{ij}}, \tag{2.4}$$

$$\phi_{ij} = \tan^{-1}\left\{\frac{\mathrm{Im}[\gamma_{ij}]}{\mathrm{Re}[\gamma_{ij}]}\right\}, \tag{2.5}$$

where Im and Re denote the imaginary and real part of the interferometric phase ϕ_{ij}, respectively.

A *Cholesky decomposition* is applied in the next step, which is the decomposition of a Hermitian, positive-definite matrix into the lower triangular matrix and its transpose:

$$C_3 = LL^T, \tag{2.6}$$

where L is the lower triangular matrix. The lower triangular matrices L are used to introduce the correlation coefficients γ in the white SLC pixel vector $z \sim \mathcal{CN}(0, I)$, where I is a 3 × 3 identity matrix. The simulated vector is then computed by multiplying the lower triangular matrix L with the pixel vector z:

$$y = L \cdot z. \tag{2.7}$$

An estimate of the covariance matrix in Eq. 2.2 can be derived from

$$\hat{C}_3 = \{\gamma_{ij}\} = \frac{1}{N_l} \sum_{i=1}^{N_l} y_i y_i^H, \tag{2.8}$$

where N_l denotes the number of looks that is used to average (multilook) the quantity in Eq. 2.8. In

order to derive estimates of the expected phase closures, the quantity

$$\hat{\Phi}_{ijk} = \angle(\hat{\gamma}_{ij} \cdot \hat{\gamma}_{jk} \cdot \hat{\gamma}^*_{ik}), \tag{2.9}$$

is computed for N_s realizations of the simulated matrix $\hat{C}_3$. The standard deviation $\hat{\sigma}_{\hat{\Phi}}$ over all simulations of $\hat{\Phi}_{ijk}$ is calculated for a given multilook factor (N_l) and stored in a look-up table. The look-up tables are used to derive a nearest neighbour interpolation of the expected closure phase from three observed coherence values. The model is assessed over a range of coherence value combinations and different numbers of looks. In order to illustrate $\hat{\sigma}_{\hat{\Phi}}$ as a function of coherence magnitude, a simplified case is considered for only a single coherence value instead of three, described by the coherence matrix

$$\Gamma_1 = \begin{bmatrix} 1 & \gamma^* & (\gamma^*)^2 \\ \gamma & 1 & \gamma^* \\ \gamma^2 & \gamma & 1 \end{bmatrix}. \tag{2.10}$$

Fig. 2.2 illustrates the relation for different multilooks.

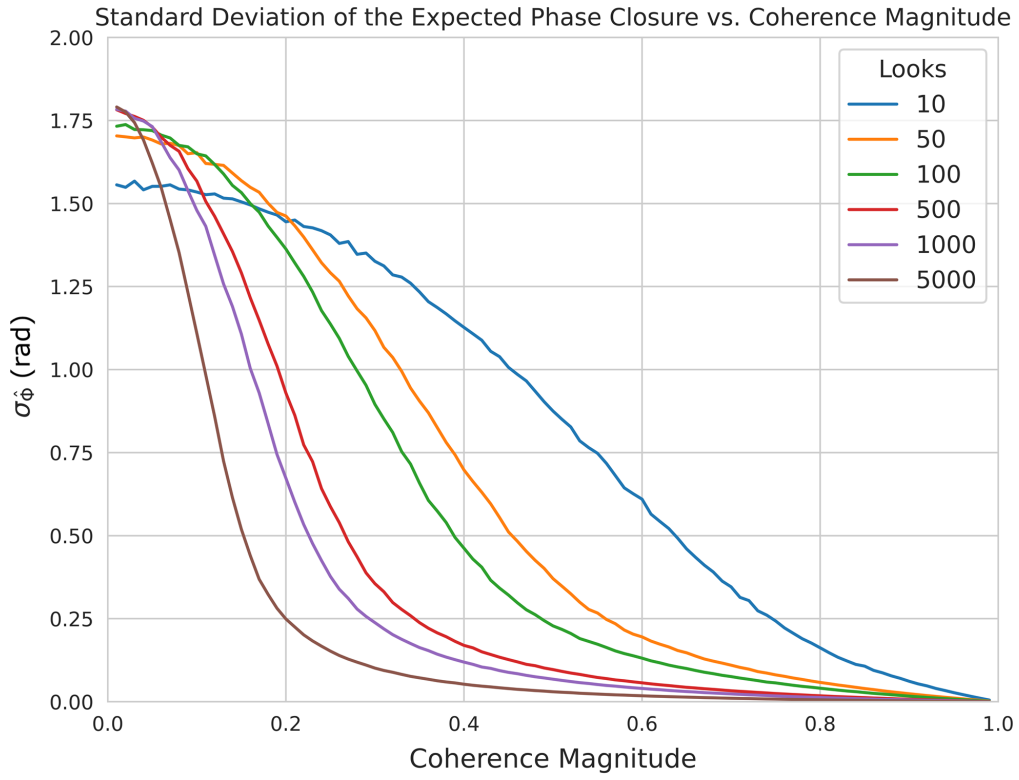


Figure 2.2: Expected Standard Deviation $\hat{\sigma}_{\hat{\Phi}}$ as a function of coherence magnitude for different looks. As one can see, the expected standard deviation is modelled higher for a smaller degree of multilooking and for lower coherence magnitudes. As the coherence or the number of looks increases, the expected standard deviation becomes smaller. (selfmade - modified from M. Manne)

3

Methodology and Data

Most recent studies have focused on soil moisture retrieval from closure phase (De Zan & Gomba, 2018; Gomba et al., 2021), however, a general characterisation of closure phase with respect to different land cover types has not been carried out yet. A better understanding of the spatial and temporal characteristics of closure phase, however, is important in order to use it for vegetation and soil monitoring in the future. Therefore, the approach of this study aimed to provide essential steps to derive a closure phase that is spatially aggregated based on land cover and standardized through a numerical significance test. The flow charts in Fig. 3.3 and Fig. 3.4 present the main steps that were required to derive a significance ratio. The tasks, in short, were the following:

1. Choosing a large enough region that featured many different land cover classes and high resolution land cover maps.
2. Acquiring and pre-processing SAR data for coherence magnitudes, amplitudes and interferometric phases.
3. Preparing and adapting the land cover maps.
4. Restoring the complex interferogram from coherence and amplitudes, averaging it (through spatial multilooking) over an homogeneous land cover area, and then recomputing the coherence and phase. Afterwards, computing the closure phase from the complex interferogram.
5. Determining the parameter-inputs for the closure phase significance test based on statistical properties of coherence and number of looks.
6. Performing the closure phase significance test over the complete research area.

The subsequent sensitivity assessment of the observed closure phase and significance ratio to spatial averaging aimed to demonstrate the effect of multilooking on the values and consistency of closure phase and to provide essential conclusions for a subsequent spatial and temporal characterisation. The influence that local changes of the target have on the linear similarity between time series of significance ratio is assessed by comparing Pearson's R between fields at different spatial distances. Lastly, the spatial characteristics of the significance ratio from different land covers are evaluated and compared with spatial signatures from amplitude backscatter. The following sections explain the methodological tasks in detail and highlight important decisions made in this study.

Table 3.1: Settings used in *RIPPL* in order to derive pre-multilooked amplitudes, coherences and interferograms.

RIPPL settings	
Setting	
Product Type	SLC
Sensor Mode	Interferometric Wide Swath
polarization	VH, VV
Digital Elevation Model	SRTM-1
Coordinate System	UTM
Pre-Multilook	20x20 m
Temporal Baseline	12 days

3.1. Study Region

The study was carried out for the region Occitanie in southern France with an area of ~75,000 km^2 and Toulouse in the center. The climate in the region is mild, and generally warm and temperate with a significant amount of rainfall even for the driest months. The air temperature over the study period (January 2017 to December 2019) was measured between 3°C and 29°C, with an average temperature at 14°C and an annual average precipitation at 823 mm water equivalent. The area is climatically classified as *Cfb* (oceanic climate) by Köppen and Geiger.

This region was selected for two main reasons: Firstly, it has a very broad spectrum of different land cover classes with agricultural areas and adjacent natural ecosystems. This enables the analysis and comparison of fine-scale heterogeneous areas (mostly agricultural sites) and homogeneous areas (primarily natural vegetation). Secondly, the availability of satellite data, land cover maps and supplementary data is abundant for the region of Southern France.

3.2. Pre-processing of SAR data in *RIPPL*

The simultaneous operation of the Sentinel-1 satellites A and B enables radar image acquisitions every 6 days. SAR Level-1 SLC (Single Look Complex) with a spatial resolution of 5 m by 20 m (range x azimuth) are downloaded and pre-processed using *RIPPL* ([R]adar [I]nterferometric [P]arallel [P]rocessing [L]ab), developed by G. Mulder (personal communication). The Python package enables the processing of SAR SLC data to create stacks of processed SAR data and is TU Delft's in-house Sentinel-1 InSAR processing tool. The SAR data used within the framework of this master's thesis are topographically corrected using the digital elevation model SRTM-1, then pre-processed to a 20x20 m grid resolution, using a conventional boxcar filter for spatial multilooking, and exported as ortho-images in UTM/WGS84 projection. The temporal baseline is set at 12 days to facilitate the composition of 6-day and 12-day interferometric products which is necessary to compute consecutive phase triplets and coherences as model inputs. An overview of the settings that were applied within *RIPPL* are presented in Table 3.1. Satellite track 110 (descending) has the greatest ground coverage and data availability for Occitanie, and will be the primary source of satellite-data for this thesis. Intermediate steps applied in *RIPPL* to derive the SAR observables are presented in the flow chart in Fig. 3.2.

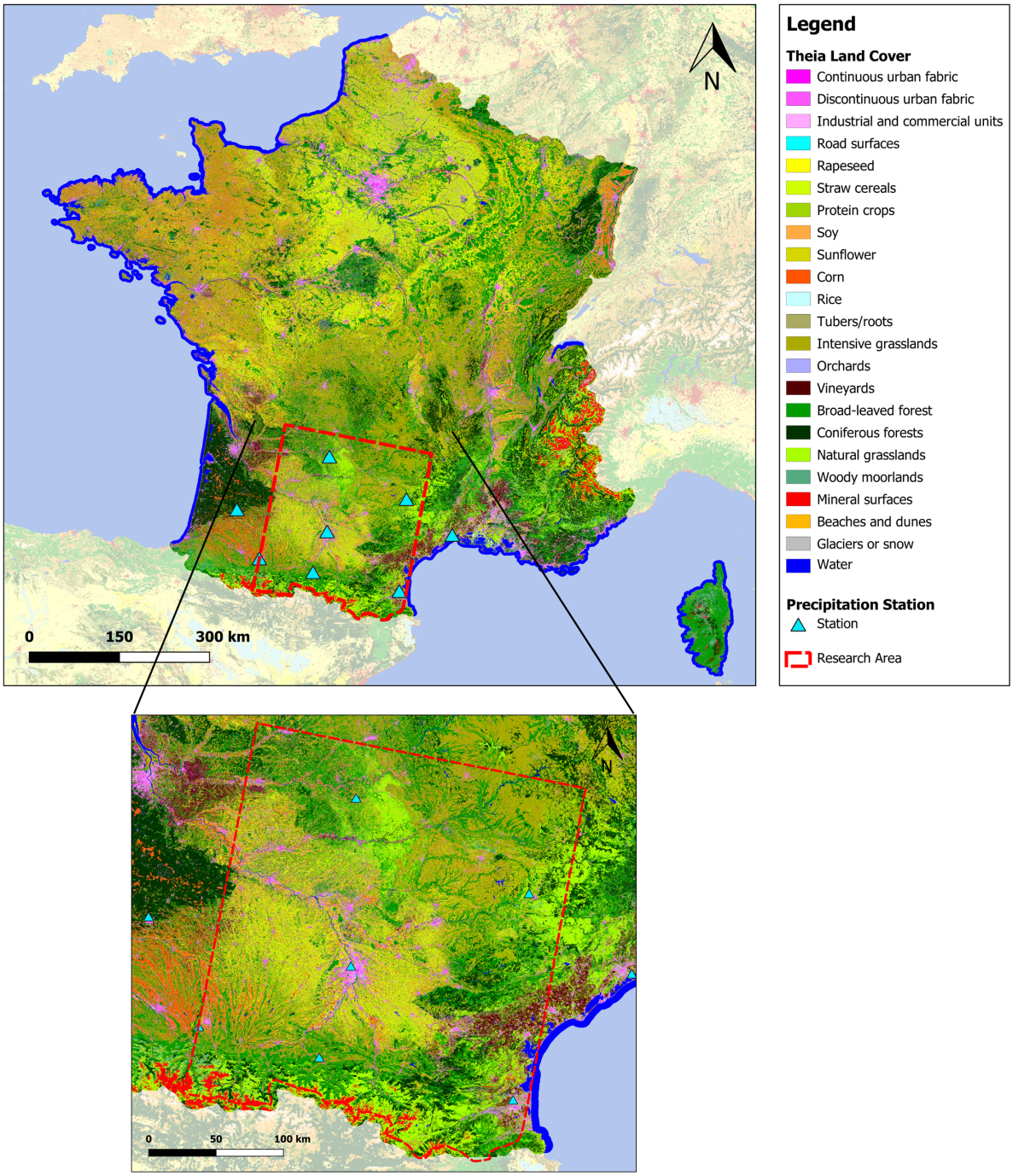


Figure 3.1: Land Cover Map (2018) for France showing the 23 different Theia land cover classes, the research area (red frame) that is covered by Sentinel-1 observations and the precipitation stations for that region.

3.3. Spatial Aggregation Based on Land Cover

3.3.1. Theia Land Cover Maps

The Theia Land Data Centre (theia.cnes.fr/) produces yearly updated land cover maps for France and other regions worldwide using Sentinel-2A and Sentinel-2B data at a resolution of 20 m (for vector formatted data). For this study, land cover maps for 2017 and 2018 were downloaded for several departments in southern France in order to aggregate the SAR pixels based on the land cover fields.

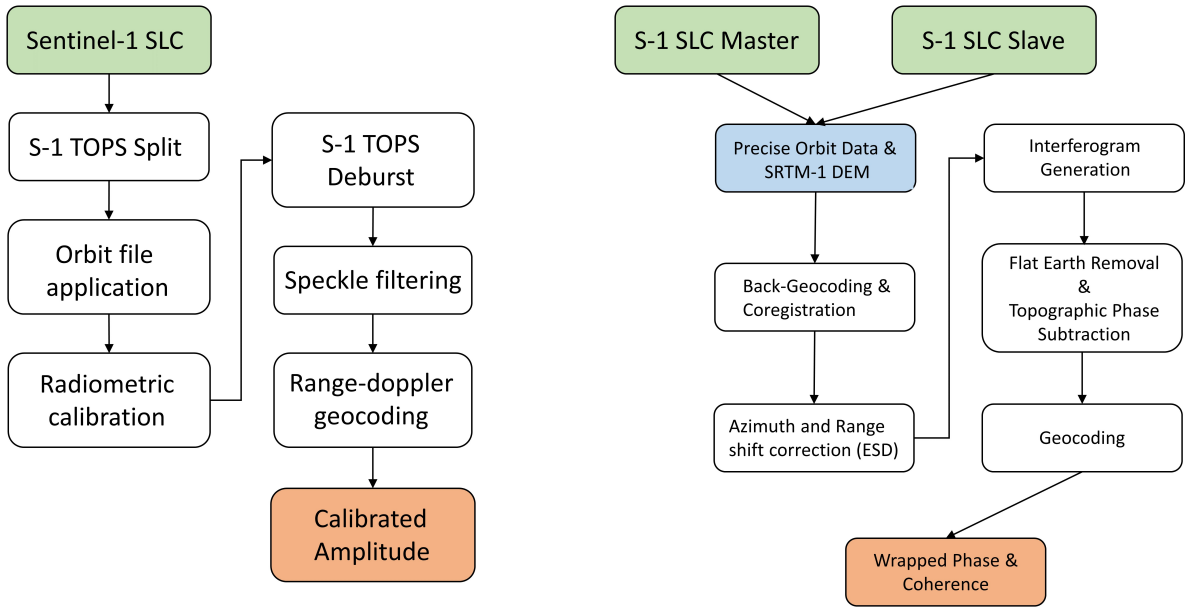


Figure 3.2: Flow chart for InSAR processing steps. Left: Processing steps applied in order to derive calibrated amplitudes. Right: Processing steps applied to derive coherence magnitudes and interferograms.

The nomenclature for 2017 features 17 land cover classes - since 2018, a 23-class nomenclature was applied. According to Theia, the nomenclature used since 2018 is fully compatible with the 17 classes employed for the 2017 land cover map. The land cover nomenclature for 2017 and 2018/2019 in comparison is presented in Table B.1 in the Appendix. For an overview of the research area with the land cover classification refer to Fig. 3.1.

The size of individual (multi-) polygons from the vector data ranges from a few meters up to several kilometers, in particular over homogeneous land cover areas. This was problematic for two reasons: Firstly, when exploring the interferogram images from *RIPPL's* output, atmospheric artifacts on the order of several kilometers were clearly visible (see Fig. 3.3, step 3). Assuming that no significant deformation of the area has occurred over short temporal baselines between 2017 and 2019, the majority of the interferometric fringes are attributable to atmospheric effects (Parizzi et al., 2006). In order to mitigate the conservation of atmospheric effects during the subsequent spatial averaging of the complex interferogram, it was necessary to ensure spatial averaging only over small regions. Secondly, averaging over larger areas could lead to the assumption that potential variabilities of an observable, when viewed spatially, are induced by a geophysical cause that is observed for the whole area of one land cover polygon while the effect could just be very local within this area and lesser in other parts of the polygon. The spatial resolution simply suffers from individual large multi-polygons.

In order to mitigate the above-mentioned effects for potential misinterpretations of the data, it has to be assured that each polygon only has a certain maximal diameter. For this research it is determined that each polygon should not be larger than 1 km in diameter, i.e. the spatial aggregation of pixels should not exceed this distance. This decision was based on a visual inspection of the interferograms from *RIPPL* to ensure that areas which exceed one phase cycle (due to atmospheric effects) are not aggregated spatially. The majority of the polygons within the land cover map fulfill this criteria of $<$ 1x1 km due to the many agricultural sites, however areas with homogeneous land cover such as forests and grasslands contain many multi-polygons that can reach 10+ km in diameter. A mesh-grid with a size of 1x1 km is applied to the whole land cover map in order to split those large multi-polygons

to the extend of each grid cell. The described problematic is illustrated in Figure 3.3 in step 3, where the two map zoom-ins show an example of two large multi-polygons (outlined in red) and the grid overlay that split these into smaller fields. As the pre-multilook of the SAR images is ~4 effective looks (ENL), the maximum number averaged pixels per polygon is 2,500, referring to ~10,000 looks.

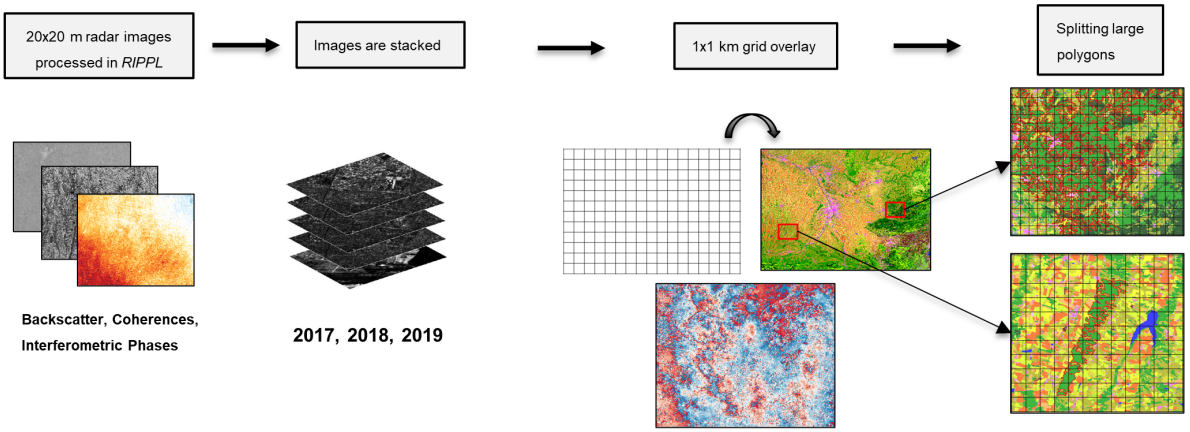


Figure 3.3: Processing routine for the data processing in *RIPPL* and the preparation of land cover maps.

3.3.2. Spatial Averaging of SAR Observables

SAR imagery from *RIPPL* for coherences, interferometric phases and calibrated amplitudes are stacked per year for all available dates between 2017 and 2019. The land cover maps that have been prepared previously are used to aggregate the pixel values based on each land cover field. As each year has its individual land cover map, the stacks are treated separately. The last date at the end of year 2017 and 2018 and the first date in 2018 and 2019 is copied respectively for the previous or following year to ensure continuous time series for the phase triplets where the last date of each year marks the “middle” date interferogram. No acquisitions were collected by Sentinel-1 for 05.02.2017, 29.08.2018, 01.02.2019, 07.02.2019 and 07.06.2019. Each pixel value from the SAR stacks is assigned to a polygon-ID from the land cover maps for further processing.

The spatial averaging approach over the perimeter of land cover fields that is applied here has two primary advantages. Firstly, the quality of the interferometric coherence estimate improves as the number of averaged pixels increases and the quality of the estimation corresponds to the extent of fields with similar land cover (Nikaein et al., 2021). Secondly, aggregating pixels over homogeneous fields, i.e. pixels that are assigned to the same land cover type, takes into account the spatio-temporal differences between land covers. This is in particular relevant for distributed scatterers, which usually refer to pixels with many small scatterers of similar size (Even & Schulz, 2018). For resolutions of some 10 meters, this is the case for most natural scatterers, such as forests, croplands, or bare soils (Even & Schulz, 2018), however there are differences depending on the land cover type and spatial resolution. Differences in the spatial properties of the target area are disregarded when multilooking is based on averaging windows (Hussain et al., 2013), as it is common practice in InSAR processing. Therefore it only made sense to aggregate groups of adjacent pixels, that share the same scattering mechanism, which was required to efficiently mitigate noise. Averaging of pixel-groups bounded to the extent of homogeneous fields eventually increased the quality of the significance test and allowed to draw conclusions about the behaviour and characteristics of closure phase over different land covers. The detailed averaging procedure for each SAR observable is explained in the following sections

(López-Martínez & Pottier, 2007; Touzi & Lopes, 1996):

Backscatter [dB]

1. Backscatter in decibel ($dB = 10\log_{10}(\frac{I}{I_0})$) is converted to intensities I.
2. The intensities are averaged based on all pixel values per polygon and converted back to backscatter [dB] for further analysis. The amplitudes A ($A = \sqrt{I}$) are preserved for the averaging of the complex interferogram.

Coherence

1. Coherence magnitudes are treated separately for 6-day and 12-day pairs during the spatial averaging. The same applies to the phases from the interferograms.
2. The coherence is defined as

$$\gamma = \frac{E[u_1 u_2^*]}{\sqrt{E[|u_1|^2]}\sqrt{E[|u_2|^2]}}; |\gamma| < 1; \angle\gamma = \phi_0, \tag{3.1}$$

where E is the expected value of the random variable x and $|u_1|$, $|u_2|$ are the amplitudes of the two initial images. The argument of γ, ϕ_0, equals the expected value of the interferometric phase. The coherence magnitudes are deconstructed to complex valued coherences using the phase angles from the interferograms and the preserved intensities:

$$\gamma_{complex} = |\gamma| \times e^{j\phi_i}, \tag{3.2}$$

where $|\gamma|$ is the coherence magnitude and ϕ_i is the phase of the interferogram.

3. The complex coherence is then multiplied with the amplitudes whereby the respective amplitudes are assigned to a master and a slave image that coincides with the 6-day and 12-day coherence dates:

$$\gamma_{renorm} = \gamma_{complex} \times A_{master} \times A_{slave}. \tag{3.3}$$

4. The master and slave intensities are derived by squaring the amplitudes. The intensities and γ_{renorm} are averaged.
5. The normalized averaged coherence is calculated by joining each averaged value referring to Eq. 3.1:

$$\bar{\gamma} = \frac{\bar{\gamma}_{renorm}}{\sqrt{\bar{I}_{master} \times \bar{I}_{slave}}}, \tag{3.4}$$

where $\bar{\gamma}$ and $\bar{I}$ denotes the averaged value of the respective instance. Eventually, the coherence magnitude is calculated by taking the absolute value of $\bar{\gamma}$, $|\bar{\gamma}|$.

Closure Phase

The closure phase is obtained from the triple product of three spatially averaged (consecutive) interferograms. The triple product is first calculated from the stored normalized complex and averaged coherence product $\bar{\gamma}$ where the asterisk denotes the complex conjugate:

$$\Gamma_{ijk} = \bar{\gamma}_{ij} \cdot \bar{\gamma}_{jk} \cdot \bar{\gamma^*}_{ik}. \tag{3.5}$$

Since the interferometric phase is an argument of the coherence, the closure phase is equal to the phase angle of the averaged normalized complex coherence triplet product:

$$\Phi_{ijk} = \angle\Gamma_{ijk}. \tag{3.6}$$

3.3.3. Supplementary Data

Theia VHSR Soil Moisture

The soil moisture product VHSR (Very High Spatial Resolution) from Theia is used as an additional measure in order to assess research objective IV. Theia in collaboration with CESBIO provides soil moisture maps at parcel scale on several sites in France. The applied approach from Theia is based on the synergic use of Sentinel-1 (radar sensor) and Sentinel-2 (optical sensor) data. For a description of vegetation, the NDVI (Normalized Differential Vegetation Index) was computed from Sentinel-2 imagery. The workflow applied to derive soil moisture estimates is described as follows (El Hajj et al., 2017):

In a first step, the agricultural areas were extracted based on Theia's land cover maps and the French RGP (Registered Geographical Parcels). NDVI mosaics derived from Sentinel-2 were used to partition the agricultural sites into homogeneous segments by applying a segmentation algorithm called Mean-Shift (Cheng, 1995). For each homogeneous segment, the mean incidence angle, mean backscattering coefficient in VV-pol, and the mean NDVI were computed. Neural networks were developed for different soil conditions (dry to slightly wet and very wet soil conditions). Lastly, the adequate neural network technique, selected based on in-situ rainfall data, was used to invert the radar signal and estimate the soil moisture.

Within a Pearson correlation analysis in this study, the correlation for the soil moisture time series from fields at different spatial distances to one another is computed. The results are compared to significance ratio and amplitude backscatter. A more detailed description of this analysis in provided later in section 3.6. A visual comparison between soil moisture maps and maps of SAR observables in this project is hampered due to the fact that the Sentinel-1 data used to estimate the soil moisture is derived from track 132 which collects images a couple of days later then Sentinel-1 over track 110 (SAR data used in this study).

Leaf Area Index (Copernicus Global Land Service)

The Leaf Area Index (LAI) is a dimensionless measure and is defined as half the total area of green elements of the canopy per unit horizontal ground area. It quantifies the amount of leaf material of all the canopy layers, including the understory which particularly contributes for forest vegetation (Copernicus Global Land Service, 2021). LAI is defined as

$$LAI = \frac{Leaf\ Area\ [m^2]}{Ground\ Area\ [m^2]}. \tag{3.7}$$

LAI-maps are provided by Copernicus Global Land Service (https://land.copernicus.eu/global/products/lai) with a spatial resolution of 300 x 300 m. The product is based on Sentinel-3/OLCI data and has been validated following the guidelines proposed in Morisette et al. (2006). The LAI data consists in an inter-comparison with other existing global products and a direct comparison with ground-based reference maps.

In order to obtain consistency with the other observables, LAI maps are resampled to a 50 x 50 m resolution using a bilinear interpolation algorithm, and aggregated based on Theia's land cover maps.

In the context of this research, the LAI will be used as an indicator for periods of enhanced vegetation coverage and growth and to characterize and discriminate different vegetation types. The LAI can be viewed as the counterpart to soil moisture, as the sensitivity of SAR observables to soil moisture is significantly influenced by vegetation.

Precipitation Data
Precipitation data is provided by the European Climate Assessment & Dataset (ECA&D, https://www.ecad.eu/) for in total eight stations distributed over southern France. The data will be used for indicating rain events in certain regions at certain times in order to better understand potential spatial variabilities in the SAR observables.

3.4. The Triplet Processing Toolbox (TPT)

The Triplet Processing Toolbox (TPT) was developed in-house at the GRS department of TU Delft (M. Manne, personal communication). The toolbox provides a practical application to compute closure phases from the complex interferogram and enables to perform the numerical closure phase significance test. The outputs from *TPT* are estimates of the Expected Phase Closure Standard Deviation $\hat{\sigma}_{\hat{\Phi}}$ and the Significance Ratio Ψ_{ijk} which is derived from the ratio of the observed closure phases $\bar{\Phi}_{ijk}$ divided by $\hat{\sigma}_{\hat{\Phi}}$.

3.4.1. Model Parameter and Toolbox Adaptations

The toolbox was initially built as an add-on for *RIPPL* where the outputs of *RIPPL* being the inputs to *TPT*. The issue at hand is that when SAR observables are multilooked to a grid using a conventional boxcar filter - which is the case with Geotiff outputs from *RIPPL* - it ignores the spatial context of the area over which the observables are spatially averaged. As the advanced method applied for this research was multilooking based on land cover polygons, *TPT* had to be adapted to meet the requirements of the new data format. With the adaptions, the significance test can be applied to phase closures averaged on a per field basis over different land cover types.

The model is run for different number of looks referring to different field sizes and different combinations of three coherence values that simulate the coherences referring to the phases in the phase triplet. The toolbox therefore required appropriate measures for three input parameters before applying the significance test.

In order to run the model for different combinations of three coherence values, a measure had to be established to determine the number of evenly spaced coherence values between 0 and 1 for which the model estimates $\hat{\sigma}_{\hat{\Phi}}$. This input parameter is referred to as the *Coherence Interpolator Steps*. The *Number of Looks* as the second input was derived from the number of pre-multilooked pixels per land cover field and is statistically linked to the *Coherence Interpolator Steps*. Lastly, an appropriate *Number of Model Realizations* of the expected phase closure for the computation of $\hat{\sigma}_{\hat{\Phi}}$ was determined with respect to the different number of looks per land cover field.

Given these three parameters, the toolbox estimates $\hat{\sigma}_{\hat{\Phi}}$ over different combinations of three coherence magnitudes and the number of looks per land cover field. The resulting estimates of $\hat{\sigma}_{\hat{\Phi}}$ are saved to look-up tables for a nearest neighbour search of the expected closure phase given three measured coherence magnitudes corresponding to the phase triplet observed over different land covers. The initial set-up of *TPT* performed the significance test based on only a single input value for each parameter which was applied to the whole closure phase SAR image. Since the adapted toolbox has to deal with phase closures averaged over land cover polygons, and hence different numbers of looks

per field over the whole research area, the modified significance test was performed with individual parameter settings for each land cover polygon.

The applied methods to determine the three described input parameters are discussed in detail in the following sections.

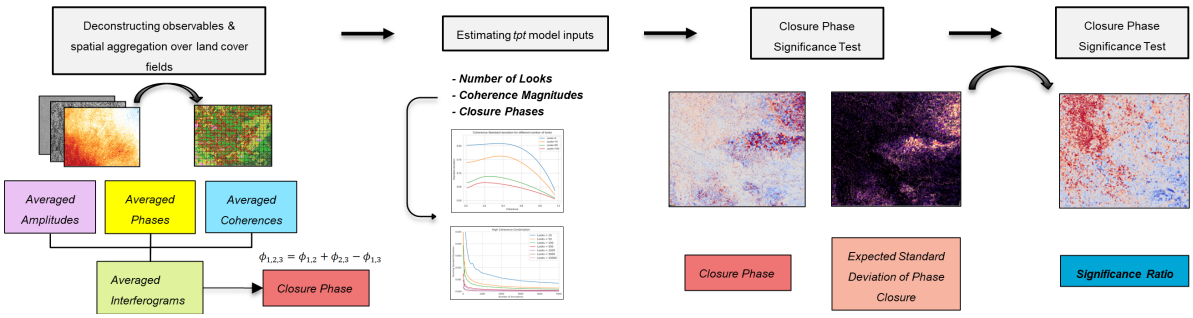


Figure 3.4: Processing routine for deriving a significance ratio and additional SAR observables. The figure shows the deconstruction of SAR observables in order to derive multilooked phase closures and the steps involved in adapting *TPT* and performing a significance test.

Number of Looks

The number of looks as an input to the significance model is given by the number of pixels averaged for the respective land cover polygon multiplied by the multilook of a single pixel. The pixels are counted per field during the averaging procedure. With a ground resolution of 5 x 20 m (range x azimuth) for the SLC, each individual pixel is pre-processed and multilooked to a 20 x 20 m grid resolution. Hence, the looks for a single pixel are ~4, the maximum number of looks is 10,000 due to the bounded polygon size of 1x1 km as defined by the grid overlay. The effect of an increasing number of looks on $\hat{\sigma}_{\hat{\Phi}}$ diminishes with a growing multilook as illustrated in Fig. 2.2. Statistically, there is no need to estimate $\hat{\sigma}_{\hat{\Phi}}$ for every single look above a certain level of multilooking. This is also relevant considering model runtimes and computational resources. Therefore, all fields with looks between 4 and 1000 (in increments of 4) are processed within the significance test; 1,000 to 5,000 looks are rounded to the tens digit, and 5,000 to 10,000 to the hundreds digit. Those “binned” looks provide the basis for a statistical derivation of the coherence interpolator steps and the number of model realizations.

Coherence Interpolator Steps

The coherence steps refer to the increments in coherence for the different combinations that are used to estimate $\hat{\sigma}_{\hat{\Phi}}$. For example, an input parameter of 10 refers to 10 evenly spaced coherences values between 0 and 1 ($x \in \mathbb{R}, 0 < x < 1$). In this study it is determined that the interpolator coherence steps should not be smaller than the standard deviation in coherence magnitude which depends on the degree of multilooking. Touzi et al. (1999) investigated the accuracy and precision of coherence estimation as a function of the coherence map resolution. Coherence estimates are shown to be significantly biased towards larger coherence estimates, in particular under low coherence conditions whereby the bias reduces with an increasing number of looks. The coherence bias is calculated in Eq. 3.9 and visualized for different number of looks in Fig. 3.5.

In order to determine appropriate coherence steps, Touzi et al. (1999) formulated a set of equations from which the standard deviation of coherence estimates can be calculated as a function of looks:

$$E(|\hat{\gamma}|^2) = \frac{\Gamma(L)\Gamma(2)}{\Gamma(L+1)} \times_3 F_2(2, L, L; L+1, 1; |\gamma|^2) \times (1 - |\gamma|^2)^L, \tag{3.8}$$

$$E(|\hat{\gamma}|) = \frac{\Gamma(L)\Gamma(1+1/2)}{\Gamma(L+1/2)} \times_3 F_2(3/2, L, L; L+1/2, 1; |\gamma|^2) \times (1-|\gamma|^2)^L, \tag{3.9}$$

where ${}_pF_q$ is the generalized hypergeometric function, $|\hat{\gamma}|$ the estimated sample coherence magnitude, L the number of looks, $|\gamma|$ the actual coherence magnitude, and the moment of order given by k. The standard deviation of $|\hat{\gamma}|$ can be calculated by

$$std(|\hat{\gamma}|) = \sqrt{E(|\hat{\gamma}|^2) - E(|\hat{\gamma}|)^2}, \tag{3.10}$$

where $E(|\hat{\gamma}|^2)$ is given by 3.8 and $E(|\hat{\gamma}|)$ by 3.9. Eventually, the standard deviation is calculated for the

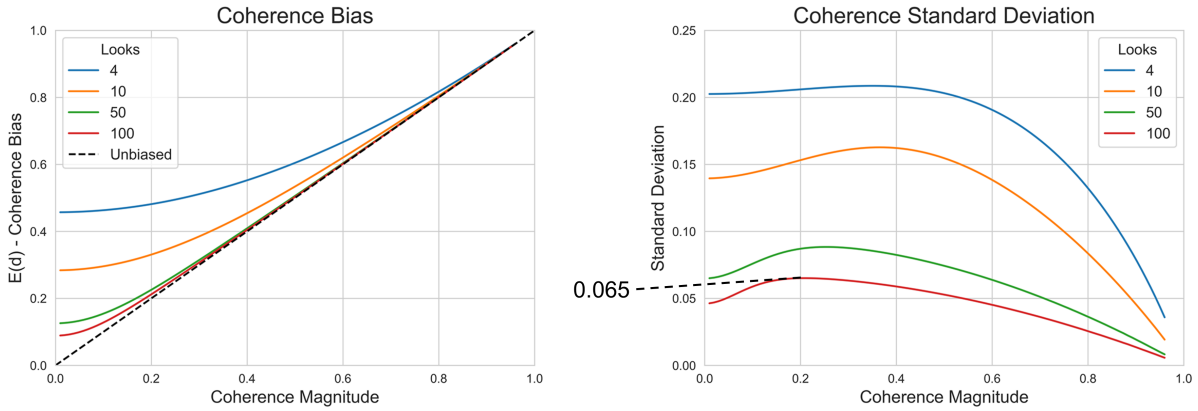


Figure 3.5: Coherence bias (left) and coherence standard deviation (right) as a function of coherence magnitude for different looks. The figure indicates that the standard deviation in coherence decreases 1) as the number of looks increases or 2) as the coherence magnitudes increase.

given set of coherences by considering the number of looks (L) with respect to each land cover polygon. Since the standard deviation varies for different coherence magnitudes, the highest standard deviation computed is considered as a measure for appropriate interpolator steps. For the example plot given in Fig. 3.5, the red line refers to a field with a multilook of 100. The maximum standard deviation is 0.065 ($|\hat{\gamma}|$ ~0.20) which is further considered an appropriate interpolator step size. With 1/0.065 = 15.38, the parameter-input is rounded down to the next integer (= 15). For this study it was determined that the maximum number of coherence steps is 100 as the number of possible combinations, and more importantly the runtime, grows exponentially. Fig. 3.6 visualizes the coherence steps as a function of possible combinations where step sizes of 0.01, i.e. 100 steps between 0 and 1, equal ~160,000 possible combinations.

Number of Model Realizations

From a given set of coherence combinations and the number of looks, the model performs several Monte Carlo simulations of the multilooked expected phase closures. Thereafter, the standard deviation $\hat{\sigma}_{\hat{\Phi}}$ over all model realizations of the expected phase closures is computed; the result is an estimate of $\hat{\sigma}_{\hat{\Phi}}$ for each given coherence combination. The coherence combinations ranged between "low" combinations, e.g. 0.1, 0.1, 0.1, and "high" combinations, e.g. 0.9, 0.9, 0.9. In order to understand the effect of consecutive model realizations on the final estimate of $\hat{\sigma}_{\hat{\Phi}}$, 20,000 model realizations of the expected phase closure were computed with different multilooks. A running standard deviation then estimated the standard deviation along with the mean in a single pass of all model realizations. The results for a high and a low coherence combination and different looks is illustrated in Fig. 3.7. What can be deduced from the figure is the point when the running standard deviation converges, meaning that additional model realizations did not change the final estimate of the expected $\hat{\sigma}_{\hat{\Phi}}$ significantly.

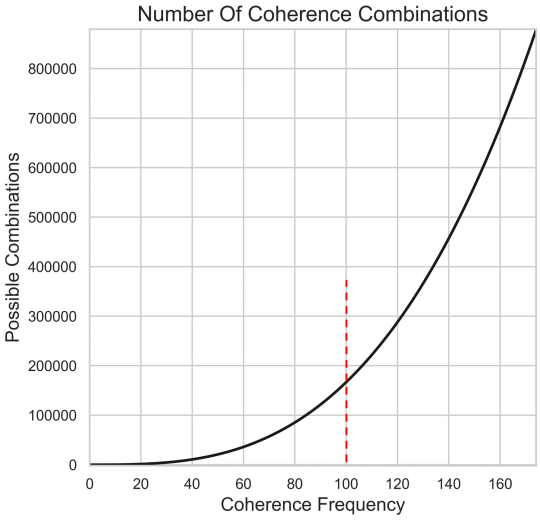


Figure 3.6: Number of possible coherence combinations as a function of coherence frequency. The number of combinations grows exponentially as the coherence interpolator steps decrease, i.e. the coherence frequency increases.

This point depends on both the number of looks and the coherence combination considered for the estimation. The number of looks is important here since it already decreases the standard deviation of the estimated phase closures considerably. Since the simulation grows exponentially with the number of coherence combinations, the degree of multilooking and the number of simulations, finding a reasonable number of model realizations was therefore crucial to avoid unnecessary and prolonged simulation runtimes. Generally, the running standard deviation converges faster for higher numbers of looks and higher coherence magnitude combinations (e.g. 0.9, 0.9, 0.9). Based on the statistics presented in Fig. 3.7, it was determined that the point when the running standard deviation for the lowest coherence combination reaches 0.01 radians (~0.57 degrees) is sufficient. The respective number of realizations is assessed over all "binned" looks that have been defined above. It was further determined that the minimum number of simulations should not be less than 500.

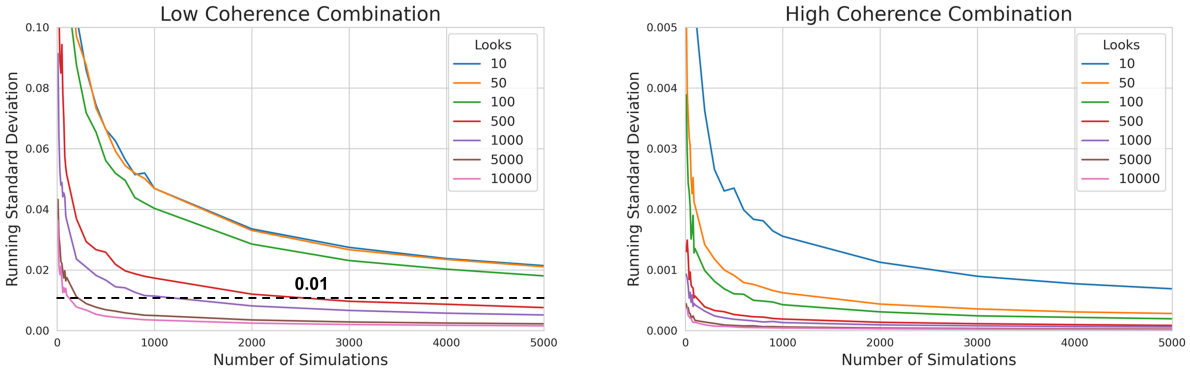


Figure 3.7: Running Standard Deviation as a function of Monte Carlo simulations for different multilooks. With an increasing degree of multilooking, the decrease in standard deviation in the estimates converges faster. Left: Running standard deviation for a low coherence combination (0.1, 0.1, 0.1). Right: Running standard deviation for a high coherence combination (0.9, 0.9, 0.9). Note that the y-axis for the high and the low coherence combination are scaled differently.

3.4.2. Significance Test within *TPT*

A fixed relationship between the model-parameters as an input to *TPT* can be established based on the definitions of the parameters in the previous sections. Fig. 3.8 presents the number of coherence

steps as a function of multilooks and the number of simulations as a function of multilooks. The number of realizations decreases as the number of looks increases, while the standard deviation in coherence improves with increasing multilooks, leading to smaller coherence steps. Given this fixed parameter relationship, 750 look-up tables have been computed and stored in total.

In a first step the number of looks for each input polygon field was either used directly or was assigned to a look-bin according to the definitions in the previous section. The number of coherence steps and number of model realizations was assigned to each individual polygon-field following the statistical relationship that was described above. Interpolator files for each look with all possible coherence combinations are computed from the look-up tables. Each polygon-field had 61 coherence-triplets (per year), i.e. coherence magnitudes were acquired every 6 days. The interpolator applies a nearest neighbour search for the closest coherence combination to the observed coherence-triplet for each time-step and returns the corresponding estimate of $\hat{\sigma}_{\hat{\Phi}}$. The significance ratios were calculated from the observed phase closures and $\hat{\sigma}_{\hat{\Phi}}$ for each time-step. A brief overview of the essential steps applied to derive the significance ratio is presented in Fig. 3.4.

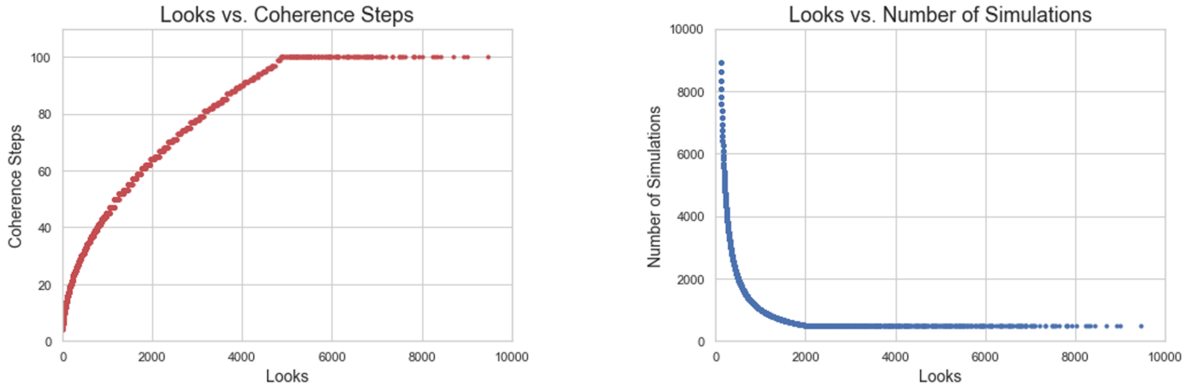


Figure 3.8: Left: Coherence steps as a function of multilook. The minimum number of steps is at 4, the maximum is limited to 100 which is reached at about 5000 looks. Right: Number of model realization as a function of multilook. The number of necessary model realizations is estimated as high as 200,000 realization for single pixel, but rapidly decreases as the number of looks increases. The minimum number of model simulations is kept at 500.

3.5. Choice of Land Cover Classes

In this study, the spatial and temporal characteristics of the closure phase and its significance ratio have been investigated mainly for three land cover classes *Broad-leaved forest*, *Natural Grassland*, and *Agricultural Crops*. Agricultural sites are further divided in *summer*, *winter*, and *perennial* (Intensive Grassland) crops. Those land covers are chosen with respect to characteristic properties, including annual phenology, vegetation density, type and location, in order to explore the effects of different vegetation properties on the closure phase. The different land covers are characterized as follows (Corine Land Cover Nomenclature, 2018):

- *Broad-leaved Forest* composes in principal deciduous trees, but also shrub and bush understory, where broad-leaved vegetation predominates. The crown cover density is at least 30% or a minimum of 500 subjects/ha in density. The cover type includes primarily deciduous tree species but also evergreens, palm trees and herbaceous vegetation (grasses and herbs), if < 25 ha.
- *Natural Grasslands* include low vegetated areas and pastures and are in comparison with intensive grassland and agricultural areas not used for perennial or annual crop growth.

- *Annual Crops* include rainfed agriculturally used land parcels that are harvested and are usually under a crop rotation system. Theia further divides annual crops in summer and winter crops for the 2017 nomenclature and splits summer and winter crops in individual crop types for 2018 and 2019. A detailed overview is provided in the Appendix.
- *Intensive Grasslands* are characterized by perennial agricultural use. The majority of patterns classified as intensive grassland are located in between agricultural sites of annual crops and regions that are used as vineyards.

3.6. Statistical Analysis

3.6.1. Pearson Correlation Coefficient (r) Analysis

The Pearson correlation coefficient is a measure of the strength of a linear similarity between two variables and is denoted by *R*:

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{3.11}$$

where *x* and *y* are two vectors of length *i* and $\bar{x}$, $\bar{y}$ corresponds to the means of *x* and *y*, respectively. The guidelines in Table 3.2 indicate the strength of the similarity with respect to *R* value ranges.

Table 3.2: Pearson's *R* guidelines evaluating the performance of the correlation analysis.

Pearson's R	
Strength of similarity	**Range**
very weak	0 - 0.2
weak	0.2 - 0.4
moderate	0.4 - 0.6
strong	0.6 - 0.8
very strong	0.8 - 1.0

In this thesis, Pearson's *R* is used as a measure of the linear relationship between significance ratio time series. The similarity is assessed as a function of 1) spatial aggregation, and 2) spatial distance between the time series.

Spatial Aggregation Assessment

In order to assess the effect of spatial averaging on the similarity between the significance ratio time series, time series which are derived from different multilooks are binned with bin sizes of 100 looks for $\leq$ 1,000 looks, and 1000 looks for $\leq$ 10,000. Pearson's *R* is calculated for all pairs within each bin.

Distance Assessment

In order to counter any effects due to different degrees of spatial aggregation between the time series, the data is filtered to only include the ones that are $>$ 8000 looks, since the phase noise expectably decreases with further averaging.
The center location of each of those polygons is calculated. After, the euclidean distance between each

center is estimated and grouped for increasing intervals of 20 km. Pearson's *R* is estimated for each distance pair.

3.6.2. Root Mean Squares (RMS)

The Root Mean Square (RMS) is defined as:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{3.12}$$

The RMS was used to analyse the effect of spatial averaging on the magnitudes of closure phase. Its values were computed over time series of closure phase with respect to different degrees of multilooking and over different land cover types. To ensure a large enough sampling for each multilook, the land cover fields were binned with respect to their degree of spatial aggregation in increments of 100 looks.

The reason why the RMS was chosen over the mean or other statistical measures for this assessment requires a brief discussion:

The phase closures are negatives and positives such as sinusoids. The mean of the phase closures would just reveal the influence on its central tendency. Phase closures that are solely produced by phase noise, as predicted by the significance model, would necessarily have zero mean and a significance ratio of |1|. With zero mean, the values for RMS and standard deviation are equal. However, if the mean of a phase closure would be consistently at a value e.g. above zero, which indicates a geophysical signal component, the standard deviation would be small and dismiss the fact that there is a consistent closure phase signal. Therefore it was more meaningful to look at the powers. Due to the squaring of the individual phase closures, peak values receive more weight which is a favorable feature as closure errors with a large signal-to-noise ratio are particularly interesting. The RMS is therefore rather specific and mainly tied to electrical engineering and signal sciences.

3.7. Work Environment, Runtimes and Processing Bottlenecks

In terms of work environment, a project space with 25 TB was provided by TU Delft as well as a HPC account for the operation and storage of programs and scripts. The pre-processing part of the thesis research alone took 3 months in total for SAR data from a satellite track with 75,000 km^2 ground coverage and three years of image acquisitions. A successful processing in *RIPPL* firstly depends on the availability of orbit files and DEM data, accessibility to the Alaska Satellite Facility (ASF) for the Sentinel-1 SLC products, and the right choice and batch-script settings for computing nodes at the HPC cluster from TU Delft. As the time or space dimensions of the required data increase, the chance that those essential files are either temporally not accessible, fail to download or fail to process correctly is very high. For successful processing of the data it is advised to monitor the use of resources on the HPC cluster closely, and inform the cluster manager before attempting to process with *RIPPL*. Patience, nerves of steel and a good oversight are necessary in order to successfully process those observables.

4

Results

This chapter presents the results obtained from the closure phase significance test and illustrates spatial and temporal characteristics of significance ratio and closure phase over different land cover types.

Section 4.1 first shows the results from the significance test in relation to the observed phase closures and coherence magnitudes. The impact that the standardization had on the observed phase closures is important to better understand its characteristics over different land covers. This section also aimed to provide a first link between land cover and the observable characteristics. Section 4.2 demonstrates the effect that different degrees of spatial averaging had on phase closure and significance ratio values, and the correlation between its time series. In Section 4.3, the conclusions from the previous section were considered to show the characteristics of significance ratio time series in comparison with LAI and coherence magnitudes over different land covers. The section continues with an assessment of the correlation between time series of significance ratio as a function of spatial distance, in comparison with results for amplitude backscatter and Theia's soil moisture product. Based on the results obtained from the time series comparison, the section concludes by comparing the spatial characteristics between significance ratio and amplitude by considering their sensitivity to soil moisture over low vegetated soils.

4.1. Comparison of Closure Phase and Significance Ratio

The observed phase closures $\bar{\Phi}_{i,j,k}$ consist of a noise-induced signal component and a possible geophysical component which could contain information about vegetation and soil properties over the observed area. The closure phase significance test was applied to each measured $\bar{\Phi}_{i,j,k}$ with the intention to standardize it with an estimate of the expected phase closure standard deviation $\hat{\sigma}_{\hat{\Phi}}$. From $\bar{\Phi}_{i,j,k}$ and $\hat{\sigma}_{\hat{\Phi}}$, the significance ratio $\Psi_{i,j,k}$ was computed as the ratio of both values with the aim to provide a measure of the signal-to-noise ratio of the observed phase closures.

The goal of this section was to compare the spatial signatures of closure phase before and after standardization and to visualize the relation between $\hat{\sigma}_{\hat{\Phi}}$ and the coherence magnitude ($|\gamma|$) with respect to different land cover types. Understanding the relation between the different observables is essential for results presented in the following sections.

A land cover map over a sub-region of the research area in southern France is illustrated in Fig. 4.1.

The maps of the observables in this section refer to the same region. Fig. 4.2 shows the observed closure phases $\bar{\Phi}_{i,j,k}$ computed from three interferograms between 26.11.2017 and 08.12.2017 and the significance ratio $\Psi_{i,j,k}$. Choosing a phase triplet from interferograms in November and December implies that the vegetation cover is generally lower, in particular over agricultural areas where crops have been harvested. The $\hat{\sigma}_{\hat{\Phi}}$ and the three coherence magnitudes $|\overline{\gamma}|$ referring to the phase triplet are presented in Fig. 4.3. The $\bar{\Phi}_{i,j,k}$- map indicates different phase levels across the scene with consistently higher or lower phase closures in certain regions. Since the model predicts $\hat{\sigma}_{\hat{\Phi}}$ based on the respective coherence magnitudes $\gamma_{i,j}$, $\gamma_{j,k}$ and $\gamma_{i,k}$, regions that suffer more from decorrelation are associated with higher estimates for $\hat{\sigma}_{\hat{\Phi}}$ as those regions are expected to be have higher levels of phase noise (see Fig. 2.2). To account for the spatial variability of coherence across different areas, $\Psi_{i,j,k}$ provides an indication of patterns of phase closures that are large with respect to $\hat{\sigma}_{\hat{\Phi}}$. Significance ratios that are low in absolute terms, i.e. $\sim$ +/- 1, indicate that the observed closure phase can be justified by decorrelation phase noise. Comparing $\bar{\Phi}_{i,j,k}$ and $\Psi_{i,j,k}$ in Fig. 4.2 clearly shows the effect that the standardization from the significance test has on the spatial characteristics, in the attempt to reveal “true” phase closures. A few observations can be made when comparing the closure phase to the significance ratio:

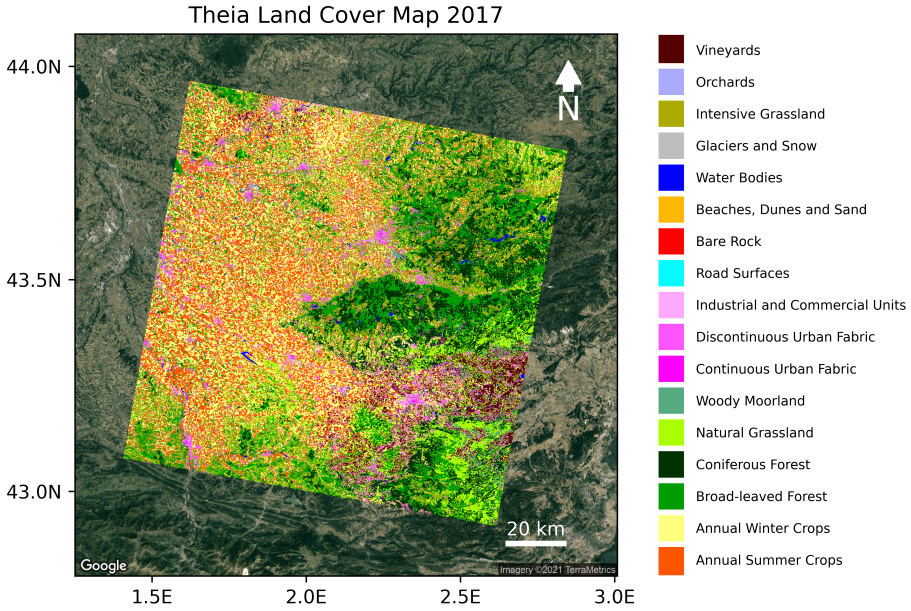


Figure 4.1: Theia land cover map for 2017 provided for a region in southern France.

1. In Fig. 4.2 (a), areas that show higher magnitudes of $\bar{\Phi}_{i,j,k}$ (indicated by ‘1’) obtain lower significance ratios (closer to +/- 1) as shown in (b) while areas with lower $\bar{\Phi}_{i,j,k}$ obtain noticeably larger significance ratios ($> |1|$) (‘2’ respectively in the maps). The region in the upper right corner of the $\bar{\Phi}_{i,j,k}$- map seems to have exceptionally large phase levels compared to the upper and lower left part. Areas of large $\bar{\Phi}_{i,j,k}$ are associated with low coherence magnitudes which receive high estimates for $\hat{\sigma}_{\hat{\Phi}}$ from the significance test as illustrated in Fig. 4.3 (a). Because of that the significance ratio is comparably low and it can be assumed that the large $\bar{\Phi}_{i,j,k}$ mostly originate from phase noise. Areas with a relatively high significance ratios are consistent with regions where the coherence is higher.

2. Relatively large phase closures in the $\bar{\Phi}_{i,j,k}$- map in Fig. 4.2 (a), indicated with ‘1’, appear purple rather than distinctively blue or red as the colorbar indicates negative and positive phase closures.

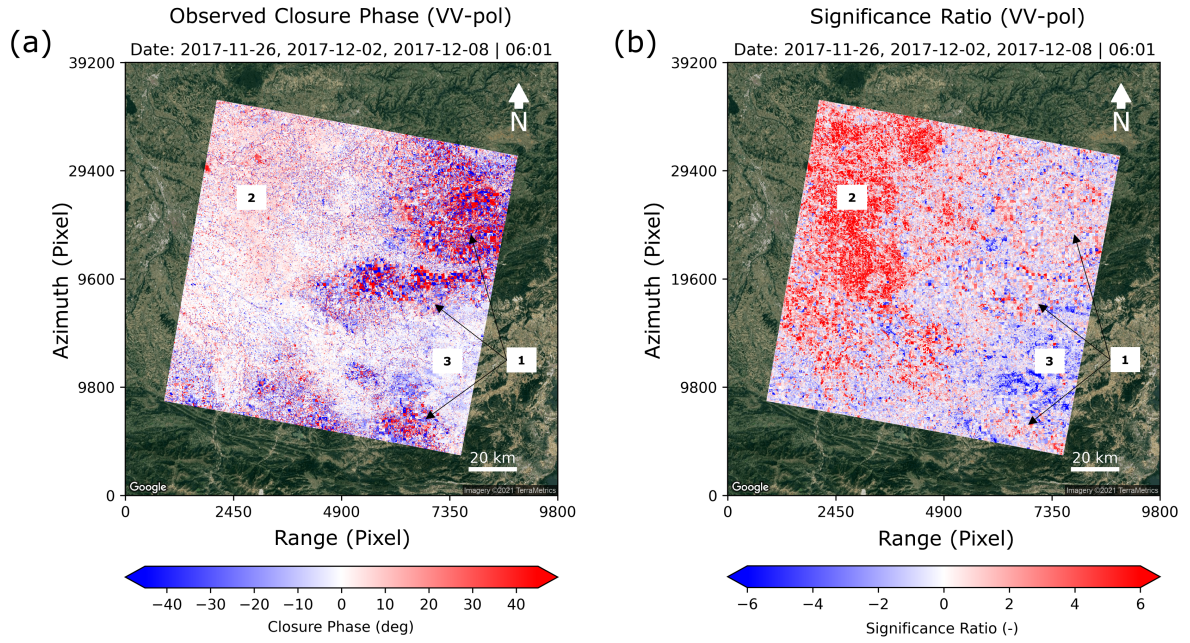


Figure 4.2: Spatial signatures of (a) the observed phase closure $\bar{\Phi}_{i,j,k}$ and (b) the significance ratio $\Psi_{i,j,k}$ for a region in southern France.

Zooming in and investigating the individual pixels and their neighborhood shows that the phase closures alter between positives and negatives (red and blue) which causes the salt and pepper like purple appearance and the associated low spatial consistency. High magnitudes and varying signs of the closure phase associated with low significance ratios indicate that it is unlikely to be due to a geophysical cause. Those regions are also consistent with low coherence magnitudes as shown in Fig. 4.3 (b).

3. The distinctive patterns of high $\Psi_{i,j,k}$ magnitudes ('2' in Fig. 4.2 (b)) appear to be spatially more consistent with visually separate areas of negative and positive signs in contrast to the higher variability of neighboring pixels in the $\bar{\Phi}_{i,j,k}$ image. This effect in particular is noticeable for regions with larger $\bar{\Phi}_{i,j,k}$.

4. Comparing the land cover map in Fig. 4.1 with the observables, one can notice that the general patterns match with the land cover boundaries. The forested regions ('1' in map) are associated with larger phase closures $\bar{\Phi}_{i,j,k}$ and low significance ratios while the southeastern part, primarily covered with grassland ('3' in map), shows lower, here negative closure phases and significance ratios. The western part of the image, mostly covered by annual and perennial crops (indicated by '2'), also shows low phase closures that appear relatively consistent in space. In comparison, this region obtains the highest significance ratios. The natural grassland areas show larger significance ratios as well while the forested areas indicate less significant phase closures in comparison.

The comparison between $\bar{\Phi}_{i,j,k}$ and $\Psi_{i,j,k}$ has demonstrated the importance of applying a closure phase significance test to separate the phase noise component associated with low coherences from phase closures that obtain a high significance ratio. Those phase closures are particularly interesting as they indicate that a large fraction of the closure phase can not be explained by decorrelation noise which suggests evidence for an underlying geophysical origin. The effect of different vegetation covers

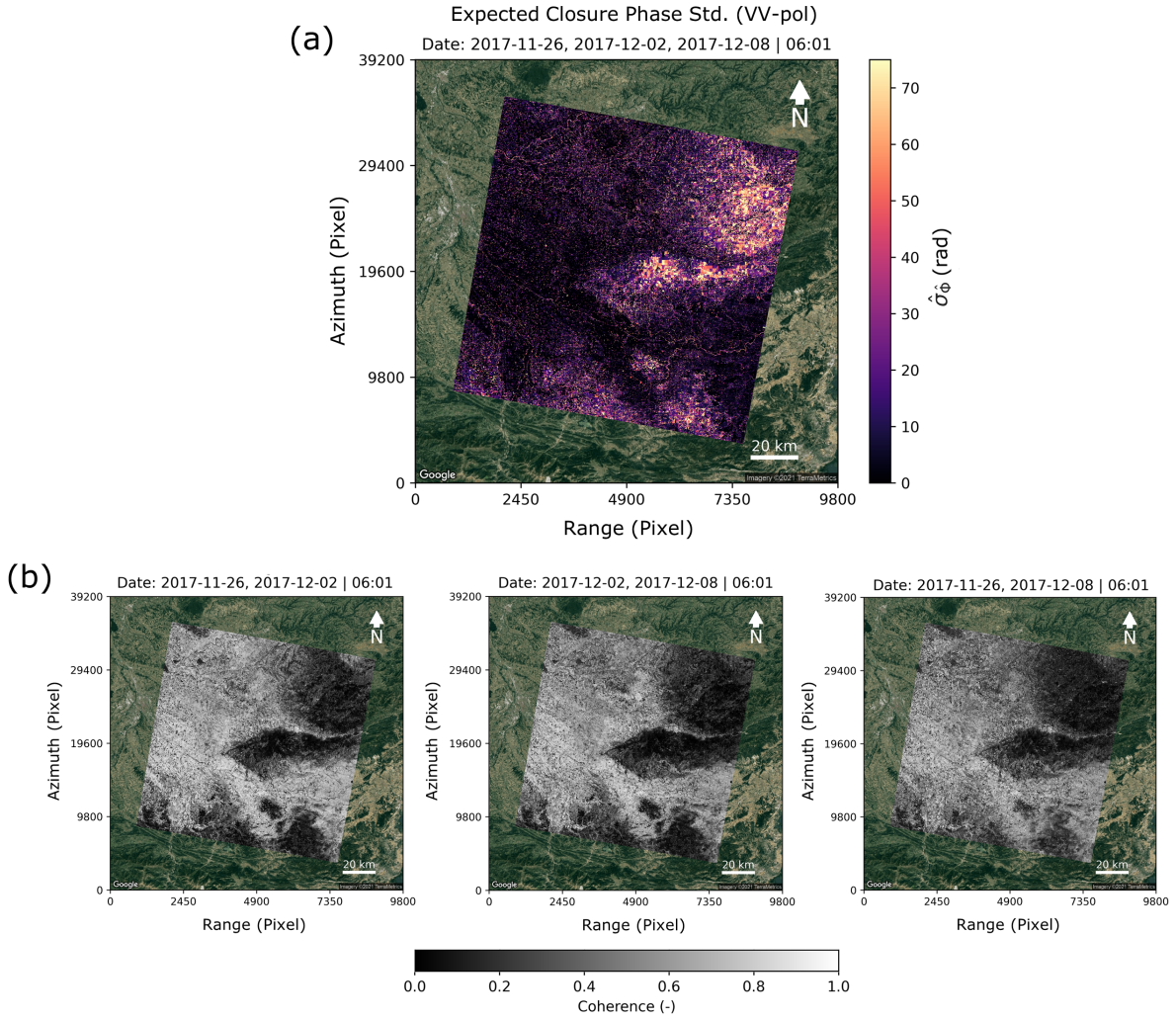


Figure 4.3: Top: Expected Closure Phase Standard Deviation ($\hat{\sigma}_{\hat{\Phi}}$) for a region in Southern France. Bottom: Coherence magnitude maps that where required as an input to the significance test in order to estimate $\hat{\sigma}_{\hat{\Phi}}$. Areas of low coherence are associated with higher estimates for $\hat{\sigma}_{\hat{\Phi}}$ and vice versa.

on the coherence and associated phase noise is reflected in $\bar{\Phi}_{i,j,k}$: For the forested regions it is argued that a large portion of $\bar{\Phi}_{i,j,k}$ is attributable to phase noise induced by volume scattering within the vegetation canopy which applies to those regions with extensive small-scale variability of large positive and negative closures. This is captured in low coherences and large $\hat{\sigma}_{\hat{\Phi}}$ estimates. The consistently low significance ratios suggest weak evidence of a (large) geophysical component that contributed to this phase triplet mismatch. Crop areas showed lower $\bar{\Phi}_{i,j,k}$ which one could prematurely conclude that a geophysical component is small, however, the significance ratio revealed that the phase closure is rather large with respect to $\hat{\sigma}_{\hat{\Phi}}$. Due to the absence of crops in November and December, the radar primarily senses the ground surface, hence the phase closures most likely indicate local soil moisture changes.

Obviously those maps only present a single triple product of the respective observable, however, the visual similarity with land cover can not be denied and suggests first evidence of a geophysical origin of phase closures that has not yet been explored with respect to different land cover types. The number of looks (per field) as an essential parameter for the significance test, indicates the degree of spatial averaging for fields of closure phase and will be subject of the following section.

4.2. Effects of Spatial Aggregation

The existence of phase closure errors only results from spatially averaged interferograms. Estimates of the expected phase closure standard deviation ($\hat{\sigma}_{\hat{\Phi}}$) are correlated with the number of looks per field, i.e. an increasing degree of multilooking (spatial averaging) is associated with lower $\hat{\sigma}_{\hat{\Phi}}$ and vice versa (see Fig. 2.2). In this sub-chapter, the sensitivity of significance ratio to different degrees of spatial averaging is analysed. Further, the effect on the linear similarity between individual time series of significance ratio is assessed by comparing Pearson's *R* as a function of multilooking. The assessment was carried out for significance ratios and the observed closure phases in comparison, for different land cover types and both VV and VH polarized data. This sub-chapter aims to address research objective III.

Fig. 4.4 presents three maps of significance ratio (b) for the research area (a), filtered to only show a certain range of multilooking. As the pixels are aggregated based on land cover fields with different sizes, the number of looks across the maps varies spatially. The map to the left is filtered for looks between 4 and 250, the middle map shows 250-1,000 looks and the map to the right shows 1,000-10,000 looks. The grid overlay of 1x1 km that is applied to the land cover maps before aggregating the pixel values per land cover field, bounded the range for the number of looks (per field) to a minimum of 4 (20 x 20 m, single pixel) and a maximum of 10,000 (1 x 1 km). That means that the outline of the grid cells is visible when zoomed-in on an area with relatively homogenous land cover as shown in Fig. 4.4 (c). From the map comparison in Fig. 4.4 one can derive at least three important takeaways:

1. The magnitudes of the significance ratio are the lowest for the map on the left with the smallest degree of multilooking. The magnitudes increase noticeably for the medium range multilooking and are the greatest for the map on the right with looks up to 10,000. This clearly indicates the positive influence that further spatial aggregation has on increasing the signal-to-noise ratio of the geophysical closure phase.

2. The consistency of the spatial signatures suffers for low degrees of spatial averaging which is partly due to the lower values of the significance ratio. Distinct patterns are better visible as multilooking increased to 250-1,000 looks, while the map with the highest looks shows very distinct spatial patterns.

3. The zoom-in on the significance ratio (c) shows different field sizes and highlights the effect and importance of splitting larger polygons into separate fields. The crop area in the left part of the zoom-in indicates consistently positive significance ratios (red pattern) even though the fields are all separated and rather small. This is a good indication of a geophysical origin which caused a similar phase closure over the whole area. The forested area in the center and to the right of (c), however, consists of only a few (multi-)polygons (as shown in the land cover map) but are bounded to 1x1 km fields before averaging spatially. If the pixel values were aggregated over the whole area, this would have given the impression that the area has a consistent phase closure signal. Looking at the inconsistency of the significance ratios over the forested area implies that this is clearly not the case.

The map comparison could show that increased spatial averaging has a positive effect on both increasing the significance ratio of phase closures and further highlighting its spatial signatures. This could be valuable for distinguishing between different land cover types using the values of the significance ratio. To further investigate its the spatial characteristics, it is beneficial to apply a filter to those

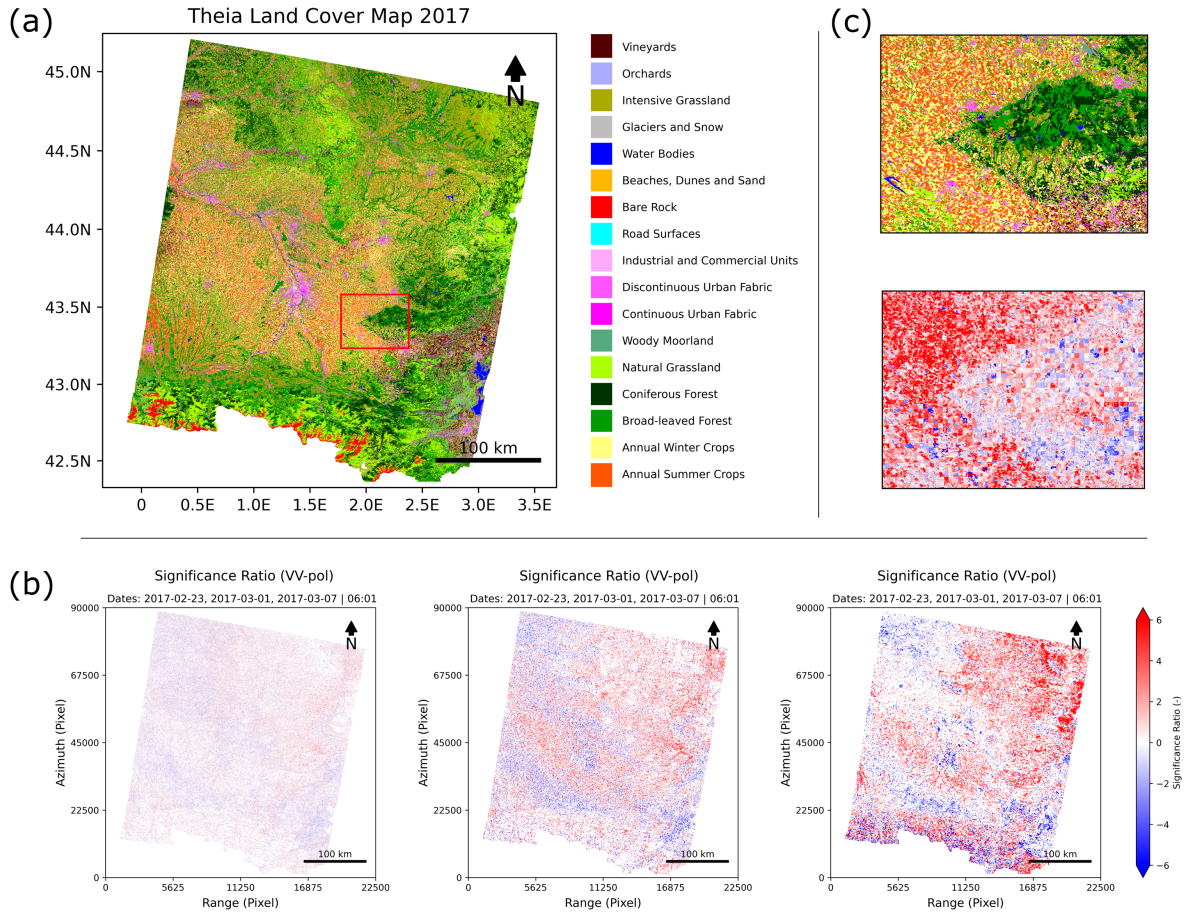


Figure 4.4: (a) Land cover map of the research area. (b) Three maps of significance ratio over the same area filtered for different ranges of multilooks. The map on the left shows all land cover polygons with 4-250 looks, the middle is filtered for 250-1,000 looks, and the map to the right shows all fields with 1,000-10,000 looks. (c) Zoom-in on an area with homogenous land cover.

maps and mask areas with significance ratios below a certain threshold or determine a minimum degree of multilooking before applying the significance test. The following sections will further elaborate on that.

4.2.1. Root Mean Square as a Function of Spatial Averaging

The results from the previous section indicated a positive relation between spatial averaging and the values of the significance ratio. The effect of spatial averaging on both the observed closure phase $\bar{\Phi}_{i,j,k}$ and the significance ratio $\Psi_{i,j,k}$ values with respect to different land covers and polarizations is assessed by comparing the evolution of root mean squares (RMS) as a function of multilooks (Fig. 4.5 and 4.6). The goal of this comparison was to indicate which land covers are associated with the largest significance ratios, how that relates to the observed phase closures and if the impact depends on different polarizations. The effect that spatial averaging has on the observed phase closures is introduced first, also to better understand the signal-to-noise ratio over different land covers. In order to ensure a minimum number of fields per multilook, the fields are binned for bin sizes of 100 looks. The RMS is computed over the time series within each look-interval and averaged per interval afterwards.

Observed Phase Closures

The RMS of the observed phase closure in Fig. 4.5 shows the greatest closure phases for looks $< 2{,}000$ and a relatively steady evolution thereafter as the number of looks increases per bin. For crops (summer/winter crops, intensive grassland), the quadratic means are smaller and the convergence to a steady trajectory is slower (at ∼2,000 looks) than compared to natural vegetation including grassland and forests (at ∼1,000 looks). In general, the RMS are greater for VH-pol closure phases for all land covers than compared to VV-pol. The greatest RMS obtain the forest types with an average of about 50 to 60 degrees for VV and VH respectively. Generally higher phase levels over forests were also observed in Fig. 4.2, and were presumably attributed to vegetation induced phase noise. The lowest RMS receive the crop types with around 20 degrees. Natural grassland has an average RMS of about 30 to 40 (VV/VH) that converges comparably fast with multilooking as well. Above ∼2,000 looks, there is seemingly no effect from further multilooking on the phase values across all land covers.

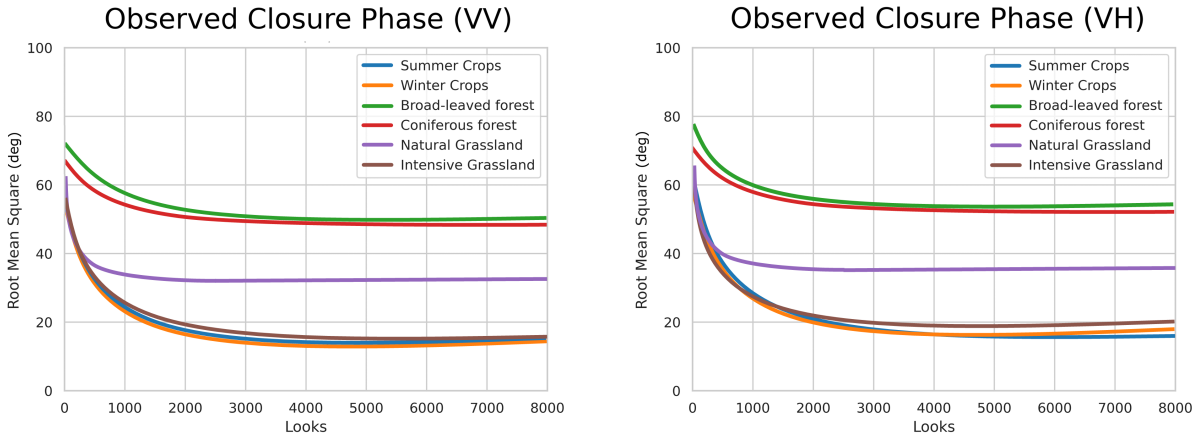


Figure 4.5: Root Mean Square (RMS) computed from time series of closure phase for different land cover types as a function of multilooks. The average RMS from all time series is illustrated per multilook with respect to different land cover types.

Significance Ratio

The trajectories of the RMS of the significance ratios in Fig. 4.6 show different characteristics with respect to polarization and land cover. If the phase closures were fully justified by decorrelation phase noise, the RMS would be 1, i.e. $\hat{\sigma}_{\Phi}$ and the observed closure phase were equal, which is indicated by the dotted black line. RMS > 1 point to a geophysical phase closure.

For the significance ratios, multilooking has a very significant effect on increasing the RMS for all land cover types. The minimum RMS averaged over time series up to 100 looks is at ∼1.2/1.4 for VH and VV-pol respectively, showing that even a small multilook can increase the SNR of geophysical phase closures. The steepest increase is noticeable up to 1,000 looks, thereafter the individual curves increase relatively steady with the degree of spatial aggregation. The highest RMS are obtained for crops and natural grassland; intensive grassland and forest vegetation show the lowest sensitivity to spatial averaging.

While the two forest types and natural grasslands would have fairly similar land cover (and therefore coherence magnitudes) throughout the year respectively, the crops (winter and summer) and intensive grassland (to a lesser extent) are characterised by seasonal vegetation cover, i.e. they are bare in winter and vegetated in spring, summer and fall respectively. The different RMS trajectories can be partly explained by this seasonality, i.e. the high RMS of significance ratio over the crop types are due to the periods, where the soils are bare. Natural grasslands, in comparison, do not have a distinct

period of bare soils from harvest but are characterised by the lowest vegetation throughout the year in accordance with Theia's land cover classification, meaning that the radar primarily senses the ground as well. The sensitivity to soil moisture variabilities is higher over sparsely vegetated or bare soils, therefore, the RMS of significance ratio over winter and summer crops, and natural grassland are all comparably high. Vegetation cover, on the other hand, introduces volume scattering associated with an increase in phase noise and hence lower significance ratios. The forest types and intensive grassland are characterised by higher vegetation cover and hence generally lower significance ratios.

Comparing VV-pol versus VH-pol, the trend and increase of the RMS for forest types and intensive grassland is similar while there is a noticeable gap for VV-pol phase closures over crops and natural grassland which is not as pronounced for VH-pol. The higher sensitivity of VV-pol to averaging over those areas, i.e. natural grassland and crops, shows that the geophysical signal in the VV-pol significance ratio is greater (and the phase noise likely smaller). Due to the propagation advantage of VV-pol to reach the ground surface (Flores-Anderson et al., 2019), the gap between the VV and VH trajectories for crops and natural grassland are attributable to greater significance ratios over low vegetated land cover types. During those times, when soils are sparsely vegetated,the radar is more sensitive to soil moisture variabilities which primarily contributed to the geophysical phase closures. For VV-pol, the RMS over natural grasslands and winter crops increased by about 4.5 times at 8000 looks, while the forests types and intensive grasslands increased by about 3 times. The sensitivity to spatial averaging was lower for VH-pol, yet the RMS over winter crops and natural grasslands increased by about 3 times, and by about 1.6-2.3 times over forests and intensive grasslands. The fact that forests in particular obtain relatively low RMS significance ratios first indicates that phase closures are potentially more sensitive to soil moisture dynamics rather than moisture changes in the leaves and trunks of trees as speculated by De Zan and Gomba (2018). When comparing the observed closure phase and the significance ratio, the generally higher phase closures $\bar{\Phi}_{i,j,k}$ over forests obtained the lowest significance ratios while the lowest phase closures $\bar{\Phi}_{i,j,k}$ from crops received the highest significance ratios. The level of phase noise seems to be significantly higher over densely-vegetated land covers, then it is over crops and natural grasslands. Volume decorrelation which is linked to forests and vegetation of a few meters height leads to enhanced phase noise (Martone et al., 2016) which explains the large RMS for $\bar{\Phi}_{i,j,k}$ and rather low RMS for the significance ratios over broad-leaved/coniferous forests and intensive grasslands. The comparison further showed that besides the vegetation type and its density, the polarization has a significant influence on the sensitivity to spatial averaging. Considering that the significance ratios seem to further increase when multilooking is $>$ 8000 could be a motivation to explore the sensitivity to further spatial aggregation.

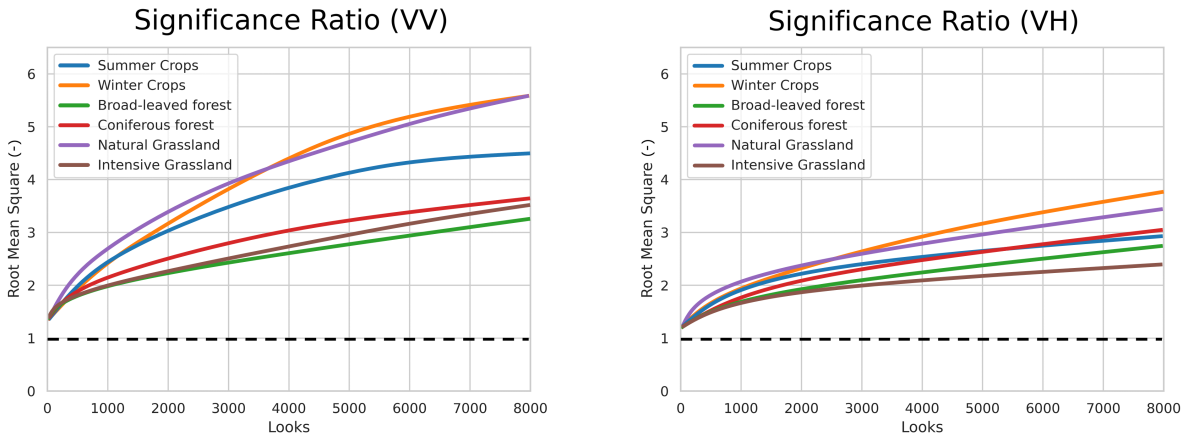


Figure 4.6: Root Mean Square (RMS) computed from time series of significance ratio for different land cover types as a function of multilooks. The average RMS from all time series is illustrated per multilook with respect to different land cover types.

Spatial Comparison of Root Mean Squares

The map comparison for 2017, 2018 and 2019 in Fig. 4.7 shows the RMS of the significance ratio computed over the time series of each polygon-field. The presented scene covers the full research, the same area as already illustrated in Fig. 4.4. Areas with < 1000 looks are masked to accentuate spatial signatures. The boxplots show the RMS values per land cover type to facilitate a comparison with the underlying land cover in Fig. 4.4.

The maps of RMS of the significance ratio capitalize on the distinct differences in RMS for VV-pol significance ratios over different land cover types. The comparison shows that the spatial signatures of RMS are spatially consistent between the years, and consistent with respect to land cover, suggesting that the annual magnitudes of significance ratio can be exploited for land cover mapping or to identify annual changes in land cover. The highest RMS (purple) are obtained by natural grasslands and crops (purple/blue), green patches mostly indicate intensive grassland and yellow-orange is associated with forest vegetation which is consistent with the land cover map. Note that the crop area in the center of the map (purple) appears similar for 2018 and 2019, but different for 2017 which is owed to the updated land cover nomenclature. The effect of spatial averaging is very distinct here, as the land cover polygons for 2017 (summer and winter crops) are generally larger than the further divided nomenclature (2018 and 2019) for individual crop types. Therefore, the aggregated crop-fields for summer and winter crops generally obtained greater RMS values due to the sensitivity of significance ratio to multilooking.

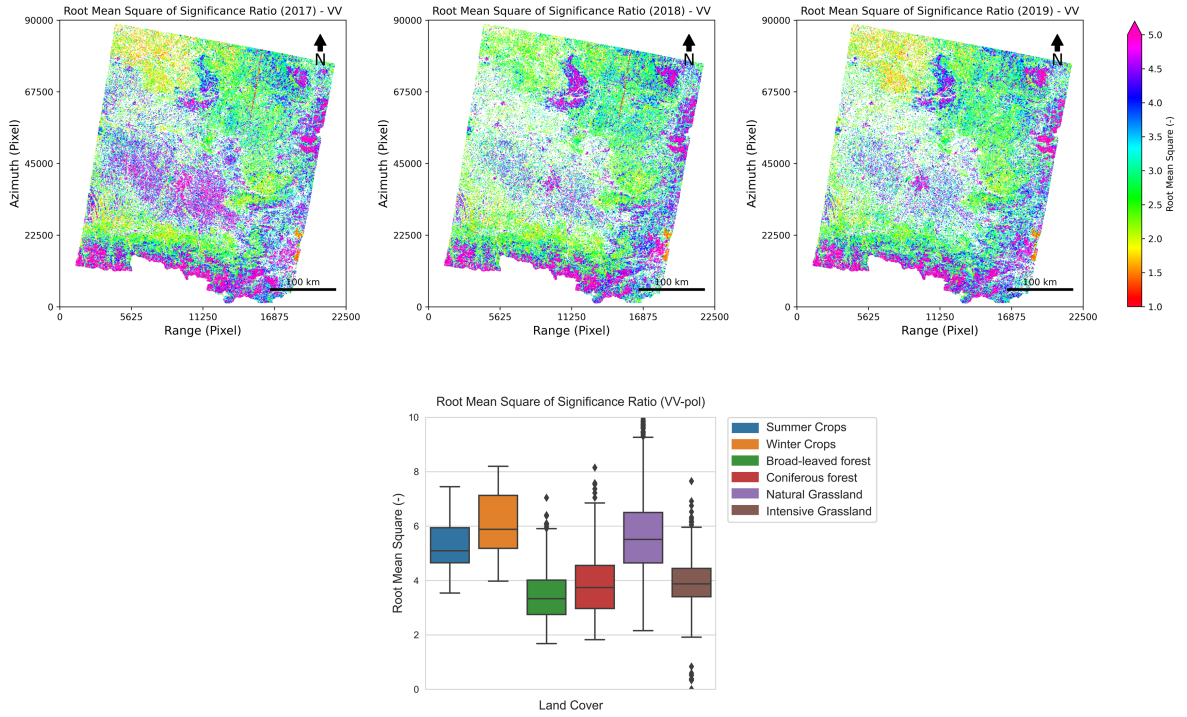


Figure 4.7: Root Mean Square (RMS) of significance ratio over the whole research area for 2017, 2018 and 2019 time series. Land cover fields smaller than 1000 looks are masked to accentuate the spatial signatures. The distinct patterns show that the RMS of significance ratio area consistent with land cover. The boxplot facilitates a comparison with the underlying land cover.

4.2.2. Pearson's *R* as a Function of Spatial Averaging

It was shown that the degree of spatial aggregation has a significant effect on the magnitudes of $\Psi_{i,j,k}$ but that does not demonstrate that it has any effect on the correlation between individual time series. An assessment of the similarity between spatially aggregated fields was performed by calculating Pearson's *R* for time series of significance ratio with a maximum distance of 25 km to each other in order to mitigate possible effects caused by spatial distance. The looks were binned logarithmically in increments of 100 looks up to 1,000, and in increments of 1,000 till 10,000 looks. The reason was to achieve a better illustration of the effect for smaller multilooking, which was possible because smaller fields outnumbered the larger ones. Histograms of the sampling of time series per look-bin are provided in the Appendix (Fig. A.1, A.2, A.3) for each land cover type. The boxplots of Pearson's *R* in Fig. 4.8 to 4.10 illustrate different land cover types with cross-polarized (VH) significance ratio on the right and co-polarized (VV) significance ratio on the left respectively. An increasing, positive effect of the *R* boxplot trajectories for significance ratio would generally imply that the impact of a geophysical cause, e.g. soil moisture variabilities, is consistent on phase closures.

Broad-leaved and Coniferous Forests

Fig. 4.8 show Pearson's *R* boxplots for broad-leaved and coniferous forest respectively. The trajectory of the boxplots indicate that an increase in spatial aggregation seems to have very limited effects on the correlation between time series for both forest types, broad-leaved forest in particular. The median of Pearson's *R* only slightly increases for low looks to 10,000 looks from 0 to about 0.12 for broad-leaved forest and 0.20 for coniferous forest. There is no noticeable difference between VV and VH significance ratio for both forest types.

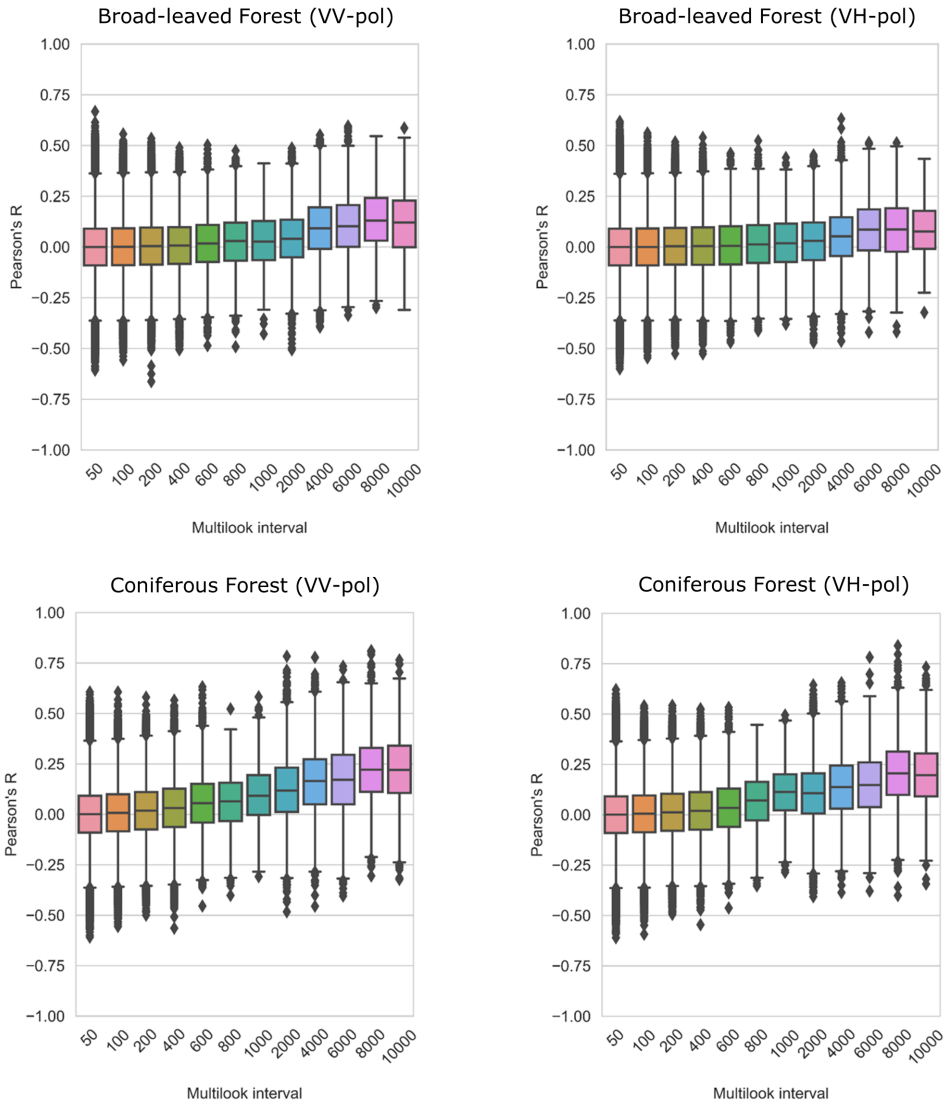


Figure 4.8: Boxplots of Pearson's R for broad-leaved and coniferous forest $\Psi_{i,j,k}$ time series as a function of multilooking. The line shows the median value. The bottom and top of the box the 25th and 75th percentiles, respectively. The upper tail is the 75th percentile plus 1.5 times the difference between the median and 75th percentile. The lower tail is the 25th percentile minus 1.5 times the difference between the median and 25th percentile. Outliers are plotted as diamonds.

Natural Grasslands

Fig. 4.9 shows box plots for natural grassland. Since it is often situated in rough areas with uneven terrain and steep slopes, the correlation between time series could be affected depending on where the data is derived from. Masking regions in the Pyrenees and Massif Central from the data, three regions that are relatively homogeneously covered with natural grassland are chosen for this analysis (Fig. 4.9).

An increase in multilooking has a significant effect on the correlation between grassland time series. R increases from 0 at < 100 looks up to 0.72 at 10,000 looks for both VV and VH polarized significance ratio with no noticeable difference between the boxplot trajectories. A higher count of outliers can be found for lower look-bins which can be explained from the higher number of time series with lower multilooks as presented in Fig. A.2 in the Appendix. The increase in correlation between the time series is greater over the first 1,000 multilooks compared to multilooking between 1,000 and 10,000, which seems to relate to the steeper increase in significance ratio magnitudes over the first 1,000 looks as

observed in the previous section.

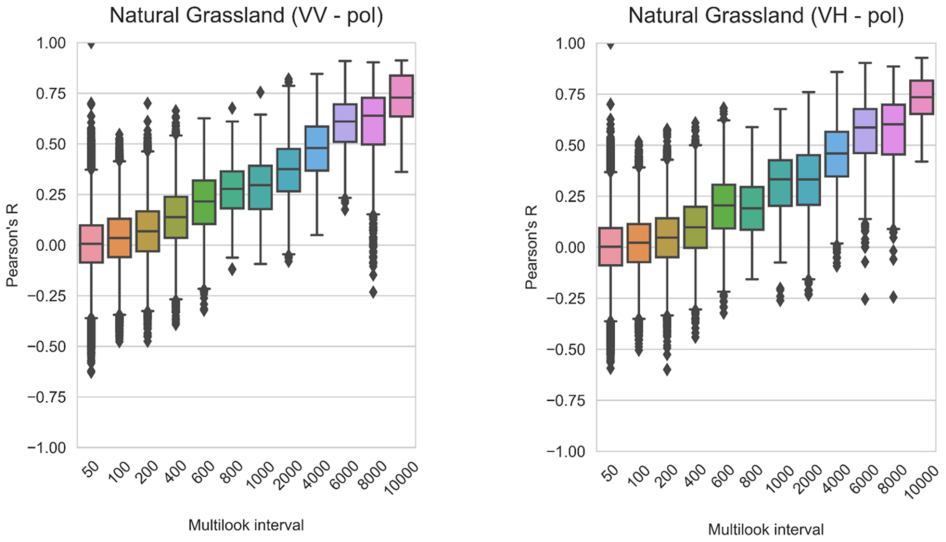


Figure 4.9: Boxplots of Pearson's R for natural grassland $\Psi_{i,j,k}$ time series as a function of multilooking. The boxplot evolution indicates a significant sensitivity of multilooking on the similarity between individual time series. The line shows the median value. The bottom and top of the box the 25th and 75th percentiles, respectively. The upper tail is the 75th percentile plus 1.5 times the difference between the median and 75th percentile. The lower tail is the 25th percentile minus 1.5 times the difference between the median and 25th percentile. Outliers are plotted as diamonds.

Agricultural Crops

Fig. 4.10 differentiates time series for fields of summer, winter and perennial crops (intensive grassland). The significance ratios from fields of summer crops show an increase in correlation, while the median R reaches 0.46 for VV and 0.52 for VH polarization at 8,000 looks. The trend is similar for winter crops with an increase to 0.51 for both polarizations. Considerable differences between VV- and VH-pol are not noticeable for both winter and summer crops. The vegetation characteristics of the two cover types are different, as winter crops primarily comprise of straw cereals and summer crops are primarily corn crops over southern France. However, the impact of increasing multilooking is very similar for both cover types.

For intensive grassland, R increases from 0 for $<$ 100 looks to 0.68 (VV) and 0.43 (VH) at 10,000 looks. The boxplot trajectories for VH and VV in comparison show a noticeable difference. While the sensitivity of VV-pol significance ratio leads to a relatively steady increase in R, VH-pol for the same fields seems to be relatively insensitive to spatial averaging up to 1,000 looks. Drawing definite conclusions from this observation is non-trivial, as intensive grassland as a land cover class comprises different types of vegetation. Therefore, the lower sensitivity of VH to spatial aggregation is most likely owed to the fact that those areas are covered with different types of vegetation all year which leads to non-similar time series of significance ratio. Both, the sensitivity of VV- and VH-pol closure phase to soil moisture are attenuated by variabilities in the vegetation cover, but presumably less pronounced for VV-pol due to its better capability to monitor soil.

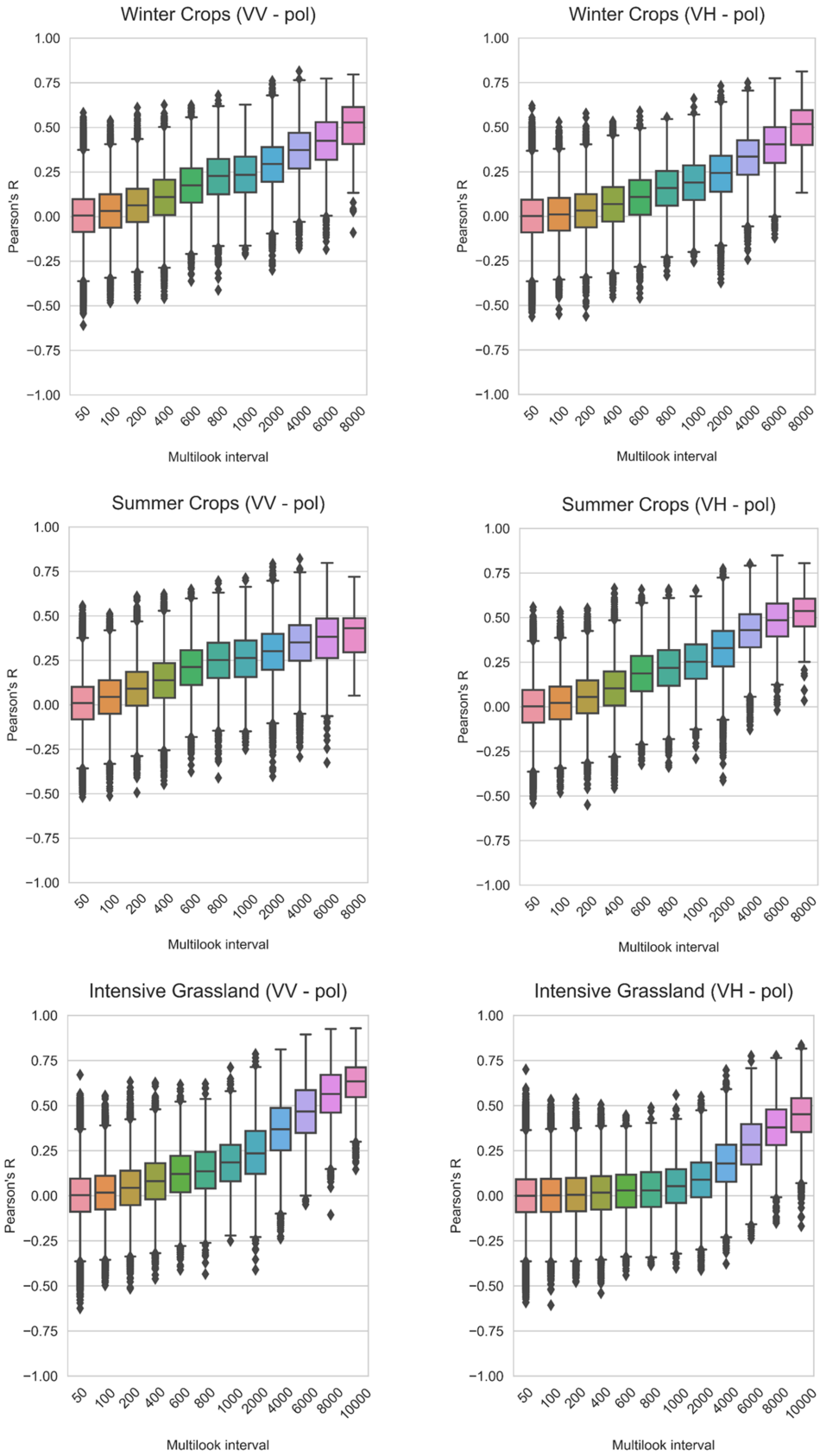


Figure 4.10: Boxplots of Pearson's *R* for winter crops (left), summer crops (right) and intensive grassland (bottom) $\Psi_{i,j,k}$ time series as a function of multilooking. The boxplot evolutions indicate a good sensitivity to spatial aggregation as the degree of multilooking increases. The line shows the median value. The bottom and top of the box the 25th and 75th percentiles, respectively. The upper tail is the 75th percentile plus 1.5 times the difference between the median and 75th percentile. The lower tail is the 25th percentile minus 1.5 times the difference between the median and 25th percentile. Outliers are plotted as diamonds.

Differences between the Land Covers

The comparison of Pearson's *R* as a function of multilooking revealed that significance ratios over different land cover types are differently sensitive to multilooking. A good sensitivity was found for the grassland land cover types, intensive and natural, and agricultural crops. On the other hand, a minor effect of spatial aggregation was observed over forest vegetation. The most obvious difference between forests and the other land covers is the above ground biomass. If the radar primarily senses the ground surface in the absence of vegetation, the influence of soil moisture dynamics on the signal would dominate. Since soil moisture is spatially variable, one would expect somewhat similar ground properties for adjacent fields. For low vegetated land cover types such as the annual crops (outside crop cycles) and natural grassland, it can be assumed that a large fraction of the phase closure is caused by soil moisture changes. This would explain the relatively large correlation coefficients, in particular as the degree of multilooking increased. Over forests, however, the radar primarily senses the forest canopy and overstory layer rather than the ground surface, in particular at C-Band. The associated volume scattering affects the coherence, and adds to the noise component in phase closures which impacts the consistency of a geophysical signal. Variations of the dielectric constant linked to tree moisture as De Zan and Gomba (2018) suggested might be somewhat consistent for close fields, but the correlation between time series over forests was found to be very low even for spatial averaging over $>$ 9,000 looks.

Interestingly, the trend of the boxplot trajectories, similarly to the RMS comparison, indicate that even further spatial averaging would have a positive effect on the correlation between individual time series, and could therefore further emphasize the temporal characteristics of significance ratio, in particular over low vegetated land covers.

4.3. Closure Phase Characteristics in Time and Space

The previous section has shown that the SNR of the geophysical closure phase increases with the degree of spatial averaging. The objective of this section was to highlight the temporal characteristics of significance ratio with respect to different land cover types. An overview of eight maps in Fig. 4.11 intends to illustrate differences of the spatial signatures of significance ratio obtained during different times of the year. Presented are both significance ratios from VV-pol and VH-pol data. This sub-chapter aims to address research objective IV.

Firstly, it is noticeable that the magnitudes of significance ratio for VV-pol data are greater than VH-pol while the spatial signatures are similar across all maps. The examples of significance ratio during the spring and summer months show that the ratios are generally lower than compared to winter and fall. Looking at the map for summer in particular, one can see that the significance ratios are lower in magnitude and a spatial consistency with land cover is difficult to see. The scene looks noisy except for areas of natural grasslands in the east of the map with greater magnitudes of significance ratio. The winter and fall maps, however, are characterized by great significance ratios and distinct spatial patterns. The spring-map already indicates a fading of spatial signatures, in particular over agricultural areas in the center of the map. The lower magnitudes of significance ratio in the summer and spring months indicate that increased vegetation cover due to crop growth and leaf-out in forest vegetation increases the noise component in phase closures significantly, leading to lower significance ratio. Another noticeable feature are the two separate patterns over natural grasslands in the center-north of the fall-map, indicated by a red circle, which likely point to rain cells associated with wetting of the area (blue in between). This is important to notice as rain events are the primary driver of soil

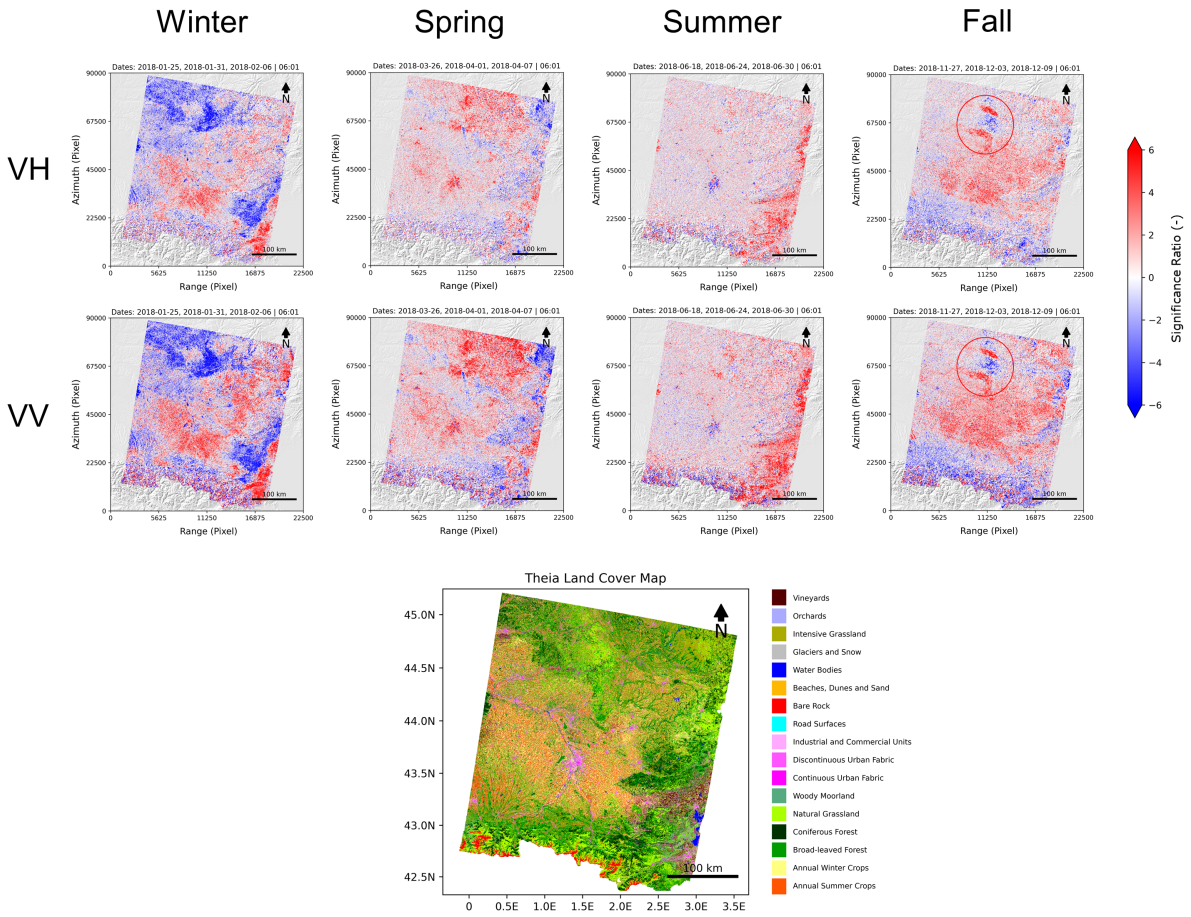


Figure 4.11: Comparison of eight maps showing the annual differences in significance ratio derived from VV-pol and VH-pol.

moisture changes, hence the closure phase will be influenced by such events, which likely shows in time series of significance ratio.

4.3.1. Time Series Comparison

The significance ratio time series in Fig. 4.13 to 4.16 present 100 time series for 2017, 2018 and 2019, illustrated as scatterers, in comparison with coherence, and LAI, as a function of different land covers. The LAI as a measure of vegetation density shows the annual changes in vegetation while the coherence is used to indicate the quality of the interferometric phase during periods of increased vegetation cover. The LAI and coherence magnitudes for VV- and VH-pol are plotted to the right of each figure for the same 100 fields. The error bars in the coherence time series indicate the standard deviation of the values per date. Areas with difficult terrain (Pyrenees in the south and a small area over the Central Massif) are masked from the data. Since the results in the previous section suggested that the degree of spatial averaging has a great effect on the significance ratio, the 100 polygons with the largest multilook (9,500 to 10,000 looks) are selected from different regions. This also allows to adequately compare between land cover types, without having to consider different levels of spatial aggregation. Empty columns indicate dates where no acquisition was collected by Sentinel-1. The reasons to illustrate the time series altogether are the following:

Firstly, to visualize the significance ratios together might dismiss the individuality of each time series, but the central idea here was to potentially show consistency of significance peaks for certain

dates that might reveal a sensitivity to vegetation dynamics despite the spatial distance between the polygons. Secondly, to assess a potential similarity between 2017, 2018 and 2019 time series that is consistent with dynamics in coherence and LAI, and thirdly to compare the general temporal characteristics between different land cover types and polarizations, also with respect to quantity and frequency of peaks of significance ratio.

This section is subdivided for the three main land cover classes *natural grassland*, *agricultural crops* including *intensive grassland*, and *broad-leaved forest*. Those land cover types are chosen with respect to characteristic properties, including annual phenology, vegetation type and density.

Natural Grasslands

Fig. 4.12 presents the significance ratios for 2017, 2018 and 2019 time series over areas of natural grasslands. The peaks of the significance ratios suggest a relatively equal distribution throughout the respective year, while VV-pol significance ratios are consistently greater. Most notable are a few larger numbers of very high significance ratios up to 20, which appear a few times for all years. Considering vegetation cover, the majority of grassland time series has a leaf area index between 0.5-2 throughout the year with a peak in June for all years (LAI$\approx$2). This month is associated with decorrelation, whereas coherence magnitudes are at an average of 0.3-0.35 for VH and 0.46-0.51 for VV. Due to the relative low vegetation throughout the year, the significance peaks largely represent the dielectric variabilities from soil moisture changes rather than phase contributions from volume scattering within the vegetation canopy. The majority of grassland areas, where those time series are derived from, are in the northern part of the research area as well as a few hundred kilometers in the eastern part along the coast. With a great sensitivity to soil moisture, single large peaks most likely represent rain events in those regions. The 2017 and 2018 rain data in Fig. C.2 and Fig. C.3 for the stations *Perpignan* and *Millau* (see Fig. C.1) in the Appendix indicate a few heavy rain events that coincide with dates of the most significance peaks, e.g. noticeable between January and March, the rain event in mid-October for 2017 and the large rain events between mid-October and December in 2018. Of course one can not relate each individual rain peak to a single cluster of significance ratios, however, the distribution of the peaks generally indicate a pattern similar to rain events throughout the year, as a number of high significance ratios is followed by periods with lower significance ratios. The consistently low vegetation cover makes areas of natural grassland a good candidate to explore the sensitivity of closure phase to soil moisture, comparable with croplands after harvest.

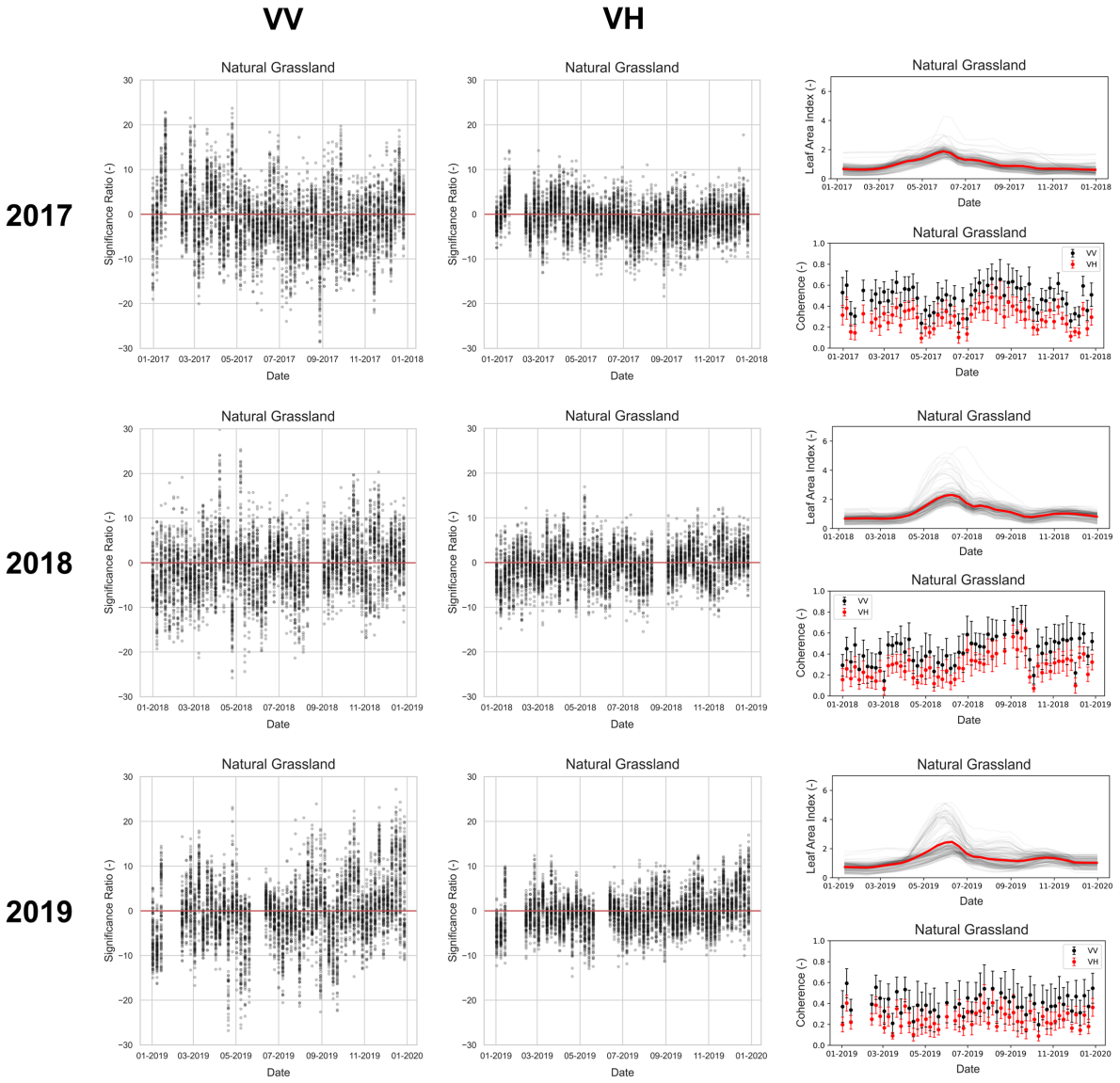


Figure 4.12: Natural grassland time series plots of $\Psi_{i,j,k}$ for 100 polygons with looks between 9,500 and 10,000 for 2017, 2018 and 2019. The plots on the left show the significance ratios for VV-pol, the center plots for VH-pol. The plots on the right indicate the temporal variability of LAI (top) and coherence magnitude (bottom) for 2017, 2018 and 2019 respectively.

Agricultural Crops

Fig. 4.13 presents the land cover type winter crops (2017) which is compared to straw cereal (2018 and 2019) since the land cover maps from 2018 and onward divide winter and summer crops in individual representatives. Straw cereal has by far the largest share in the group of winter crops, and is therefore picked for the comparison.

The time series from all years show a large number of significance ratios (up to 30) between July and December while the ratios are mostly $<$ |5| between March and July. Compared to the LAI and coherence time series, this period is associated with an increase in LAI and low coherence magnitudes which is indicative for the phenological cycle of winter crops. Leaf emergence of winter crops starts in early March and reaches its maximum mid-April to early May (LAI-peak) while the harvest takes place between June and July. Following the crop growth period, the number of peaks abruptly increases and remains large for the rest of both years from July onward, however the significance ratios are not

consistently greater than during the crop cycle period, only sometimes. Comparing both polarisations, VV-pol phase closures show a larger number and generally greater significance than VH-pol. Noticeable is the fact that the majority of large ratios are positive. This observation holds true for 2017, 2018 and 2019 significance ratios.

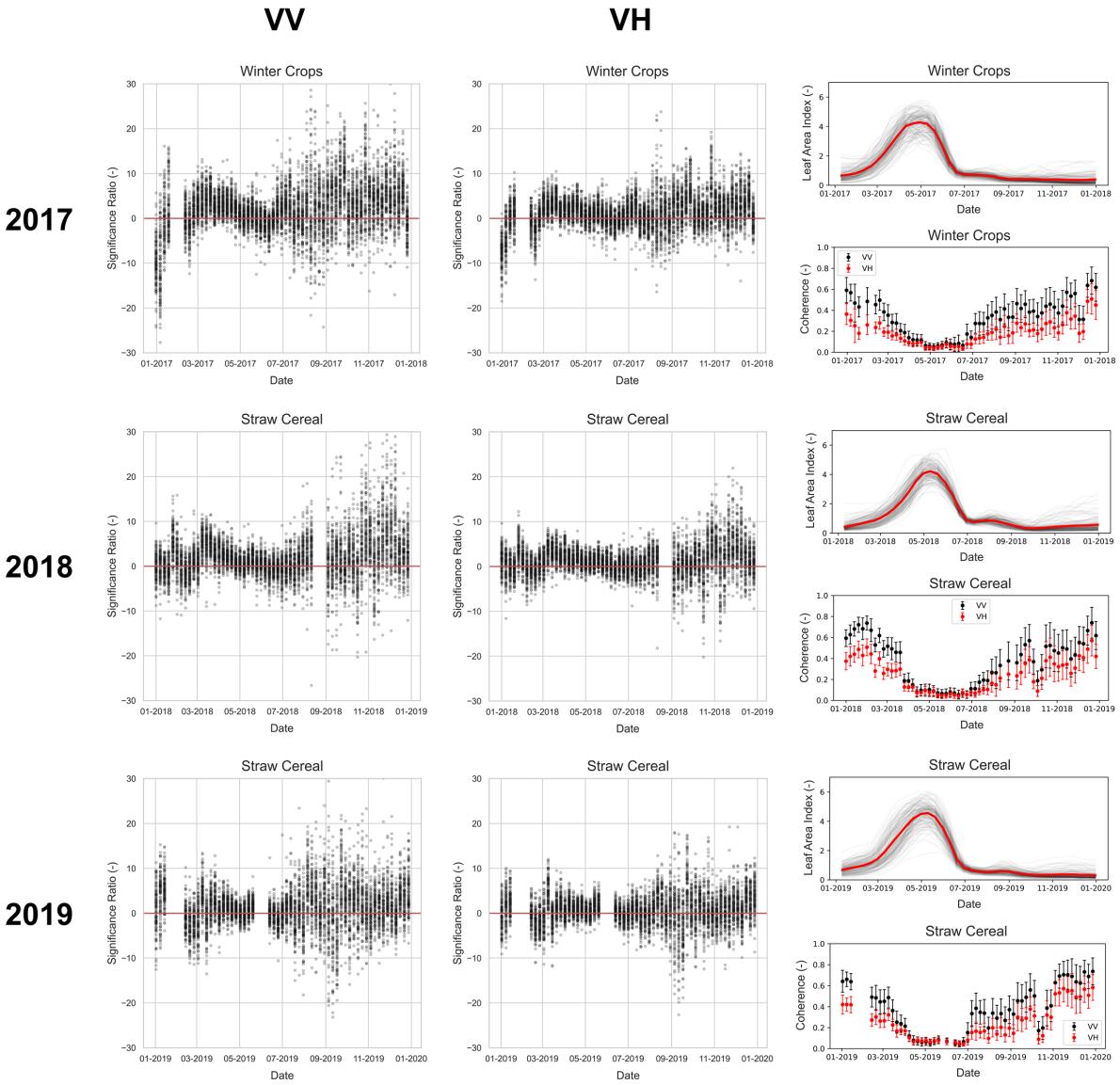


Figure 4.13: Same setting as in Fig. 4.12 but for winter crops (2017) and straw cereal (2018 and 2019).

In Fig. 4.14 summer crops are compared with Corn, the most dominant representative for summer crops from the 2018 and 2019 land cover classification. Over summer crops, the greatest significance ratio are between January-June and October-December for time series of all years while most significance ratios not exceed |5| between June and October. In contrast to winter crops, summer crops have their vegetative stages between late-May and June, indicated by increasing LAI and decreasing coherence magnitudes, and are harvested between September and October (decreasing LAI, increasing coherence). Similarly to winter crops, the period following the harvest event is associated with large ratios for all years. All 2017, 2018 and 2019 time series show the highest count and largest peaks for the period between January and May before the crop emergence begins.

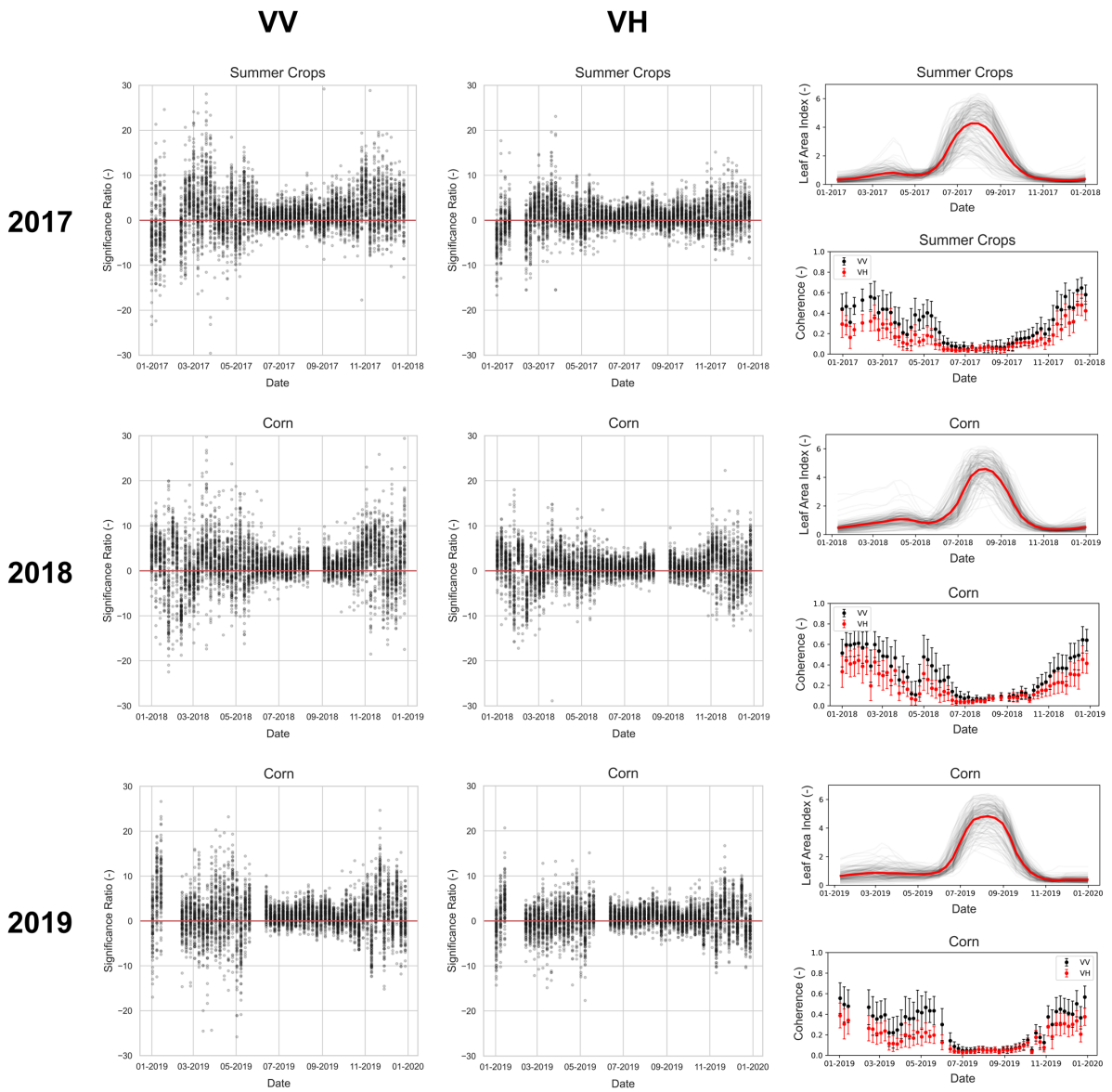


Figure 4.14: Same setting as in Fig. 4.12 but for summer crops (2017) and corn (2018 and 2019).

Intensive Grassland in Fig. 4.15 shows different temporal characteristics between 2017, 2018 and 2019 respectively. Comparing the evolution of LAI between the years, one can notice that 2017 indicates two distinctive growth periods while this is not as pronounced for 2018 and 2019. Consequently, major crop growth periods coincide with low coherence magnitudes between March and November. Related to the temporal characteristics of coherence and LAI is the occurrence of significance ratios: High peaks in January to April for 2017, 2018 and 2019 are observed when the LAI is < 1 and the coherence is between 0.3 and 0.5 for VH and VV respectively. The LAI remains > 1 over 2017, for 2018 the LAI decreases to about 1 at the end of September, and for 2019 remains > 1 which is in keeping with only a few significance peaks for 2017 and a noticeably larger number for 2018 and 2019 between September and December. Differences in the LAI trajectory for 2017, 2018/2019 time series over intensive grassland can be explained by different crop types with different growth periods that perennial crops feature. It seems that 2017 involves at least two crop types with distinctive growth cycles while one crop type dominates for 2018. Because of that the second half of 2017 has significantly larger LAI,

associated with low coherences, which led to a lower number of significance ratios when compared to 2018 and 2019. The fact that the LAI is consistently > 1 (excluding February and March) compared to summer and winter crops is because perennial crops do not need to be replanted each year, they automatically grow back after harvest. Hence a delimited period of “high peaks” and “low peaks” of significance ratio is absent. One can also notice that the difference between VV-pol and VH-pol coherence is mostly < 0.1 for those periods while the different polarisations have a strong effect on the closure phase significance. There are only a very few dates that caused high significance peaks for VH-pol, about 1% with $> |10|$ magnitude, compared to VV-pol.

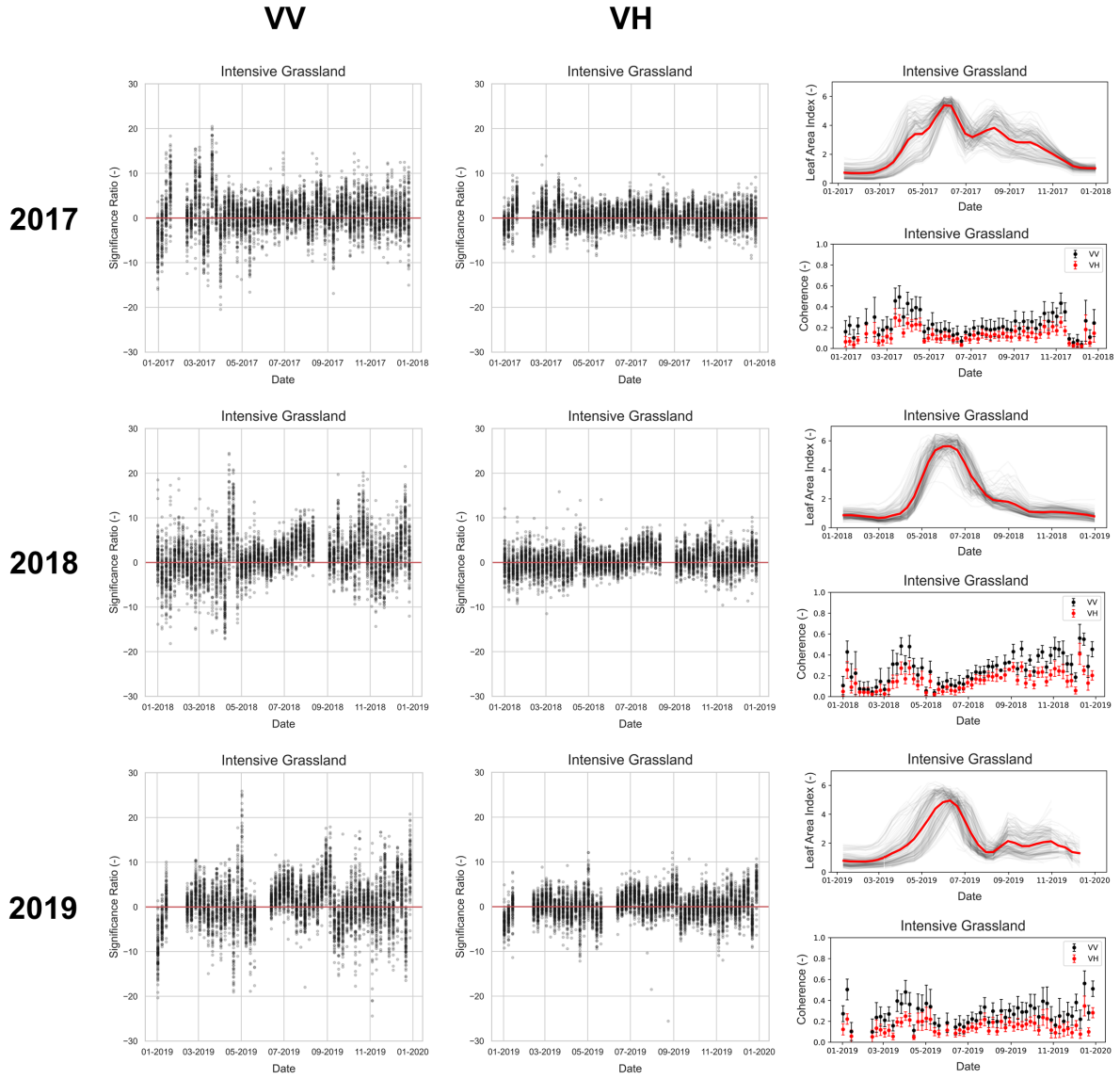


Figure 4.15: Same setting as in Fig. 4.12 but for intensive grassland.

The crop cycles for winter, summer and perennial crops (intensive grassland) are associated with sudden increases and decreases in vegetation density. Vegetation coverage leads to volume decorrelation during the spring and summer months due to the presence of multiple scatterers within a single resolution cell which increases the phase uncertainty (Martone et al., 2016). For periods of vegetation cover, indicated by higher LAI values (and therefore lower coherence magnitudes), the significance

ratio tends to be low (> -5 and < 5) and consistently low, i.e. the observed closure phase values are most likely to be noise rather than of geophysical origins.

Outside those periods, soils of winter and summer crops are barely vegetated, if at all, as the LAI indicates < 0.5 values for both summer and winter crops, and the coherence magnitudes are generally higher. Thus, the sensitivity of closure phase to dielectric changes from soil moisture is higher during those periods. Even though the significance ratio is much higher than during periods of vegetation cover (high LAI), it also varies a lot over time. At certain dates, the magnitudes of significance ratio are as low as during the high LAI period, which means that the closure phase can only be treated as an indicator of a geophysical origin for dates where the significance ratio is higher. In the summer/spring period, the highest possible significance ratio seems to be confined to about |5| for summer and winter crops, and <|10| for intensive grassland. Projecting that on the ability of closure phase to distinguish between areas that are covered with vegetation or not, or to monitor vegetation growth from time series, those values can be considered as adequate thresholds. If the significance ratios consistently exceed those thresholds it can be assumed that the vegetation cover of the soils has decreased and vice versa.

Looking at winter crops in Fig. 4.13, one can also notice a negative trend from primarily positive significance ratios (up to 10) at the beginning of the crop cycle in early March until the maximum LAI at end of April. Such a trend is not noticeable over summer crops which are characterized by a period of crop growth and a bare soil as well. A possible explanation for this trend could be found in differences of the crop type structure and its associated LAI trajectories. The LAI trajectory of winter crops shows a steady and elongated increase in contrast to the LAI evolution of summer crops which is rather steep in comparison. Hence, high significance ratios (over bare soils) abruptly decrease due to the sudden increase of the corn-LAI and the sudden decrease of the coherence magnitude, i.e. the closure phase values are noise rather than of geophysical origins. Over winter crops (straw cereal), however, the steady increase in LAI and decrease of the coherence magnitude along with a steady decrease of the significance ratio indicates that the closure phase is still sensitive to ground surface properties, and the effect of volume decorrelation as the vegetation density increases is delayed rather than sudden. The fact that the significance ratios are primarily positive (before its magnitudes decrease with increasing LAI for winter crops), generally seems to hold for periods when the soils are sparsely vegetated or bare, i.e. before the leaf development of crops and after harvest.

Worth noticing are also the large (negative) cluster of peaks in January of 2017 over winter and summer crops. As Mironov et al. (2017) and Zwieback et al. (2011) indicate the similar dielectric response of drying and freezing soils, those large significance ratios are likely linked to frozen soils in early January. This is not an objective of this study but potentially interesting to explore in future research.

Broad-Leaved Forests

The significance ratios over broad-leaved forests in Fig. 4.16 show that the recurrence and magnitude of its values is generally lower when compared to the other cover types. The coherence magnitudes are consistently low as well, mostly $\sim$0.1. The LAI indicates that leaf development (leaf-out) starts in the beginning of April and reaches its peak mid-May. The LAI remains at its peak of about 6 and decreases in October, when the leaf-fall has set in. Outside this period, the LAI is mostly < 1 and not zero due to the presence of grass or other low vegetation on the forest floor. The significance ratios are generally the lowest for the period of maximum canopy development, except a few single peaks. Worth mentioning is the accumulation of higher ratios, in December 2018 where both polarizations indicate peak-cluster. This could be caused by soil wetting since the sensitivity to soil moisture increases when the trees are leafless. Moreover, scattering contributions from double-bounce with bare tree trunks enhance

the closure phase signal. De Zan and Gomba (2018) describe that each bounce causes a phase delay which adds to the final phase triplet. Two bounces with an equal effect on the interferometric phase would lead to a doubled phase effect in the interferogram and hence the closure phase would also double. This could be a possible explanation for high ratios in periods where the LAI is < 1. In comparison to the other land cover types, the time series lack the distinctive seasonality observed for winter and summer crops, i.e. a distinct period where the soils are bare.

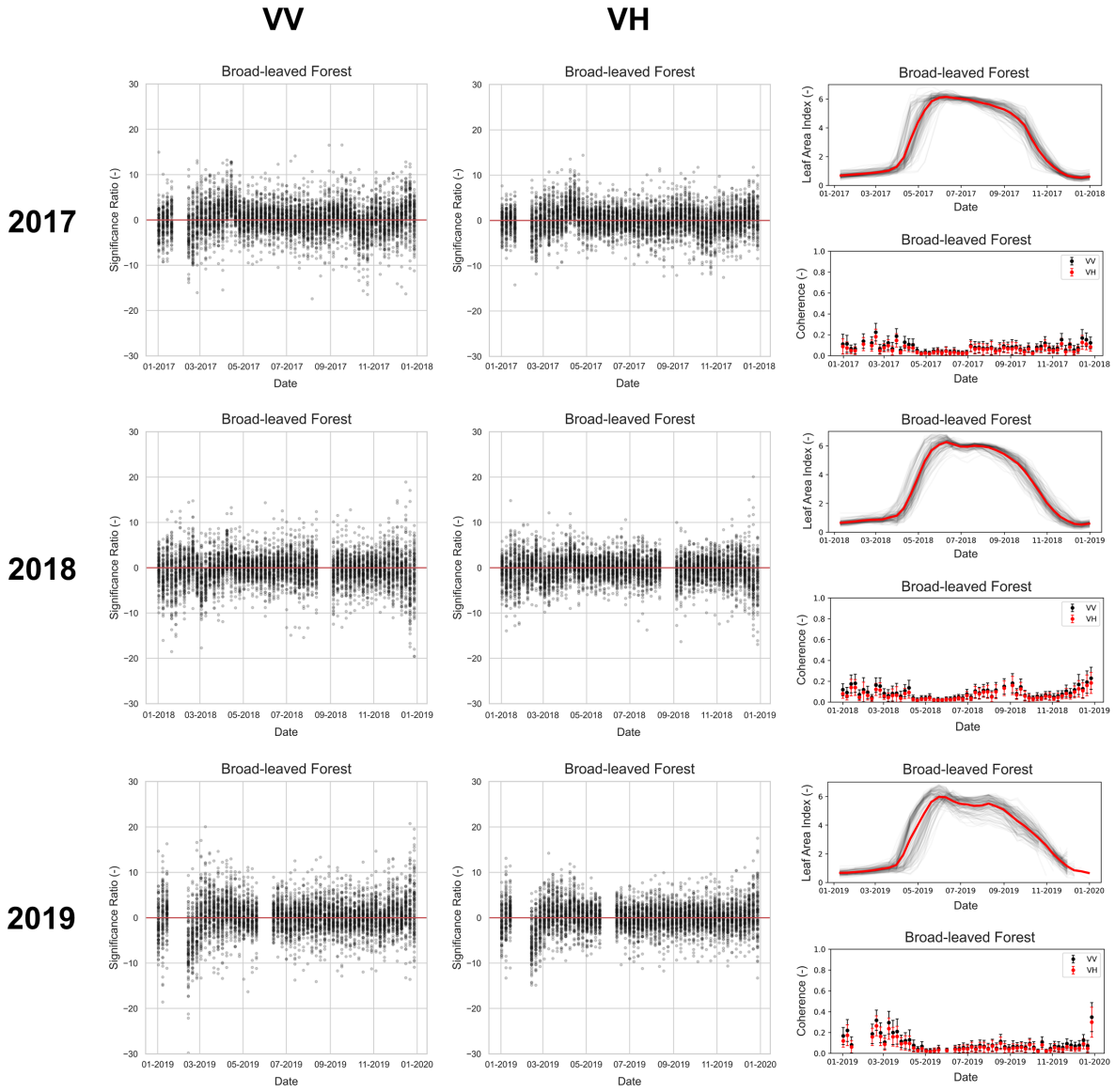


Figure 4.16: Same setting as in Fig. 4.12 but for broad-leaved forest.

4.4. Spatial Consistency of Significance Ratio

In this section, the correlation between significance ratio time series from 2017 and 2018 is assessed as a function of spatial distance between the polygon-fields. The objective of this assessment was to find out how dependent the consistency of the significance ratio is on locality, i.e. if phase closures are primarily sensitive to local target changes, such as soil moisture variabilities and how this changes with spatial distance. The assessment was performed through Pearson's *R* correlation analysis in compari-

son for different land cover classes. The assessment was also carried out for amplitude backscatter to find out, if both significance ratio and amplitude have a similar sensitivity to distance. Including Theia's soil moisture product in this comparison served the purpose to investigate if the sensitivity to distance between the soil moisture product, significance ratio and amplitude is similar. The sensitivity analysis of significance ratio to spatial averaging proved that time series from different fields (maximum distance 25 km) had a moderate to strong R over crops and grassland. As the level of spatial aggregation was a limiting factor for the similarity between significance ratio time series, only fields with a size of > 8000 looks were considered for this analysis. Pearson's R was then estimated over pairs of fields that are grouped in bins with a distance of < 20 km to one another. It has to be mentioned that due to the heterogeneous land cover distribution over the research area, the maximum distance between individual land cover polygons is at 250 km for grasslands and forests and kept at 210 km for crops to guarantee a sufficient number of polygons per distance bin.

4.4.1. Agricultural Crops

Fig. 4.17 compares the boxplots trajectories of Pearson's R over agricultural crops for 2017 and 2018, to analyse if a sensitivity to distance largely depends on the "exact" land cover type, or rather on the general spatio-temporal characteristics and similarities in vegetation structure. E.g., corn and sunflower (2018, 2019) are different types of crops, but are characterised by a similar crop cycle (see Fig. 4.14), and therefore both classified as summer crop for the 2017 nomenclature. Due to the shortage of fields for sunflower, summer crops are compared to corn fields. The upper panel shows winter and summer crops for 2017, the lower panel presents straw cereal (winter crop) and corn (summer crop) for 2018. The boxplot trajectories for both years indicate that the distance between the time series has a significant effect on the correlation between significance ratios. Straw cereal has a moderate Pearson's R of 0.51 at short distances < 20 km for both polarisations, which decreases to 0.18 (VV) and 0.13 (VH) at 210 km. Winter crop (2017) has a similar evolution, however, slightly lower R at 0.39 (VV, VH) and a relatively steady decrease to about 0.10 (VV) and 0.12 (VH). Summer crop (2017) shows a similar trajectory to winter crop with moderate R at 0.40 (VV) and 0.42 (VH), decreasing to significant correlation at 210 km (about 0). The boxplot evolution of corn (2018) shows a more pronounced sensitivity to distance as the median R decreases from 0.51 to 0.12 over 70 km, and remains constant with increasing distance. No significant differences between VV- and VH-pol are noticeable for each crop type.

The similarity analysis for amplitude backscatter in comparison shows a similar sensitivity to distance when comparing the boxplot trajectories to significance ratio, however, the correlation between the time series is generally higher across the crop types, independent of the polarisation. The correlation at small distances is between 0.9 and 0.8 for summer crops, winter crops and straw cereal, and at 0.66 for Corn. With increasing distance between the polygons, the correlation decreases similarly as observed for significance ratio. The boxplots of Theia's soil moisture product show a steady decrease in correlation with increasing distance as well. The general trends of the trajectory are similar to the ones observed for significance ratio and amplitude. Over straw cereal, the correlation is at R=0.96 for distances < 20 km and decreases to R=0.52 at 210 km. Over summer crops, R decreases from 0.95 to 0.35 at 210 km.

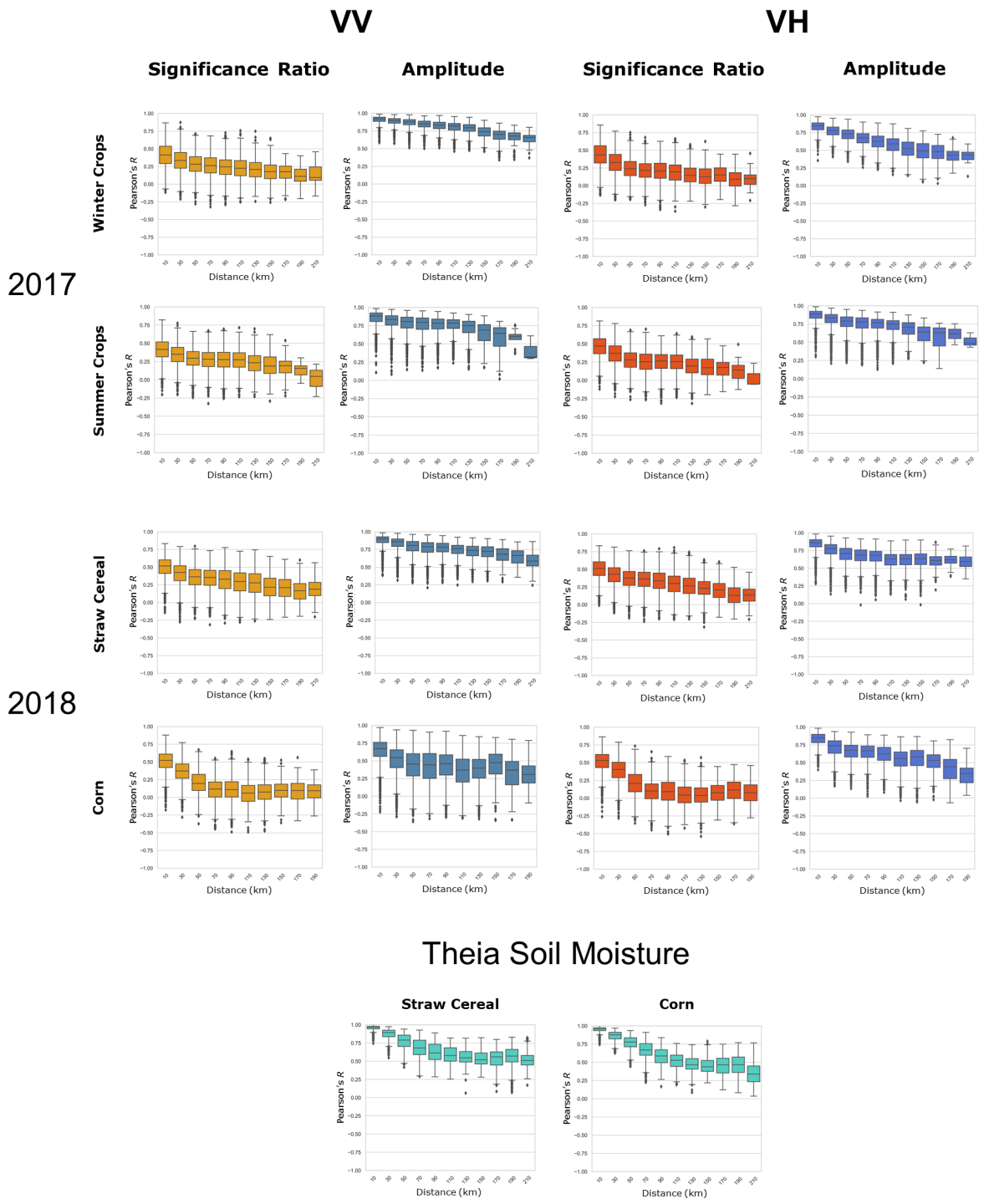


Figure 4.17: Boxplots of Pearson's *R* for 2017 and 2018 time series of significance ratio over crops in comparison with amplitude and Theia's soil moisture product as a function of distance between the time series. The line shows the median value. The bottom and top of the box the 25th and 75th percentiles, respectively. The upper tail is the 75th percentile plus 1.5 times the difference between the median and 75th percentile. The lower tail is the 25th percentile minus 1.5 times the difference between the median and 25th percentile. Outliers are plotted as diamonds.

4.4.2. Grasslands and Broad-Leaved Forest

Fig. 4.18 compares Pearson's *R* for significance ratio for the land cover types natural and intensive grassland, and broad-leaved forest with backscatter amplitude. The boxplot trajectory of broad-leaved forest can not effectively indicate the effect of distance on the similarity as the initial similarity for ad-

jacent time series is low to very low (VH: 0.25, VV: 0.21). The trajectories for both polarisations show a further decrease in correlation to ~0 for VV and 0.10 for VH. Natural grassland however suggests a similar sensitivity to distance as observed over crops. Pearson's *R* is moderately high at 0.52 (VV) and 0.45 (VH) and decreases to 0.15 for both VV- and VH-pol over 70 km distance. Afterwards, the similarity between the time series does not seem to be affected significantly by further distancing. The correlation for intensive grassland decreases from a moderately high *R* at 0.42 (VV) and 0.37 (VH) to 0.11 (VV, VH) at 110 km distance. Interestingly, the similarity seems to increase again between the 110 km and 170 km distance lag, before it decreases thereafter. One has to consider the characteristics of this land cover type here. As intensive grassland features different perennial crops, it is argued that the initial decrease in correlation is caused by the effect of distance over different vegetation types while the crop types at a distance of 110-170 km to one another are largely the same.

Comparing to the boxplots of backscatter amplitude, the trajectory show noticeable differences. For natural grassland, the spread in *R* is significantly larger when compared to significance ratio. The suspicious buckling in the evolution of intensive grassland boxplots is also noticeable for backscatter, particularly for VV-pol. In contrast to the significance ratio, boxplot trajectories of backscatter over broad-leaved forest indicate a high initial correlation at 0.76 (VV) and 0.93 (VH) which decreases steadily to low correlations of 0.36 (VV) and 0.51 (VH) respectively. This highlights the ability of amplitude backscatter to monitor vegetation, leading to similar time series even if the distance between the fields is large. The generally higher correlations between VH-pol time series over broad-leaved forest compared to VV-pol is an indication for the better ability of VH-pol to monitor vegetation.

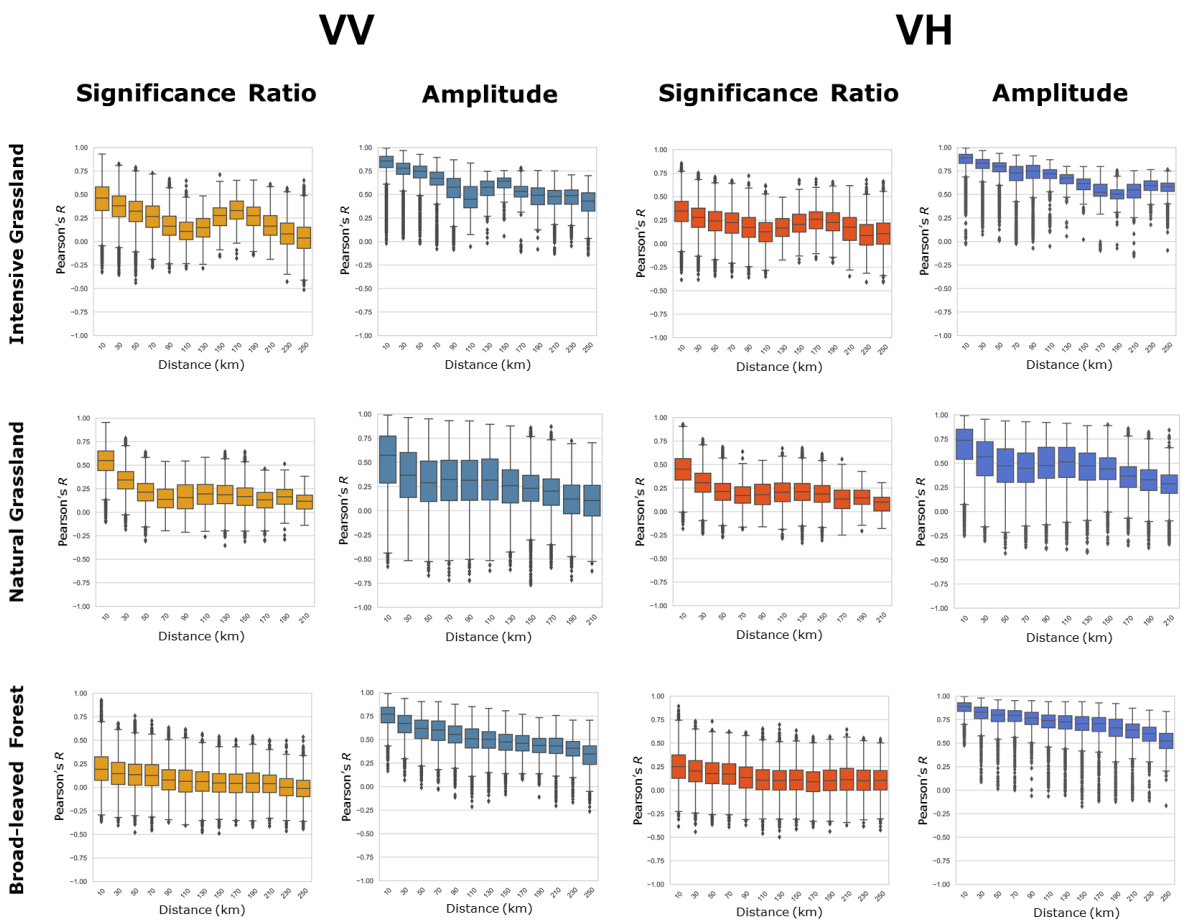


Figure 4.18: Boxplots of Pearson's R for time series of significance ratio over forest and grassland $\Psi_{i,j,k}$ in comparison with amplitude as a function of distance between the time series. The line shows the median value. The bottom and top of the box the 25th and 75th percentiles, respectively. The upper tail is the 75th percentile plus 1.5 times the difference between the median and 75th percentile. The lower tail is the 25th percentile minus 1.5 times the difference between the median and 25th percentile. Outliers are plotted as diamonds.

4.4.3. Effect of Spatial Distance

The correlation analysis revealed that land cover types with lower vegetation in particular, including winter/summer crops and natural grassland, show very similar results as Pearson's R decreased with increasing distance between the land cover polygons. This suggested that geophysical phase closures were primarily initiated by local events associated with e.g. precipitation which impacts soil moisture dynamics. Backscatter amplitude which is a well established observable to measure soil moisture changes, largely coincided with the boxplot trajectories from significance ratio, indicating the same underlying contribution linked to moisture changes. However, R was consistently larger than for significance ratio across all distances. Boxplot trajectories for Theia's soil moisture product showed a similar decline in correlation with distance. The striking observation of the buckling in the trajectory of intensive grassland boxplots needs some further thoughts: Intensive grassland is primarily located between summer and winter crops in an area surrounding Toulouse. As indicated by Fig. 4.15, the vegetation cover is significantly higher throughout the year compared to natural grassland and annual crops, and the evolution of LAI for 2017 indicates two distinctive peaks of crop growth. The alternation of large significance ratios and periods of crop growth (associated with lower significance ratios) at two periods in Fig. 4.15 originates from different crops at presumably different regions. Because of that it was argued that the buckling is owed to largely the same vegetation cover at this distance, therefore,

R was sensitive to the changing magnitudes of significance ratios related to different crop cycles which happened to be consistent at 110-170 km distance. The comparably high correlation between time series for this land cover type also suggests that the closure phase sensitivity to soil moisture changes is robust to moderate vegetation cover of about 3-4 LAI. Another reason why this observation is quite valuable is the fact that the buckling is also reflected in the amplitude trajectory which further confirms that amplitude and phase closure are sensitive to similar target properties.

4.4.4. Spatial Comparison between Significance Ratio and Amplitude

The Pearson correlation analysis between amplitude backscatter and significance ratio showed that both observables were similarly sensitive to increasing spatial distance. The decrease in correlation between the individual time series (also in comparison with Theia's soil moisture), suggested that both observables are largely influenced by local target changes. Since amplitude backscatter is sensitive to soil moisture and significance ratio showed the largest magnitudes over periods of low vegetation cover (LAI <1), this section aims to compare both observables spatially for a period where the soils are sparsely vegetated or bare. The goal was to investigate if a sensitivity to soil moisture is reflected in the spatial signatures of both amplitude and significance ratio, and what the differences between the two were. Therefore, the land cover type winter crops was chosen for the comparison between three phase triplets and the five corresponding amplitude images between 10.08.2021 and 03.09.2021. The maps in Fig. 4.19 are filtered for fields of winter crops and additionally for significance ratios $> |2|$ to highlight spatial signatures (as shown in Fig. 4.4). The background shows exaggerated hillshading to indicate topography. The chosen period is after the harvest has taken place over the whole area and the bare soils are only exposed to rainfall rather than additional irrigation.

The significance ratio varies noticeably between the three maps, showing distinct non-random contiguous patterns. There is visual evidence that the red area in the left part of the first triplet changes from red, positive significance ratio, to blue which likely indicates a relative change in the target properties related to soil moisture changes. A few precipitation events between mid-August and September (see Fig. C.2) most likely caused the different significance ratio patterns in the maps. The significance ratio remains negative for the aforementioned area which means that the soil moisture change, presumably soil wetting due to the precipitation events, extended until the last date, otherwise the phase would have been consistent and mostly filtered out due to the predefined $< |2|$ ratios. The amplitude time series as supplementary products attempt to help understanding the relative soil moisture variabilities. The single acquisition dates help to isolate relative changes of the ground properties by simply comparing the spatial signatures of the patterns with the ones observed in the triplets. In general, the spatial patterns show a good similarity with the ones in the triplets. Obviously those analysis fall victim to subjectivity, but visual observations can be summarized nevertheless:

In the 16.08.2017 amplitude image for example, one can see that the blue area in the map center extends about 150 km to the east. The same pattern can also be found in the first and second significance ratio map. However, the corresponding significance ratio indicates another distinctive red pattern that extends in northeast direction which is not as pronounced in the 16.08.2017 amplitude image. Since linear changes in moisture have a cubic effect on closure phases (De Zan & Gomba, 2018) is likely the reason for the enhanced significance ratio signal in this area. In general, changes of amplitudes towards larger values are linked to soil wetting (Entekhabi & Moghaddam, 2007). Assuming that holds true for what is illustrated in the map, the positive significance ratio would have been caused by soil wetting. Moving forward, the amplitude acquisition on 28.08.2017 again shows blue and green patterns that match with those illustrated in the second and third significance ratio image. However,

here a relative change towards larger amplitudes is reflected in negative significance ratio signs while lower backscatter is indicated by positive significance ratios.

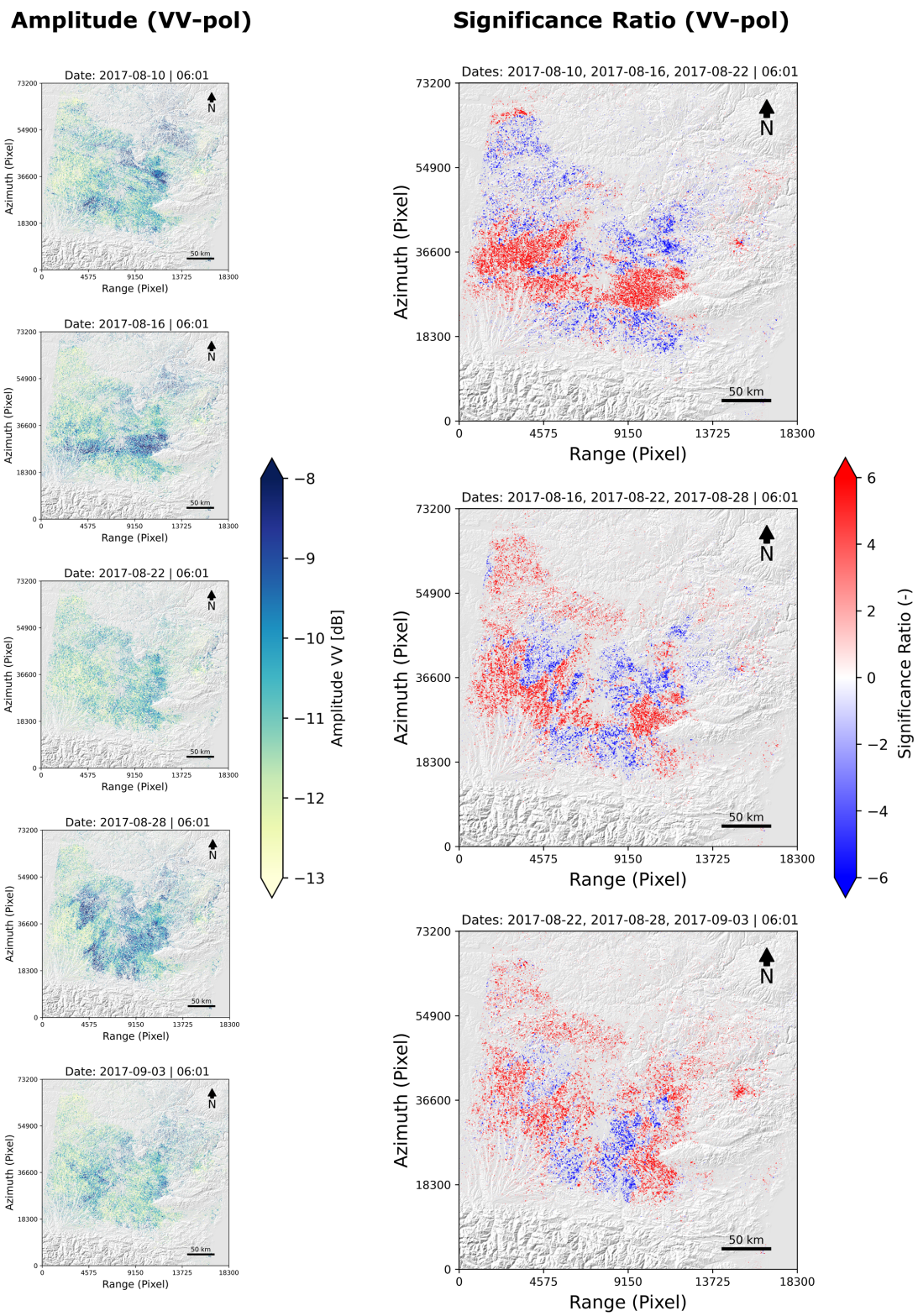


Figure 4.19: Three consecutive $\Psi_{i,j,k}$ maps with exaggerated hillshading between 10.08.2017 and 03.09.2017 and five consecutive amplitude maps for the same dates as given by the significance ratios. The maps are filtered to only represent the land cover winter crops. The dates are picked because of the bare soils after harvest which increases the sensitivity of the radar to soil moisture.

5

Discussion

The initial motivation for this study was very curiosity driven in the beginning because a characterisation of closure phase over different land covers has not been carried out in previous research. The aim of this project was to provide guidelines on the derivation of closure phase, standardized through a significance test, and essential conclusions from a spatial and temporal characterisation over different land cover types. Those are important steps towards using the closure phase for soil and vegetation monitoring in the future.

To this end, the objectives of this study were to derive closure phases over different land covers (1), to adapt the closure phase significance test and define measures for its parameter-inputs (2), and to demonstrate the effect of spatial averaging on the closure phase value and consistency (3). Important decisions and conclusions proved to be essential for the subsequent characterisation and analysis of phase closures over different land covers and in comparison with other observables (4). Based on these objectives, this chapter reflects on the main findings of this study, discusses its conclusions and limitations, and suggests recommendations for future research.

5.1. Parameter Assumptions in the Significance Test

A big part of this study ended up addressing a lot of methodological questions that were not anticipated before but were essential and valuable to derive a significance ratio and to characterize it spatially and temporally. Multilooking in this study was based on the spatial aggregation of pixel-values over homogeneous fields with similar land cover. This approach was advantageous for two reasons in particular: Firstly, the quality of the interferometric coherence estimate improves as the number of averaged samples increases and secondly the aggregation of pixels over land cover fields takes the spatial and temporal differences between land covers into account (Nikaein et al., 2021). However, for simplification we assumed that the number of averaged samples in a field is equal to the number of independent looks *n* which is wrong. Put simply, adjacent pixels in a raw SAR image become dependent once a SAR matched filter has been applied (Gierull & Sikaneta, 2002). The effect is that given *m* SAR image pixels that form one multilooked pixel, only $n < m$ are statistically independent. In practical terms, this means that fields which obtained the same number of samples due to similar spatial sizes, have different number of looks.

Besides that, the coherence estimate is biased towards higher coherences magnitudes, in particular if the a number of looks is low. The coherence bias and its estimation was explained in detail in Chapter 3 and visualized in Fig. 3.5 as a function of different looks. The definition of parameters for the coherence interpolator in the significance test was based on the statistical relation between the number of looks and standard deviation in coherence estimation. In view of the above mentioned assumptions regarding the number of looks and the uncertainties in the coherence estimate, this choice requires a critical comment. In the attempt to find a measure for appropriate step sizes in the coherence interpolator in relation to the averaged samples per field, estimates of the standard deviation in coherence magnitudes were exploited. The decrease in standard deviation as the number of looks increases, as illustrated in Fig. 3.5, was used to define the size of coherence interpolator steps. Even though differences in the quality of the coherence estimate with respect to the number of looks is principally an appropriate measure to determine the interpolator setup, the attempted accuracy in defining a “correct” step size is prone to uncertainties itself, in particular for low coherences where a lot of averaging is required to reduce the coherence bias. The uncertainty and bias involving the estimation of coherence magnitudes undermines the fine tuning of the coherence interpolator. This was at least acknowledged in choosing to confine the minimum coherence step to 0.01 referring to 100 coherence values between 0 and 1.

5.2. Spatial Averaging over Different Land Cover Types

The degree of spatial averaging proved to have a significant impact on significance ratio values and the temporal correlation between significance ratio time series. A visual comparison of three maps with different multilooking ranges indicated that the degree of spatial averaging is important to accentuate spatial signatures and increase the magnitudes of significance ratio.

In order to compare the effect on the significance values, it was decided to calculate the RMS for fields of different multilooks and compare its average across different land covers. The purpose of this comparison was to demonstrate the effect of averaging on the magnitudes of significance ratio rather than to indicate the influence on its central tendency. Large significance ratios that are not “expected” given the coherence magnitudes and degree of multilooking were particularly interesting for the understanding of phase closures over different land covers. The comparison showed that an increase in multilooking increases the magnitudes of significance ratio across all land cover types, however, the RMS over croplands and natural grasslands were consistently larger than compared to forests and intensive grasslands.

The two primary differences between the land covers that influence the effect of multilooking on the significance ratio are the vegetation density and the consistency of vegetation cover throughout the year. Over low vegetated areas such as natural grasslands and croplands (before bud and leaf development, after harvest), the radar primarily senses the ground surface while over cover types with relatively large vegetation (greater LAI), including forests and intensive grasslands, it primarily senses the vegetation canopy. Volume scattering from vegetation is one of the main sources that corrupts the interferometric phase and leads to decorrelation (Flores-Anderson et al., 2019). For C-Band more so than for longer wavelength, the sensitivity to volume scattering is high (Flores-Anderson et al., 2019) which led to a larger volume scattering component in phase closures over forests (in particular) and intensive grasslands. The differences in sensitivity can further be explained by considering both the significance ratio and the observed closure phase in comparison. The influence of phase noise was the highest over forests and almost independent of the multilook which showed in the greatest magni-

tudes of the observed closure phase. The RMS of the significance ratios indicated that a higher level of averaging was required to improve its SNR and increase its magnitudes. Worth mentioning are the characteristics of intensive grasslands. Similar to the crop types, the observed phase closures over intensive grassland showed lower phase levels while the sensitivity of the significance ratios to multilooking was comparably low, similarly as low as over forests. This emphasizes the importance of the significance test, as different magnitudes of the signal and noise component respectively can lead to similar magnitudes of the significance ratios.

Here it is argued that over land cover with low vegetation (natural grasslands and agricultural crops), changes in the dielectric properties from changing soil moisture (in the top few cm) and consistently greater coherences (see Fig. 4.13, 4.14 and 4.12) are linked to the greatest significance ratios, hence yield the greatest RMS. Those large changes of the dielectrics are not sampled by the radar over dense vegetation since the ground is shielded by vegetation. Periods of vegetation cover that are associated with volume decorrelation, influenced the degree of spatial averaging that was required to increase the SNR of the geophysical closure phase over intensive grasslands and forests, leading to lower RMS of the significance ratio as multilooking increased.

Increased multilooking also had a positive effect on the linear relation between signficance ratio time series, in particular over croplands and the grassland cover types as suggested by Fig. 4.8 to 4.17. In contrast, the effect was very low over broad-leaved and coniferous forests. The aim of this comparison was to understand, if spatial averaging not only increases the magnitudes of significance ratio across different land covers but has also a significant effect on accentuating an underlying geophysical origin: If the degree of spatial averaging enhances the signal-to-noise ratio enough to isolate a geophysical signal component, then significance ratios within short distances to one another were expected to have a similar temporal signature. The results from Pearson's correlation analysis suggests that cover types that are comparably low in vegetation yield the most consistent significance ratios. This suggests that temporal changes of the dielectrics on the ground surface provoke similar phase inconsistencies while vegetation cover attenuates a consistent signal component, even if the degree of spatial averaging is large. The results over intensive grasslands were particularly interesting as the effect of spatial averaging on the magnitudes of the significance ratio was rather low, whereas the correlation between the time series increased significantly with multilooking. This indicates that even cover types with consistent vegetation cover and relatively low significance ratios can yield consistent signals.

Considering the role of different polarizations, generally speaking, cross-polarized (VH) data are more sensitive to volume scattering of canopies whereas co-polarized (VV) data are typically more sensitive to surface scattering components (Flores-Anderson et al., 2019). The impact of using different polarizations is most pronounced in the RMS analysis of significance ratio over natural grasslands and crops for VV-pol, where a gap is noticeable to forest and intensive grassland. An advanced sensitivity of VH-pol to volume scattering can also be derived from the RMS analysis for the observed phase closures, which obtained consistently greater RMS then compared to VV-pol. The greater sensitivity of VV-pol to the surface component is reflected by a steeper increase of the RMS for significance ratio which is not as pronounced for VH-pol. Interestingly, the differences in Pearson correlation between the polarizations were marginal and not consistently lower or higher for either one of the polarizations. This indicates that even though the closure phase is more sensitive to volume scattering and its significance ratios are more influenced by decorrelation, the effect on increasing the temporal signature of the geophysical phase closure is similar.

In this study the maximum multilook was limited to 10,000 looks, i.e. fields of 1x1 km. However, both the trajectories of the respective land cover type for the RMS comparison and Pearson's correlation

analysis indicate that even further spatial averaging could increase the significance ratios as well as the correlation between individual time series, which would further highlight its temporal signatures. On the other hand, more averaging diminishes the spatial resolution which is not advantageous to potentially monitor soil moisture changes over areas that are very heterogeneous in land cover such as croplands. This tradeoff has to be considered, as the monitoring of soil moisture is particularly interesting over agricultural crops. If closure phase were used for soil moisture retrieval or utilised as a supplementary observable, the spatial resolution is an important deciding factor.

5.3. Closure Phase Sensitivity to Dynamics in Vegetation

The vegetation cover over soils seems to be the most relevant determining factor whether phase closures contain a great geophysical component or not. In a time series comparison of significance ratio, coherence and LAI over 100 fields, all land covers showed a distinctive signature for significance ratio that was consistent with characteristics observed for LAI and the coherence magnitudes. Periods of vegetation growth with peak LAI between 4 and 7 (depending on the land cover) were associated with coherence magnitudes < 0.2 and the lowest significance ratios. With reference to the time series, a significance ratio of |5| seemed to be a good threshold value to distinguish between vegetation-covered areas or to monitor vegetation growth over time using the magnitudes of significance ratio. During periods of low vegetation cover or over bare soils, the coherence was higher and the LAI were significantly lower, < 1 over all land covers. Those periods were linked to the greatest significance ratios, which suggests that most “true” phase closures primarily derive from changes in soil moisture in the top few centimeter of the ground surface. The occurrence of significance ratios that were $>> 5$ was only observed during those periods of low LAI. However, the significance ratios were not consistently above the considered threshold value of |5|. Sometimes they were as low as during periods of vegetation cover (high LAI), which means that the closure phase cannot be simply used as an indicator of a geophysical origin during those periods. Only for some dates. This suggests that a significant (or sudden) change of the soil properties, e.g. induced by rainfall, is necessary to obtain significance ratios that are much greater than |5|.

Drawing the simple and incomplete conclusion that phase closures obtained over densely vegetated areas such as forests obtain no valuable information is too hasty. Previous research by De Zan and Gomba (2018) pointed to different behaviours of the closure phase over different land cover types and suggested the investigation of those findings in future studies. They observed a strong closure phase signal and consistent retrievals of soil moisture over forests and suggested that the water status of the trees and plants could contribute to the closure phase signal. Variations of the dielectric constant of trees have been demonstrated by McDonald et al. (2002). This thesis also showed that forested areas obtain the greatest (measured) phase closures, yet among all land cover types the significance ratios were the lowest and the distinctive seasonality observed in the other cover types was not present.

5.4. Similarities between Amplitude and Closure Phase

In order to better understand the influences contributing to inconsistencies in phase closures, amplitude backscatter was chosen for a comparison with significance ratio since its sensitivity to vegetation and soil moisture is confirmed by many studies (e.g. Steele-Dunne et al. (2017) and Wagner et al. (2003)). Hence, similarities in spatio-temporal characteristics allow to draw conclusions about the mechanisms behind phase closure inconsistencies. The results from a comparison of the spatial consistency between time series of significance ratio and amplitude backscatter showed similar boxplot trajectories

as the distance between the individual fields, that were compared, increased. This led to the initial assumption that both observables are primarily influenced by local changes of the target properties linked to soil moisture.

To confirm the similarities between significance ratio and amplitude backscatter spatially, three maps of significance ratio were compared to the spatial signatures of the corresponding five amplitude images over winter crops (Fig. 4.19). The dates 10.08.2017-03.09.2017 were picked intentionally after harvest, where a significant increase in significance ratio was observed in Fig. 4.13 and soils were barely vegetated. The comparison showed that variabilities in soil moisture cause similar spatial variabilities in amplitude backscatter and significance ratio. Brightening effects from amplitude, i.e. greater backscattering coefficients, stem from wet soils and vegetation and are indicated by the blue patterns on the map. Green to yellow areas between the blue patches, i.e. lower amplitude backscatter, likely indicate active raining, where the convection system likely caused signal attenuation (Flores-Anderson et al., 2019). Presumably active raining was also found in Fig. 4.11 over an area in fall of 2018. Similar patterns have been observed for the significance ratio, where the blue areas indicate the same pattern as the blue areas in amplitude backscatter, suggesting that negative significance ratios have been caused by soil wetting. However, in an earlier triplet, the positive significance ratios are linked to greater amplitude backscatter, which was attributed to soil wetting as well. This ambiguity is owed to the inherent moisture ordering problem in closure phases. The correct underlying moisture evolution cannot be retrieved solely from the phase closure signs, as the correct ordering is subject to cyclical permutations, meaning that the same closure phase sign can be caused by different moisture evolutions. Also worth noting is the distinct red pattern in northeast direction in the 10.08-16.08.-22.08. phase triplet, which is not as pronounced in the amplitude image. Since linear changes in moisture have a cubic effect on closure phases (De Zan & Gomba, 2018) is likely the reason for the enhanced closure phase signal in this area and further suggests the advanced capabilities of closure phase to monitor soil moisture dynamics. The fact that using amplitude information for soil moisture retrieval is less complicated and constrained than using phase triplets could be exploited for solving the cyclic permutations and finding the correct moisture order.

There are also distinct differences that derive from the comparison between significance ratio and amplitude backscatter. Considering the Pearson correlation analysis, the scaling of R is generally higher for amplitude backscatter with medium-low correlation even at large distances, whereas the correlation of significance ratio time series approach R=0 at $>$ 200 km distances. Vegetation growth is generally associated with an increase in amplitude backscatter (Vreugdenhil et al., 2018) which translates into a consistent signal component. For significance ratios, however, periods of vegetation cover were associated with the lowest magnitudes and a decrease in spatial consistency, as a result of volume scattering mechanisms and the associated decorrelation. Hence, the similarity between individual significance ratios decreased significantly during those periods. Since vegetation growth cycles are relatively consistent (as illustrated by the LAI time series), an increasing amplitude backscatter over vegetation covered areas adds a relatively consistent term to the time series which contributes to the similarity between them. For closure phases, vegetation primarily increases the phase noise component which subtracts from the correlation between the time series and decreases the significance ratio. The “steady” decrease in correlation for both amplitude backscatter and significance ratio therefore is primarily caused by local soil moisture changes during periods of low vegetation cover.

Considering the time series correlation over broad-leaved forests e.g., the dense vegetation cover seems to largely suppress a potential consistent signal component due to low coherences over forests and associated phase noise (as already discussed above), whereas the correlation for amplitude

backscatter is significantly higher. However, this does not rule out the possibility that there is a consistent phase closure mechanism for certain periods, e.g. linked to seasonal vegetation growth or water uptake. The consistency of those single peaks in significance ratio across different time series, however, would not lead to high scores of Pearson's *R*. Therefore, more sophisticated measures of correlation or change detection for time series over vegetated areas might be better to assess a consistent signal. The influence of vegetation decorrelation remains a challenging factor in the understanding of closure errors.

5.5. Limitations and Recommendations

It is acknowledged that the results and interpretations presented in this thesis are subject to limitations:

- **Bounded multilook**: Pearson's *R* statistical analysis showed that the similarity between time series strongly depends on the degree of spatial averaging, with the exception of forests. The spatial aggregation of pixels for this research was kept at a maximum field size of 1x1 km, referring to 10,000 pixels. However, the results suggested that further aggregation over fields of the same land cover type could presumably further increase the SNR of the geophysical closure phase. One has to keep in mind the impeding effects due to atmospheric artefacts, wherefore the averaging of the complex interferogram should be bounded to smaller field sizes (here 1x1 km). Further averaging of phase triplets would require to 1) derive phase triplets similar to the approach applied in this project and 2) average the phases of the respective phase triplet thereafter to increase multilooking.

- **Pearson's *R* as similarity measure**: The Pearson correlation measures the linear similarity between all points in a time series and is therefore not a good measure to detect a few consistent peaks in a time series. Time series of significance ratio were characterized by significantly larger magnitudes for certain periods while vegetation cover introduced phase noise. Therefore, in the case of forest vegetation and other densely vegetated land covers (which are generally associated with decorrelation and low significance ratios), alternative measures based on e.g. change detection could provide more suitable ingredients to analyse consistent features in time series of significance ratio. A relatively new approach for time series analysis refers to "shapelets" which are described as discriminative sub-sequences of time series that best predict the target variable (Grabocka et al., 2014). Those methods could help to identify certain land covers from a few consistent peaks in time series of significance ratio.

- **Synergies with amplitude backscatter**: The comparison between single acquisitions of amplitude backscatter and significance ratio showed noticeable spatial similarities, hinting that amplitude could be a valuable observable to assist with solving the moisture order ambiguity problem in phase closures. Amplitude can provide supplementary information in two ways: Firstly, from the identification of similar patterns, one can derive between which two acquisitions the most significant moisture change occurred. Secondly, the trend of change towards increasing or decreasing amplitude values can provide valuable information about the correct moisture trajectory related to the sign of the phase triplets. As negative or positive phase closures constrain the number of possible moisture orders, intermediate amplitude data could be exploited to further constrain or even solve the ambiguity problem.

- **Clustering with respect to land cover**: Unsupervised clustering of significance ratios based on its temporal signatures was one of the initial research goals discussed for this project. However,

due to the unknown requirements of spatial averaging and the lack of knowledge on how spatial distance affects the correlation of significance ratio time series with respect to different land covers, those questions had to be answered first. The results suggested that the geophysical closure phase is primarily influenced by target changes that are very local, e.g. associated with rain events. One would have to apply several constraints on the time series of significance ratio (max. spatial distance, min. degree of multilooking) before attempting to cluster for different land covers.

6

Conclusion

The research objectives of this study were to derive a significance ratio for closure phases from a numerical significance test and to further characterize them spatially and temporally over different land cover types. The results are important steps towards using the closure phase for soil and vegetation monitoring in the future. A solution to aggregate phase closures based on land cover polygons was presented first. Necessary modulations of the closure phase significance test were discussed and the determination of its input-parameters were based on the statistical relationship between coherence magnitudes and the number of looks on a per field basis. In the context of a comprehensive sensitivity assessment, the effects of different degrees of spatial averaging on the significance ratio consistency and values were analysed. Essential conclusions were used to characterize significance ratio spatially and temporally, and in comparison with other satellite observables.

The research questions that were introduced in Chapter 1 are answered based on the results derived in this thesis.

Question 1 - *What processing steps are necessary and what parameter-based decisions are essential before performing a closure phase significance test?*

For this study, a spatial averaging approach over entire land cover fields was chosen which ensured the quality of the significance test and enabled to relate significant phase closures exclusively to the characteristics of the respective land cover. An important additional step was necessary in preparing the land cover maps: A visual inspection of the interferograms showed fringes that were related to atmospheric effects on a km-scale. In order to mitigate the conservation of those artefacts during the subsequent multilooking procedure, it was necessary to ensure spatial aggregation over bounded distances which was given by applying a 1x1 km mesh-grid to the land cover maps. For the field-averaging, the complex interferogram was restored from the coherences and amplitudes, each instance was multilooked based on the land cover fields, and the coherences and phase angles were recomputed. A phase closure value for each land cover field was derived from the complex averaged coherence triple product. The result were phase closures, strictly aggregated over one land cover type and different resolutions ranging between 4 (minimum number of looks) and $\sim$10,000 looks.

The modified significance test required three parameter inputs: *Number of Looks*, *Coherence Interpolator Steps*, and *Number of Model Simulations*. The definition of those parameters was primarily

based on the statistical relationship between the number of looks and the standard deviation in coherence estimation, derived from Touzi et al. (1999). The number of looks was given by the number of averaged pixels multiplied by four (pre-multilook of a single pixel). The coherence standard deviation depends on the number of averaged samples, i.e. the greater the multilook, the smaller the estimates for the coherence standard deviation (see Fig. 3.5). The greatest estimate for the coherence standard deviation was determined as an appropriate step size for the definition of coherence interpolator steps. Lastly, determining a minimum number of simulations was necessary in particular for fields with a multilook > 1000, because the size of the simulated array grew exponentially with the number of looks and coherence combinations (see Fig. 3.6). For a low coherence combination (i.e. 0.1, 0.1, 0.1), the running standard deviation estimated the standard deviation along with the mean in a single pass of all model realizations. The results over different looks showed that the running standard deviation converged faster as either the number of looks increased, or the coherence combination increased. From a visual inspection of the curves presented in Fig. 3.7, a threshold for the running standard deviation of 0.01 radians ($\sim$0.57 degrees) was determined, while the minimum number of simulations regardless should not be lower then 500. This was essential to obtain reasonable runtimes and prevented month on end for minor improvement of $\hat{\sigma}_{\hat{\Phi}}$. A statistical relationship between the input-parameter was established which is presented in Fig. 3.8.

Question 2 - *What is the effect of spatial averaging on the significance ratio of closure phase and how does the consistency between individual time series of significance ratio change as a function of spatial averaging?*

Spatial averaging proved to be very essential for increasing the SNR of the geophysical signal component in phase closures, while the impact was more significant over fields with lower vegetation cover. Increasing the number of looks, in principal, decreases the phase noise component in phase closures while any geophysical component of the phase closure does not change. Increasing the SNR in phase closures is the general reason behind the significance test, where the estimates of the expected phase closure standard deviation depend on the degree of multilooking and the coherence magnitudes. Due to those dependencies, the significance ratios allowed to make a comparison between fields of different coherences and different looks. A visual comparison of three maps over the same area, filtered for different look-ranges showed that increased multilooking was capable to increase the significance ratio values and highlight consistent patterns of significance ratio. A comparison of the average RMS of significance ratio time series for different multilooking degrees showed that the effect of spatial averaging was different with respect to land cover type and choice of polarization. All land cover types indicated an increase in the RMS as the number of looks per field grew while the most significant impact was observed for crops and natural grassland. The fact that those land covers on average obtained the greatest significance ratios and were most sensitive to spatial averaging was explained by the differences in vegetation cover. Volume scattering within vegetation increased the phase uncertainty and resulted in lower significance ratios over forests and intensive grasslands, as indicated by lower RMS and the differences presented in a time series comparison. Moreover, using different polarisation influenced the magnitudes of significance ratios immensely. The sensitivity of VV-pol significance ratio to multilooking was significantly higher then for VH-pol and the differences between the RMS trajectories were much greater. The differences were explained by the greater sensitivity of VH-pol to vegetation structure, i.e. volume scattering had a greater impact on VH significance ratio, whereas VV-pol was more sensitive to surface scattering. This led to the distinct gap of RMS between cover types with lower

vegetation and intensive grassland and forests.

In a second assessment, a Pearson correlation analysis was performed between time series of significance ratio as a function of multilooking. All land cover types showed a steady increase in *R* as the number of looks per field increased. Natural grasslands obtained the greatest *R* with 0.72 (VV, VH) at 10,000 looks. Summer crops *R* increased to 0.46/0.52 (VV/VH) and winter crops to 0.51 (VV, VH); over intensive grassland *R* increased to 0.68/0.43 (VV, VH). Low to very low correlations were observed for coniferous and broad-leaved forests with 0.20 and 0.12 respectively at 10,000 looks. The high correlations, in particular over natural grassland and crops, suggest that the similarity between the time series was primarily driven by soil moisture dynamics in the top few centimeters. The increasing trajectories of all boxplots indicated that further averaging could further increase the SNR of the geophysical signal in closure phase and lead to an even greater *R* between the time series.

Question 3 - *How consistent is the significance ratio spatially and temporally as a function of different land cover types?*

The results in this study showed that there are distinct differences between the significance ratio over different land cover types. An assessment of the impact of spatial averaging on the significance ratio demonstrated the importance of multilooking to increase its values and the spatio-temporal consistency. Therefore, the subsequent characterisation over different land cover types, spatially and temporally, was carried out over fields with a multilook of $>$ 8,000. A time series comparison clearly indicated the influence of vegetation on the values of significance ratio. Over agricultural crops, the significance ratios were the greatest before and after their respective crop cycle. Those characteristics were consistent between all years and with respect to different crop cycles. During the period of maximum crop development with LAI between 4 and 6, the significance ratio mostly decreased to $<$ |5|, while periods outside the crop growth obtained phase closures with magnitudes $>$|10|. However, the significance ratios at times of low LAI where not always greater than during periods of vegetation cover (high LAI), which indicates that a significant (or sudden) change of the soil properties, e.g. induced by rainfall, is necessary to obtain significance ratios that are much greater than |5|. Over natural grasslands, in contrast, the significance ratios were consistently large throughout the year, without a distinct period of exceptionally low or high significance ratios. This coincides with comparably low LAI (mostly $<$ 2) and coherence magnitudes that did not have a distinct pattern throughout the year, such as the sudden decrease during crop growth. Over broad-leaved forests, the significance ratios where generally lower, mostly $<$|5|. The characteristical seasonality and distinct periods of greater significance peaks generally allow to distinguish between time series from different land cover types, despite the fact that the time series were derived from different locations over the whole research area.

Results from the Pearson correlation analysis to assess the spatial consistency of significance ratio time series showed that the correlation between the time series decreased with increasing distance. Natural grasslands and the crop types indicated a similar sensitivity to increasing distance, with the steepest decrease in correlation observed over the first 90-100 km. The high dependence on the spatial proximity highlighted the sensitivity of closure phase to local target changes, primarily influenced by soil moisture changes. Theia's soil moisture product showed a similar boxplot trajectory with the steepest decrease in *R* over the first 90-100 km as well. The correlation of significance ratios over forest areas was generally very low (R $<$ 0.2).

Question 4 - *Is the significance ratio consistent with spatial characteristics observed for SAR amplitude?*

An assessment of Pearson's *R* for time series of significance ratio and amplitude backscatter at different distances showed that the boxplot trajectories in Fig. 4.17 and Fig. 4.18 have a similar decreasing trend for natural and intensive grasslands, and croplands. The boxplot trajectories were compared to boxplots of Theia's soil moisture product which indicated a similar sensitivity to spatial distance. This implied that both observables are similarly influenced by local changes of the ground properties associated with soil moisture changes. Over forested areas, the significance ratio generally indicated a very small correlation between the time series, whereas amplitude time series were strongly correlated at small distances and still moderately correlated at $>$ 200 km distance. The higher correlation between amplitude time series was attributed to the high sensitivity of amplitude to vegetation which resulted in a consistent signal. For significance ratio, however, periods of vegetation cover were associated with low significance scores and decorrelation. Amplitude and significance ratio were further assessed spatially over winter crops which confirmed that spatial signatures of amplitude backscatter are consistent with those of significance ratio. The period for the spatial comparison between amplitude and significance ratio was intentionally picked in late summer (after harvest) to ensure that the soils are bare and the radar senses the ground surface. The similar sensitivity of both observables to soil moisture changes suggests that amplitude backscatter could be useful to resolve the ambiguity associated with phase closures and provides an essential tool for soil moisture monitoring using the closure phase in the near future.

6.1. Added Value and Future Work

The results in this thesis provide further directions towards using the closure phase for soil and vegetation monitoring in future applications. This necessarily requires that its spatial and temporal characteristics are understood well. In the attempt to characterise the significance ratio of closure phases over different land cover types, a spatial averaging approach was applied to SAR pixels over areas with similar land cover. A closure phase significance test allowed to compare phase closures over fields with different coherences and different number of looks. Provided that the degree of spatial averaging was large enough, it was possible to relate differences in the magnitudes, and spatio-temporal signatures of the significance ratio to different land cover types. The results further pointed to a complementarity of the significance ratio with amplitude backscatter.

Those gained insights are essential to better understand the influence of soil moisture and vegetation dynamics on "true" phase closures, in particular with respect to different land covers. Future efforts should make use of the spatial and temporal features and different behaviours of phase closures over different land cover types. Eventually, the closure phase can be exploited for application including soil moisture retrieval, vegetation monitoring, or to separate their contribution from other signals, such as InSAR deformation estimates.

Bibliography

Ahern, F. J., Leckie, D. J., & Drieman, J. A. (1993). Seasonal changes in relative c-band backscatter of northern forest cover types. *IEEE Transactions on Geoscience and Remote Sensing*, *31*(3), 668–680.

Attema, E., & Ulaby, F. T. (1978). Vegetation modeled as a water cloud. *Radio science*, *13*(2), 357–364.

Autret, M., Bernard, R., & Vidal-Madjar, D. (1989). Theoretical study of the sensitivity of the microwave backscattering coefficient to the soil surface parameters. *Remote Sensing*, *10*(1), 171–179.

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., & Anderson, C. (2007). Initial soil moisture retrievals from the metop-a advanced scatterometer (ascat). *Geophysical Research Letters*, *34*(20).

Bousbih, S., Zribi, M., El Hajj, M., Baghdadi, N., Lili-Chabaane, Z., Gao, Q., & Fanise, P. (2018). Soil moisture and irrigation mapping in a semi-arid region, based on the synergetic use of sentinel-1 and sentinel-2 data. *Remote Sensing*, *10*(12), 1953.

Bousbih, S., Zribi, M., Lili-Chabaane, Z., Baghdadi, N., El Hajj, M., Gao, Q., & Mougenot, B. (2017). Potential of sentinel-1 radar data for the assessment of soil and cereal cover parameters. *Sensors*, *17*(11), 2617.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., et al. (2011). Soil moisture estimation through ascat and amsr-e sensors: An intercomparison and validation study across europe. *Remote Sensing of Environment*, *115*(12), 3390–3408.

Cheng, Y. (1995). Mean shift, mode seeking, and clustering. *IEEE transactions on pattern analysis and machine intelligence*, *17*(8), 790–799.

Cho, E., Moon, H., & Choi, M. (2015). First assessment of the advanced microwave scanning radiometer 2 (amsr2) soil moisture contents in northeast asia. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 117–129.

Copernicus Global Land Service. (2021). Lai [Available at https://land.copernicus.eu/global/, Accessed: 2021-12-13].

De Zan, F., & Gomba, G. (2018). Vegetation and soil moisture inversion from sar closure phases: First experiments and results. *Remote sensing of environment*, *217*, 562–572.

De Zan, F., Parizzi, A., Prats-Iraola, P., & López-Dekker, P. (2013). A sar interferometric model for soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(1), 418–425.

De Zan, F., Zonno, M., & Lopez-Dekker, P. (2015). Phase inconsistencies and multiple scattering in sar interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(12), 6608–6616.

Dostálová, A., Wagner, W., Milenković, M., & Hollaus, M. (2018). Annual seasonality in sentinel-1 signal for forest mapping and forest type classification. *International Journal of Remote Sensing*, *39*(21), 7738–7760.

El Hajj, M., Baghdadi, N., & Zribi, M. (2019). Comparative analysis of the accuracy of surface soil moisture estimation from the c-and l-bands. *International Journal of Applied Earth Observation and Geoinformation*, *82*, 101888.

El Hajj, M., Baghdadi, N., Zribi, M., & Bazzi, H. (2017). Synergic use of sentinel-1 and sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas. *Remote Sensing*, *9*(12), 1292.

Entekhabi, D., & Moghaddam, M. (2007). Mapping recharge from space: Roadmap to meeting the grand challenge. *Hydrogeology Journal*, *15*(1), 105–116.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., et al. (2010). The soil moisture active passive (smap) mission. *Proceedings of the IEEE*, *98*(5), 704–716.

Even, M., & Schulz, K. (2018). Insar deformation analysis with distributed scatterers: A review complemented by new advances. *Remote Sensing*, *10*(5), 744.

Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F., & Rucci, A. (2011). A new algorithm for processing interferometric data-stacks: Squeesar. *IEEE transactions on geoscience and remote sensing*, *49*(9), 3460–3470.

Flores-Anderson, A. I., Herndon, K. E., Thapa, R. B., & Cherrington, E. (2019). The sar handbook: Comprehensive methodologies for forest monitoring and biomass estimation.

Gabriel, A. K., Goldstein, R. M., & Zebker, H. A. (1989). Mapping small elevation changes over large areas: Differential radar interferometry. *Journal of Geophysical Research: Solid Earth*, *94*(B7), 9183–9191.

Gao, Q., Zribi, M., Escorihuela, M. J., & Baghdadi, N. (2017). Synergetic use of sentinel-1 and sentinel-2 data for soil moisture mapping at 100 m resolution. *Sensors*, *17*(9), 1966.

Gierull, C. H., & Sikaneta, I. C. (2002). Estimating the effective number of looks in interferometric sar data. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(8), 1733–1742.

Goldstein, R. M., & Werner, C. L. (1998). Radar interferogram filtering for geophysical applications. *Geophysical research letters*, *25*(21), 4035–4038.

Gomba, G., De Zan, F., Palmisano, D., Mattia, F., & Satalino, G. (2021). Potential for phase-based surface moisture estimation with c-band sar data. *Electronic Proceedings of the 13th European Conference on Synthetic Aperture Radar (2021)*, 1–3.

Grabocka, J., Schilling, N., Wistuba, M., & Schmidt-Thieme, L. (2014). Learning time-series shapelets. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 392–401.

Hensley, S., Michel, T., Van Zyl, J., Muellerschoen, R., Chapman, B., Oveisgharan, S., Haddad, Z. S., Jackson, T., & Mladenova, I. (2011). Effect of soil moisture on polarimetric-interferometric repeat pass observations by uavsar during 2010 canadian soil moisture campaign. *2011 IEEE International Geoscience and Remote Sensing Symposium*, 1063–1066.

Holah, N., Baghdadi, N., Zribi, M., Bruand, A., & King, C. (2005). Potential of asar/envisat for the characterization of soil surface parameters over bare agricultural fields. *Remote sensing of environment*, *96*(1), 78–86.

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of photogrammetry and remote sensing*, *80*, 91–106.

Jacob, A. W., Vicente-Guijalba, F., Lopez-Martinez, C., Lopez-Sanchez, J. M., Litzinger, M., Kristen, H., Mestre-Quereda, A., Ziółkowski, D., Lavalle, M., Notarnicola, C., et al. (2020). Sentinel-1 insar coherence for land cover mapping: A comparison of multiple feature-based classifiers. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 535–552.

Karimi, P., Bastiaanssen, W. G., & Molden, D. (2013). Water accounting plus (wa+)–a water accounting procedure for complex river basins based on satellite measurements. *Hydrology and Earth System Sciences*, *17*(7), 2459–2472.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., & Berger, M. (2001). Soil moisture retrieval from space: The soil moisture and ocean salinity (smos) mission. *IEEE transactions on Geoscience and remote sensing*, *39*(8), 1729–1735.

Khabbazan, S., Vermunt, P., Steele-Dunne, S., Ratering Arntz, L., Marinetti, C., van der Valk, D., Iannini, L., Molijn, R., Westerdijk, K., & van der Sande, C. (2019). Crop monitoring using sentinel-1 data: A case study from the netherlands. *Remote Sensing*, *11*(16), 1887.

Khalil, R. Z. et al. (2018). Insar coherence-based land cover classification of okara, pakistan. *The Egyptian Journal of Remote Sensing and Space Science*, *21*, S23–S28.

King, M. D., Platnick, S., Menzel, W. P., Ackerman, S. A., & Hubanks, P. A. (2013). Spatial and temporal distribution of clouds observed by modis onboard the terra and aqua satellites. *IEEE transactions on geoscience and remote sensing*, *51*(7), 3826–3852.

Lee, J.-S., Papathanassiou, K. P., Ainsworth, T. L., Grunes, M. R., & Reigber, A. (1998). A new technique for noise filtering of sar interferometric phase images. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(5), 1456–1465.

Li, J., & Wang, S. (2018). Using sar-derived vegetation descriptors in a water cloud model to improve soil moisture retrieval. *Remote Sensing*, *10*(9), 1370.

Li, Y.-y., Zhao, K., Ren, J.-h., Ding, Y.-l., & Wu, L.-l. (2014). Analysis of the dielectric constant of saline-alkali soils and the effect on radar backscattering coefficient: A case study of soda alkaline saline soils in western jilin province using radarsat-2 data. *The Scientific World Journal*, *2014*.

Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote sensing and image interpretation*. John Wiley & Sons.

López-Martínez, C., & Pottier, E. (2007). Coherence estimation in synthetic aperture radar data based on speckle noise modeling. *Applied optics*, *46*(4), 544–558.

Marino, A., Cloude, S. R., & Woodhouse, I. H. (2012). Detecting depolarized targets using a new geometrical perturbation filter. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(10), 3787–3799.

Martone, M., Rizzoli, P., & Krieger, G. (2016). Volume decorrelation effects in tandem-x interferometric sar data. *IEEE Geoscience and Remote Sensing Letters*, *13*(12), 1812–1816.

McDonald, K. C., Zimmermann, R., & Kimball, J. S. (2002). 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*, *40*(9), 2063–2082.

McNairn, H., & Brisco, B. (2004). The application of c-band polarimetric sar for agriculture: A review. *Canadian Journal of Remote Sensing*, *30*(3), 525–542.

Mironov, V. L., Kosolapova, L. G., Lukin, Y. I., Karavaysky, A. Y., & Molostov, I. P. (2017). Temperature- and texture-dependent dielectric model for frozen and thawed mineral soils at a frequency of 1.4 ghz. *Remote Sensing of Environment*, *200*, 240–249.

Morisette, J. T., Baret, F., Privette, J. L., Myneni, R. B., Nickeson, J. E., Garrigues, S., Shabanov, N. V., Weiss, M., Fernandes, R. A., Leblanc, S. G., et al. (2006). Validation of global moderate-resolution lai products: A framework proposed within the ceos land product validation subgroup. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(7), 1804–1817.

Morrison, K., Bennett, J. C., Nolan, M., & Menon, R. (2011). Laboratory measurement of the dinsar response to spatiotemporal variations in soil moisture. *IEEE transactions on geoscience and remote sensing*, *49*(10), 3815–3823.

Nikaein, T., Iannini, L., Molijn, R. A., & Lopez-Dekker, P. (2021). On the value of sentinel-1 insar coherence time-series for vegetation classification. *Remote Sensing*, *13*(16), 3300.

Nikolaeva, E., Sykioti, O., Elias, P., & Kontoes, C. (2015). Multi-temporal intensity and coherence analysis of sar images for land cover change detection on the island of crete. *SAR Image Analysis, Modeling, and Techniques XV*, *9642*, 964208.

Oon, A., Ngo, K. D., Azhar, R., Ashton-Butt, A., Lechner, A. M., & Azhar, B. (2019). Assessment of alos-2 palsar-2l-band and sentinel-1 c-band sar backscatter for discriminating between large-scale oil palm plantations and smallholdings on tropical peatlands. *Remote Sensing Applications: Society and Environment*, *13*, 183–190.

Ouaadi, N., Jarlan, L., Ezzahar, J., Zribi, M., Khabba, S., Bouras, E., Bousbih, S., & Frison, P.-L. (2020). Monitoring of wheat crops using the backscattering coefficient and the interferometric coherence derived from sentinel-1 in semi-arid areas. *Remote Sensing of Environment*, *251*, 112050.

Palmisano, D., Satalino, G., Balenzano, A., Bovenga, F., Mattia, F., Rinaldi, M., Ruggieri, S., Skriver, H., Davidson, M. W., Cartus, O., et al. (2019). Sensitivity of sentinel-1 interferometric coherence to crop structure and soil moisture. *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, 6219–6222.

Parizzi, A., Perissin, D., Prati, C., Rocca, F., & Ferretti, A. (2006). Accurate dem reconstruction from permanent scatterers and multi-baseline interferometry. *2006 IEEE International Symposium on Geoscience and Remote Sensing*, 157–160.

Samiei-Esfahany, S., & Hanssen, R. (2017). On the evaluation of second order phase statistics in sar interferogram stacks. *Earth Observation and Geomatics Engineering*, *1*(1), 1–15.

Santoro, M., Askne, J. I., Wegmuller, U., & Werner, C. L. (2007). Observations, modeling, and applications of ers-envisat coherence over land surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(8), 2600–2611.

Sekertekin, A., Marangoz, A. M., & Abdikan, S. (2020). Alos-2 and sentinel-1 sar data sensitivity analysis to surface soil moisture over bare and vegetated agricultural fields. *Computers and Electronics in Agriculture*, *171*, 105303.

Steele-Dunne, S. C., Hahn, S., Wagner, W., & Vreugdenhil, M. (2019). Investigating vegetation water dynamics and drought using metop ascat over the north american grasslands. *Remote Sensing of Environment*, *224*, 219–235.

Steele-Dunne, S. C., McNairn, H., Monsivais-Huertero, A., Judge, J., Liu, P.-W., & Papathanassiou, K. (2017). Radar remote sensing of agricultural canopies: A review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(5), 2249–2273.

Thoma, D., Moran, M., Bryant, R., Rahman, M., Collins, C. H., Keefer, T., Noriega, R., Osman, I., Skrivin, S., Tischler, M., et al. (2008). Appropriate scale of soil moisture retrieval from high resolution radar imagery for bare and minimally vegetated soils. *Remote Sensing of Environment*, *112*(2), 403–414.

Torbick, N., Chowdhury, D., Salas, W., & Qi, J. (2017). Monitoring rice agriculture across myanmar using time series sentinel-1 assisted by landsat-8 and palsar-2. *Remote Sensing*, *9*(2), 119.

Touzi, R., & Lopes, A. (1996). Statistics of the stokes parameters and of the complex coherence parameters in one-look and multilook speckle fields. *IEEE Transactions on Geoscience and Remote Sensing*, *34*(2), 519–531.

Touzi, R., Lopes, A., Bruniquel, J., & Vachon, P. W. (1999). Coherence estimation for sar imagery. *IEEE transactions on geoscience and remote sensing*, *37*(1), 135–149.

Ulaby, F. T., Batlivala, P. P., & Dobson, M. C. (1978). Microwave backscatter dependence on surface roughness, soil moisture, and soil texture: Part i-bare soil. *IEEE Transactions on Geoscience Electronics*, *16*(4), 286–295.

Ulaby, F. T., Moore, R. K., & Fung, A. K. (1986). Microwave remote sensing: Active and passive. volume 3-from theory to applications.

Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., & Strauss, P. (2018). Sensitivity of sentinel-1 backscatter to vegetation dynamics: An austrian case study. *Remote Sensing*, *10*(9), 1396.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldana, J., De Rosnay, P., Jann, A., Schneider, S., et al. (2013). The ascat soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteorologische Zeitschrift*.

Wagner, W., Scipal, K., Pathe, C., Gerten, D., Lucht, W., & Rudolf, B. (2003). Evaluation of the agreement between the first global remotely sensed soil moisture data with model and precipitation data. *Journal of Geophysical Research: Atmospheres*, *108*(D19).

Wegmüller, U., Santoro, M., Werner, C., & Cartus, O. (2015). On the estimation and interpretation of sentinel-1 tops insar coherence. *Proc. of FRINGE, ESA SP-731 (CD-ROM), ESA Publications Division, European Space Agency, Noordwijk, The Netherlands*.

Yebra, M., Dennison, P. E., Chuvieco, E., Riano, D., Zylstra, P., Hunt Jr, E. R., Danson, F. M., Qi, Y., & Jurdao, S. (2013). A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. *Remote Sensing of Environment*, *136*, 455–468.

Zebker, H. A., & Villasenor, J. (1992). Decorrelation in interferometric radar echoes. *IEEE Transactions on geoscience and remote sensing*, *30*(5), 950–959.

Zhang, T., Zeng, Q., Li, Y., & Xiang, Y. (2008). Study on relation between insar coherence and soil moisture. *Proceedings of the ISPRS Congress, Beijing, China*, 3–11.

Zwieback, S., Bartsch, A., Melzer, T., & Wagner, W. (2011). Probabilistic fusion of ku and c-band scatterometer data for determining the freeze/thaw state. *IEEE transactions on geoscience and remote sensing*, *50*(7), 2583–2594.

Zwieback, S., Hensley, S., & Hajnsek, I. (2015). A polarimetric first-order model of soil moisture effects on the dinsar coherence. *Remote Sensing*, *7*(6), 7571–7596.

Zwieback, S., Hensley, S., & Hajnsek, I. (2017). Soil moisture estimation using differential radar interferometry: Toward separating soil moisture and displacements. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(9), 5069–5083.

Zwieback, S., Liu, X., Antonova, S., Heim, B., Bartsch, A., Boike, J., & Hajnsek, I. (2016). A statistical test of phase closure to detect influences on dinsar deformation estimates besides displacements and decorrelation noise: Two case studies in high-latitude regions. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(9), 5588–5601.

A

Sampling Histograms

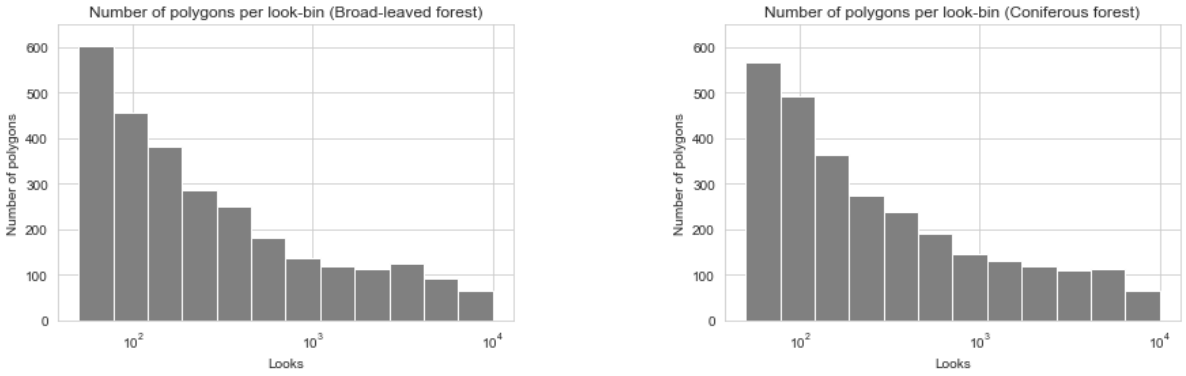


Figure A.1: Histogram for polygons per looks-bin for the land cover types broadleaved-forest (left) and coniferous forest (right)

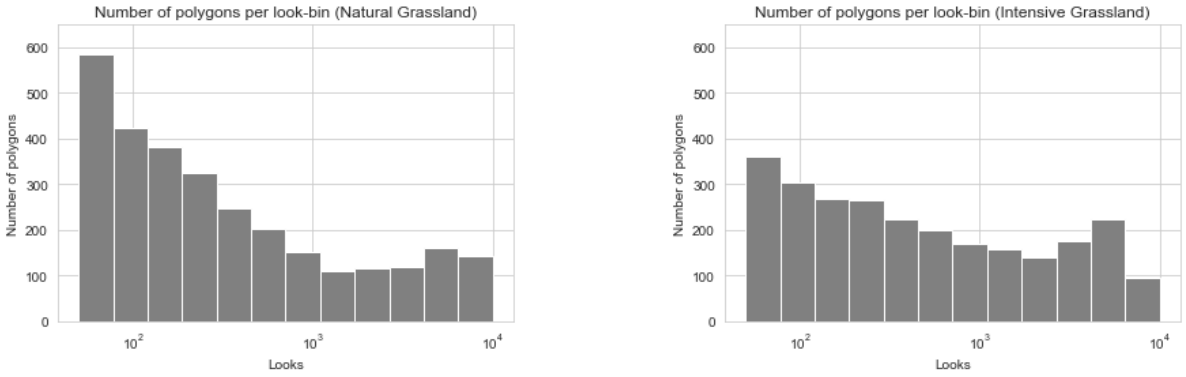


Figure A.2: Histogram for polygons per looks-bin for the land cover types natural grassland (left) and intensive grassland (right)

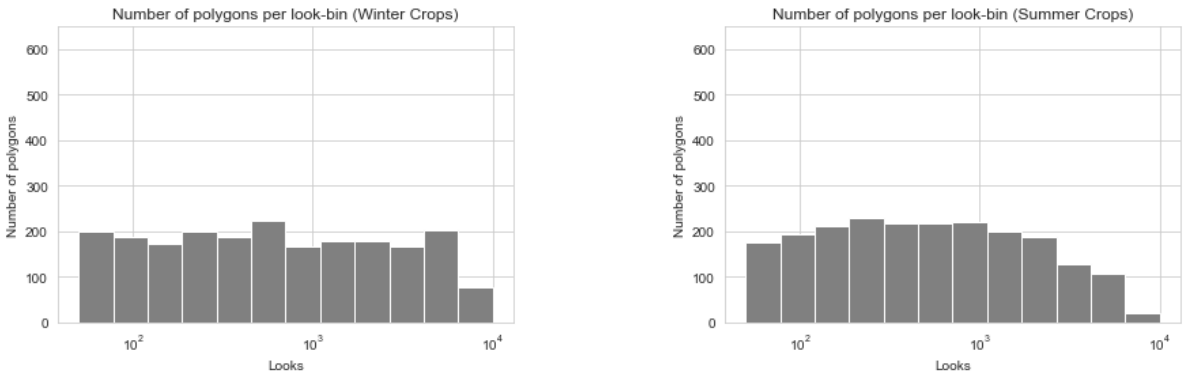


Figure A.3: Histogram for polygons per looks-bin for the land cover types winter crops (left) and summer crops (right)

B

Land Cover Nomenclature Translation

This table presents a comparison between 2017, 2018 and 2019 nomenclature used by Theia for their 2017, 2018 and 2019 land cover maps. The summer and winter crops are further divided in separate crops. The remaining land cover classes are consistent between the years.

Table B.1: Comparison between Theia's land cover nomenclature for 2017 versus 2018/2019

Theia Land Cover Nomenclature	
2017	**2018 and 2019**
Continuous Urban Fabric	Continuous Urban Fabric
Discontinuous Urban Fabric	Discontinuous Urban Fabric
Industrial and Commercial Units	Industrial and Commercial Units
Road Surfaces	Road Surfaces
Annual Winter Crops	Rapeseed
Annual Winter Crops	Straw Cereals
Annual Winter Crops	Protein Crops
Annual Summer Crops	Soy
Annual Summer Crops	Sunflower
Annual Summer Crops	Corn
Annual Summer Crops	Rice
Annual Summer Crops	Tubers/Roots
Intensive Grasslands	Intensive Grasslands
Orchards	Orchards
Vineyards	Vineyards
Broad-leaved Forest	Broad-leaved Forest
Coniferous Forests	Coniferous Forests
Natural Grasslands	Natural Grasslands
Woody Moorlands	Woody Moorlands
Mineral Surfaces	Mineral Surfaces
Beaches and Dunes	Beaches and Dunes
Glaciers or Snow	Glaciers or Snow
Water	Water

C

Precipitation Stations and Data

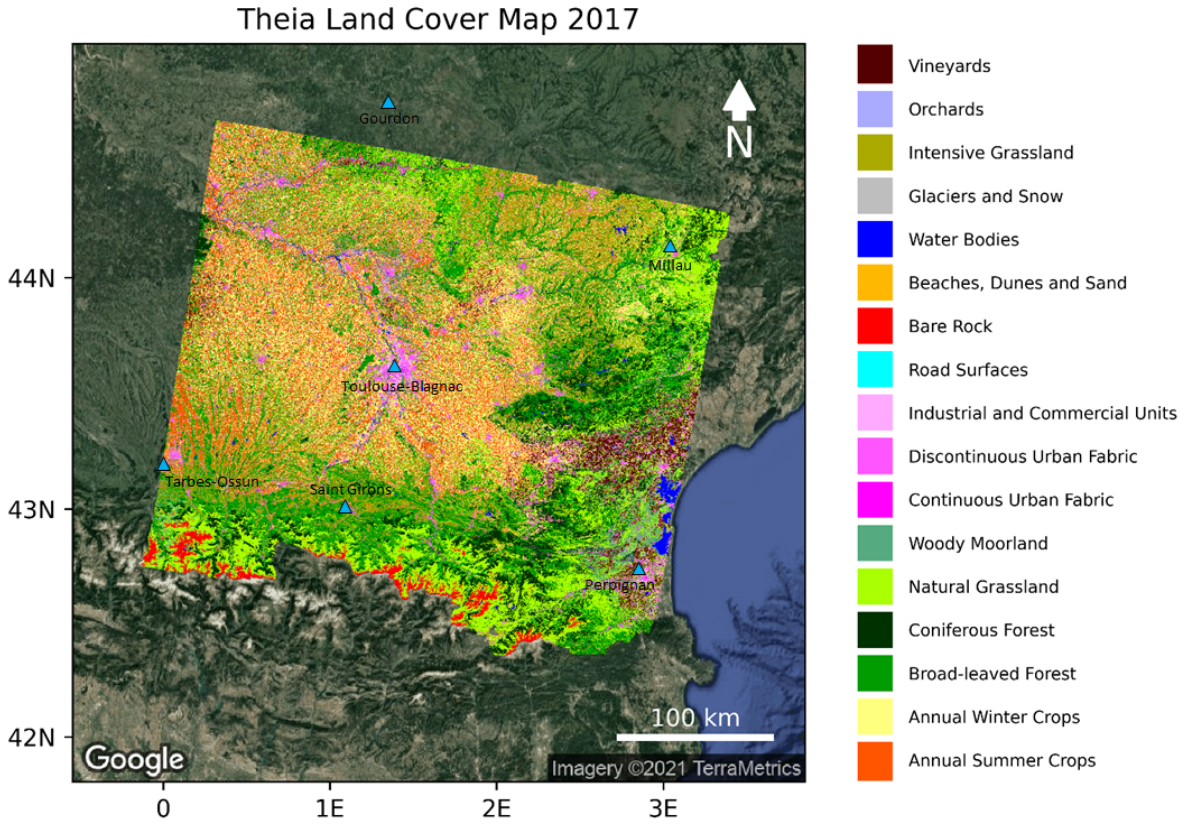


Figure C.1: Theia land cover map showing the location of all precipitation stations.

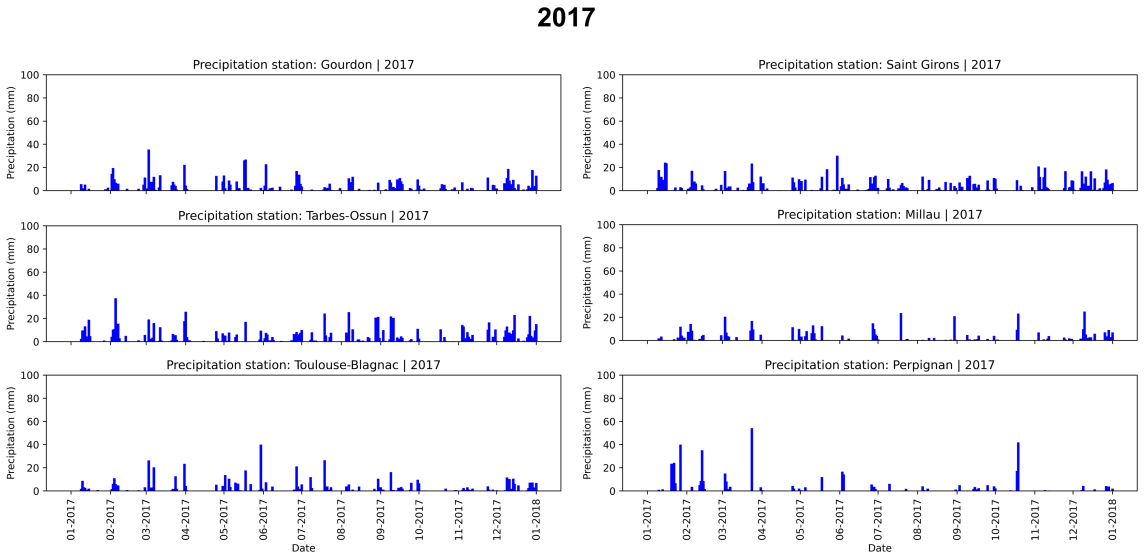


Figure C.2: 2017 daily precipitation recorded by rain gauges in Occitanie.

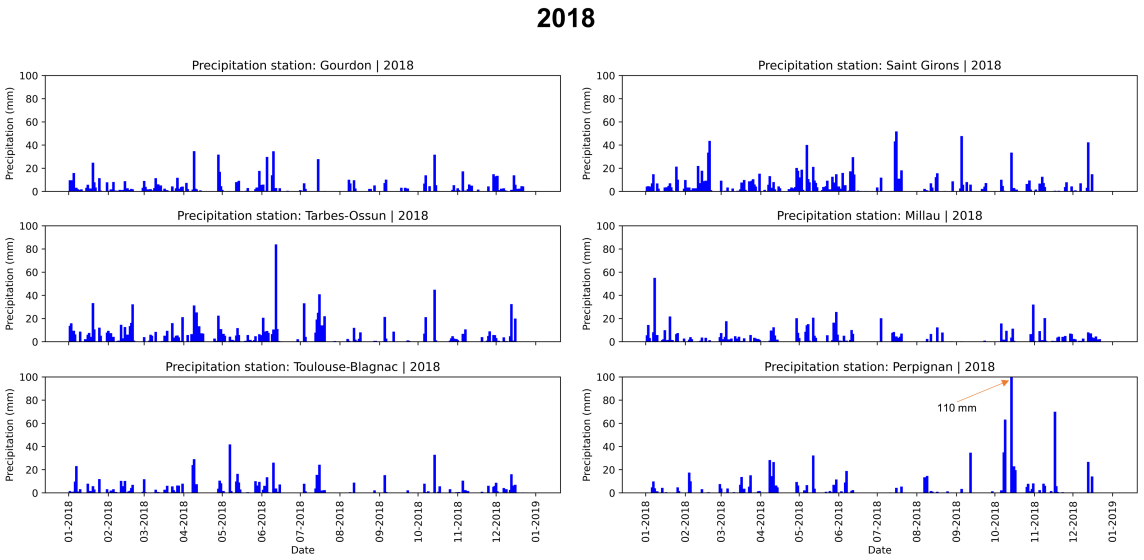


Figure C.3: 2018 daily precipitation recorded by rain gauges in Occitanie.

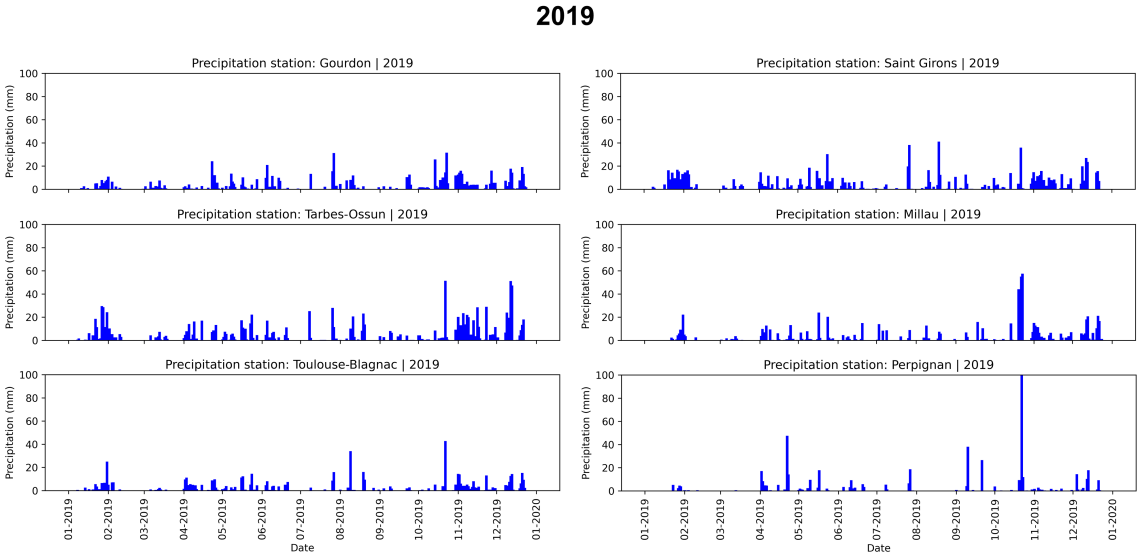


Figure C.4: 2019 daily precipitation recorded by rain gauges in Occitanie.