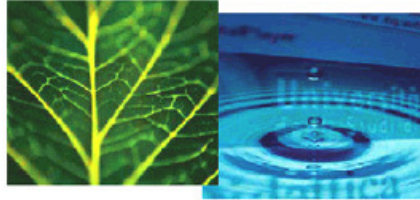


PhD Dissertation



International Doctorate School in Information and
Communication Technologies

DISI - University of Trento

ADVANCED REGRESSION AND DETECTION METHODS FOR
REMOTE SENSING DATA ANALYSIS

Davide Castelletti

Advisor:

Lorenzo Bruzzone

Co-advisor:

Begüm Demir

April 2017

Abstract

Nowadays the analysis of remote sensing data for environmental monitoring is fundamental to understand the local and global Earth dynamics. In this context, the main goal of this thesis is to present novel signal processing methods for the estimation of biophysical parameters and for the analysis icy terrain with active sensors. The thesis presents three main contributions.

In the context of biophysical parameters estimation we focus on regression methods. According to the analysis of the literature, most of the regression techniques require a relevant number of reference samples to model a robust regression function. However, in real-world applications the ground truth observations are limited as their collection leads to high operational cost. Moreover, the availability of biased samples may result in low estimation accuracy. To address these issues, in this thesis we propose two novel contributions. The first contribution is a method for the estimation of biophysical parameters that integrates theoretical models with empirical observations associated to a small number of in-situ reference samples. The proposed method computes and correct deviations between estimates obtained through the inversion of theoretical models and empirical observations. The second contribution is a semisupervised learning (SSL) method for regression defined in the context of the ε -insensitive SVR. The proposed SSL method aims to mitigate the problems of small-sized biased training sets by injecting priors information in the initial learning of the SVR function, and jointly exploiting labeled and unlabeled samples in the learning phase of the SVR.

The third contribution of this dissertation addresses the clutter detection problem in radar sounder (RS) data. The capability to detect clutter is fundamental for the interpretation of subsurface features in the radargram. In the state of the art, techniques that require accurate information on the surface topography or approaches that exploit complex multi-channel radar sounder systems have been presented. In this thesis, we propose a novel method for clutter detection that is independent from ancillary information and limits the hardware complexity of the radar system. The method relies on the interferometric analysis of two-channel RS data and discriminates the clutter and subsurface echoes by modeling the theoretical phase difference between the cross-track antennas of the RS. This allows the comparison of the phase difference distributions of real and simulated data.

Qualitative and quantitative experimental results obtained on real airborne SAR and RS data confirm the effectiveness of the proposed methods.

Keywords[Remote sensing, signal processing, biophysical parameters estimation, support vector regression, synthetic aperture radar, radar sounder, clutter detection.]

Contents

1	Introduction	1
1.1	Background and motivation	1
1.2	Objectives and novel contributions of the thesis	4
1.3	Biophysical parameter estimation	5
1.4	Clutter detection in radar sounder signals	6
1.5	Structure of the Thesis	7
2	Fundamentals and background	9
2.1	Introduction to microwave remote sensing systems	9
2.1.1	Radar	9
2.1.2	Synthetic aperture radar	11
2.1.3	Radar interferometry	13
2.1.4	Radar sounder	16
2.2	Overview on approaches for biophysical parameters estimation	19
2.2.1	Biophysical parameters estimation problem	20
2.2.2	Derivation of empirical data-driven relationships using empirical methods	20
2.2.3	Inversion of physical based theoretical models	21
2.3	Overview on clutter in radar sounder data	22
2.3.1	Clutter problem definition	23
2.3.2	Clutter suppression techniques	24
2.3.3	Detection techniques	25
3	A novel hybrid method for the correction of the theoretical model in- version in bio/geophysical parameter estimation	27
3.1	Introduction	27
3.2	Proposed hybrid method	30
3.2.1	Computation and modeling of deviations	31
3.2.2	Estimation of the final target values	33

3.3	Datasets description and experimental setup	34
3.4	Experimental results	36
3.4.1	Results on the Scatt dataset	36
3.4.2	Results on the SMEX dataset	39
3.4.3	Results on the ActPass dataset	41
3.5	Conclusion	43
4	A novel semisupervised method for the estimation of biophysical parameters based on Support Vector Regression	45
4.1	Introduction	46
4.2	Problem definition	49
4.3	Proposed SSL method for regression	50
4.3.1	Injecting priors information in the initial learning of the SVR function	51
4.3.2	Exploitation of both labeled and unlabeled samples in the learning process of the regression function	54
4.4	Dataset description and design of the experiments	58
4.5	Experimental results	60
4.5.1	Paneveggio dataset: estimation of stem diameter	60
4.5.2	Abalone dataset: age of abalone	62
4.6	Conclusion	64
5	A novel technique based on interferometry for clutter detection in two-channel radar sounder data	67
5.1	Introduction	67
5.2	Proposed technique for cross track clutter detection	70
5.2.1	Manual feature extraction and theoretical phase difference estimation	72
5.2.2	Interferogram formation for cross-channel phase difference	74
5.2.3	Comparison of theoretical and real phase difference distributions	76
5.3	Dataset description and experimental results	77
5.3.1	MARFA radar system description	78
5.3.2	Dataset 1: deep subsurface	80
5.3.3	Dataset 2: Clutter feature detection	83
5.3.4	Dataset 3: Clutter and subsurface discrimination	85
5.4	Cross track clutter detection in planetary RS	90
5.5	Discussion and conclusion	94
6	Conclusions and future work	95
6.1	Research context and summary of novel contributions	95

6.2	Concluding remarks and future developments	98
7	List of Publications	101
7.1	Journal Papers	101
7.2	Conferences	101
	Bibliography	103

List of Tables

3.1	Main characteristics of the three datasets considered in the experimental analysis.	35
3.2	Estimation performances obtained by the proposed hybrid approach with the GDB and LDB strategies and the standard inverse theoretical model (Scatt dataset).	37
3.3	Estimation performances obtained by the proposed hybrid approach with the GDB and LDB strategies and the standard inverse theoretical model (SMEX dataset).	39
3.4	Estimation performances obtained by the proposed hybrid approach with the GDB and LDB strategies and the standard inverse theoretical model (ActPass dataset).	41
4.1	The list of variables used in the experiments and their physical meanings for the estimation of stem diameter.	58
4.2	The list of variables used in the experiments and their physical meanings for the estimation of abalones ages.	59
4.3	Mean and standard deviation (std) of MAE and R^2 obtained by the standard SVR and proposed SS-SVR.	61
4.4	Results of the first step of proposed method obtained when the value of radius is estimated either by the proposed strategy or by randomly fixing its value.	61
4.5	Mean and standard deviation (std) of MAE and R^2 obtained by executing the proposed SS-SVR with and without the 2nd step.	62
4.6	Mean and standard deviation (std) of MAE and R^2 obtained by the standard SVR and proposed SS-SVR.	63
4.7	Comparison of the results of the first step of proposed method obtained when the value of radius is estimated by either the proposed strategy or by randomly fixing its value.	63

4.8	Mean and standard deviation (std) of MAE and R^2 obtained by executing the proposed SS-SVR with and without the 2nd step.	63
5.1	Typical operating parameters for MARFA RS.	79
5.2	SAD obtained between the real and the simulated (on both nadir and clutter hypothesis) histograms and mean of the simulated distributions. . .	82
5.3	SAD obtained on real and simulated (on both nadir and clutter hypothesis) histograms and mean of the simulated distributions for the selected feature.	85
5.4	SAD values obtained between the real and the simulated (on both nadir and clutter hypothesis) histograms and the mean of the simulated distributions for feature f_1	89
5.5	SAD values obtained between the real and the simulated (on both nadir and clutter hypothesis) histograms and mean of the simulated distributions for feature f_2	89

List of Figures

2.1	Monostatic radar configuration.	10
2.2	Principle of SAR.	11
2.3	SAR acquisition geometry.	12
2.4	Interferometric radar configurations.	15
2.5	Exact formula of the interferometric phase in the cross-track single-pass configuration.	16
2.6	Nadir acquisition geometry of radar sounder systems.	17
2.7	Geometry of the clutter problem.	23
2.8	Surface clutter contribution before SAR processing (light gray locus), and after focusing in azimuth direction by Doppler processing and in range by range compression (dark gray cells). This example assumes flat topography in the absence of airplane roll.	24
3.1	Block diagram of the proposed hybrid approach.	31
3.2	Scatterplots of measured versus estimated dielectric constant values obtained by (a) the inverse theoretical model, (b) the proposed hybrid approach with the GDB strategy, and (c) the proposed hybrid approach with the LDB strategy (Scatt dataset with HH23 features).	38
3.3	Scatterplots of measured versus estimated dielectric constant values obtained by (a) the inverse theoretical model, (b) the proposed hybrid approach with the GDB strategy, and (c) the proposed hybrid approach with the LDB strategy (SMEX dataset).	40
3.4	Scatterplots of measured versus estimated dielectric constant values obtained by (a) the inverse theoretical model, (b) the proposed hybrid approach with the GDB strategy, and (c) the proposed hybrid approach with the LDB strategy (ActPass dataset with the HHVV feature).	42
4.1	Block scheme of the proposed SSL method.	51

4.2	A qualitative example in a 1-D feature space: a) distribution of the biased data; the unlabeled samples are the squares in green and the labeled samples are the circles in black. b) Radius definition to estimate the weight of label samples. c) the WSVR learning function.	53
4.3	A qualitative example in a 1-D feature space. (a) the WSVR learning function obtained at the first step; (b) selection of SVs that lie on the ε -tube after elimination of samples out and far from the tube (c) selection of informative unlabeled samples having largest confidence and relevancy to be correctly estimated.	57
5.1	2-D representation of the acquisition geometry of a RS instrument.	70
5.2	Block diagram of the proposed technique.	71
5.3	Example of manual feature selection. a) The red area corresponds to the set of pixels that belong to a feature of interest. b) The yellow line is the simple feature-model used to project the feature onto the surface in the cross-track direction.	72
5.4	Two-channel sounding geometry. The red dots are reflection points on the surface. For echoes coming from nadir, the signal path length ρ_{chl} and ρ_{chr} are equivalent and $\Delta_{\Phi_t} = 0$. The red dotted line represents an example of off-nadir signal path (for a clutter feature with the same delay), in which the difference between ρ_{chl} and ρ_{chr} makes $\Delta_{\Phi_t} \neq 0$	73
5.5	Qualitative illustration of the comparison between theoretical and real phase difference distributions. The three different scenarios are shown. In red and in blue the real and theoretical phase distributions respectively are presented.	76
5.6	Location of the RS tracks collected by MARFA on the Greenland ice shell.	78
5.7	Block scheme of the MARFA radar system.	79
5.8	Deep subsurface dataset (track NAQLK/JKB2j/ZY1b); optical imagery acquired on Greenland (DigitalGlobe).	80
5.9	Deep subsurface dataset (track NAQLK/JKB2j/ZY1b): (a) Amplitude radargram obtained from a single channel (left) of MARFA radar (11.5 km long portion of the entire acquisition), (b) RS interferogram obtained with the proposed method.	81
5.10	Feature shown in Figure 5.9a: (a) theoretical phase difference histogram for nadir hypothesis, (b) clutter hypothesis and (c) real phase difference histogram extracted from the real interferogram.	82
5.11	Clutter feature detection dataset (track GOG/JKB2j/BWN01b); optical image acquired on Greenland (DigitalGlobe).	83

5.12	Clutter feature detection dataset (track GOG/JKB2j/BWN01b); (a) amplitude radargram obtained from a single channel (left) of MARFA radar (10 km long portion of the entire acquisition), (b) RS interferogram obtained with the proposed method.	84
5.13	Clutter feature detection dataset (track GOG/JKB2j/BWN01b); zoomed portion of the radargram containing the selected feature.	84
5.14	Feature shown in Figure 5.11: (a) theoretical phase difference histogram for nadir hypothesis, (b) clutter hypothesis and (c) real phase difference histogram extracted from the interferogram.	85
5.15	Clutter and subsurface discrimination dataset (track GOG/JKB2j/BWN01a); optical image acquired on Greenland (DigitalGlobe).	86
5.16	Clutter and subsurface discrimination dataset (track GOG/JKB2j/BWN01a); (a) amplitude radargram obtained from a single channel (left) of MARFA radar (8 km long portion of the entire acquisition), (b) RS interferogram obtained with the proposed method.	87
5.17	Clutter and subsurface discrimination dataset (track GOG/JKB2j/BWN01a); zoomed portion of the radargram containing the selected features.	87
5.18	Feature f_1 shown in Figures 5.15 and 5.17: (a) theoretical phase difference histogram for nadir hypothesis and (c) clutter hypothesis and (e) real phase difference histogram. Feature f_2 shown in Figures 5.15 and 5.17: (b) theoretical phase difference histogram for nadir hypothesis and (d) clutter hypothesis and (f) real phase difference histogram.	88
5.19	Theoretical phase difference behavior for a single range line in the case of a) phase difference expected at low altitude, with $H_0 = 1000$ m and $B = 19$ (MARFA airborne radar configuration); b) phase difference expected for satellite application, with $H_0 = 400$ km and $B = 10$ m.	91
5.20	Configuration of the simulation for RS orbiter case: (a) digital elevation model of the Europa icy moon and synthetic clutter feature. (b) Simulated radargram.	92
5.21	Planetary two-channel RS data simulation at Europa (10 meters baseline and 400 km orbit altitude): (a) radargram and (b) interferogram. (c) Phase difference histogram for the magenta feature (surface echo). (d) Phase difference histogram for the yellow feature (clutter).	93
5.22	Planetary two-channel RS data simulation at Europa (10 meters baseline and 400 km orbit altitude): (a) radargram and (b) interferogram. (c) Phase difference histogram for the magenta feature (surface echo). (d) Phase difference histogram for the yellow feature (clutter).	93

Chapter 1

Introduction

The aim of this first Chapter is to introduce the PhD thesis work. In particular, it presents the context in which estimation and detection methods are exploited to analyze remote sensing images. This allows to state the aim of this work and to highlight its novel contributions. Finally, the structure of the whole document is outlined.

1.1 Background and motivation

In the last decade the interest in remote sensing data analysis for the Earth Observation (EO) is increased. Techniques for EO data analysis have strong impact in many research areas having as a major concern the interaction of humans with environment. For example, the study of biophysical parameters and ice sheets dynamics are fields of study of crucial importance in the scientific community. The techniques for the estimation of biophysical terrain parameters and for the analysis of Polar regions of Earth are among the most relevant branch of EO because they contribute to provide key environmental indicators. In this scenario, the availability of a large amount of information acquired by cutting-edge active and passive sensors encourages the definition of methods to analyze the remote sensing data. In this thesis, we focus our attention on signal processing methods for the estimation of biophysical parameters and for the analysis icy terrain with active sensors. Concerning the estimation problem, we aim to provide improved solutions to tackle the scarcity of ground truth information, which is a fundamental resource for the retrieval of biophysical parameters. Regarding the analysis of icy terrains we focus on the definition of methods for the detection of clutter in Radar Sounder (RS) data. In the following we give some background on these two topics.

Biophysical parameters estimation

In remote sensing data analysis, the estimation of biophysical parameters is relevant to the

understanding of environment dynamics at local and global scales. Soil moisture content (SMC), leaf chlorophyll content, temperature, leaf area index, are some examples of the most important parameters related to surface terrain that can be retrieved by remote sensing data. Every year international space agencies and private companies launch satellite missions with innovative payloads for EO. Some examples of recent missions are: Sentinel 1 (Snoeij et al. [2010]), one of the missions of the European Space Agency' Copernicus Program; the Soil Moisture Active Passive (SMAP) mission (Entekhabi et al. [2010]); and Radarsat-2 (Rolland et al. [2012]), a Canadian commercial satellite. In more details, Sentinel 1 carries single or dual polarization C-Band Synthetic Aperture Radar (SAR) instrument to provide an all-weather, day-and-night supply of imagery relevant to several monitoring applications (e.g., Arctic sea-ice extent, the marine environment including oil-spill, land-surface for motion risks, forest, water and soil management). SMAP is a mission for soil moisture content estimation at global scale. Radarsat-2 is satellite carrying a quad polarization C-Band SAR having common characteristics with Sentinel but different operatives modes. These sensors made available new data, which are relevant to the estimation of biophysical parameters.

These parameters are continuous variables of fundamental importance in several application domains. For instance, for natural disasters prevention, such as drought, flooding and landslides (Sandholt et al. [2002], Engman and Gurney [1991]) a precise knowledge of SMC in superficial terrain layers is needed. SMC concentration is also a parameter required in precision farming, which relies on indicators estimated from satellite data to support decision for the management of the farm site with the goal of optimize returns while preventing the waste of resources (Bolten et al. [2010], Heathman et al. [2003]). From a physical perspective, the change of these parameters, which depends on the chemical and structural characteristics of a target, corresponds to a variation of the electromagnetic response that can be measured by the sensor. Absorption, transmission and reflection (Ulaby and Long [2014]) of an electromagnetic wave varies according to the characteristics of the observed target. Ultimately, the estimation of parameters can be treated as a regression problem. From an analytical view point a regression problem can be solved by designing a mapping function being capable to couple the parameter to be estimated (i.e., target or label) with a set of feature extracted from the remote sensing data (e.g., backscattered power). However, most of the regression techniques require ground truth data (i.e., reference samples) to build a reliable mapping function. As an example, supervised regression algorithms need good quality and quantity of the reference samples to model a precise and robust mapping function. In real-word applications the limited disposability of reference samples or the availability of bias ground truth information are major concerns when a regression algorithm is exploited.

Considering the criticality of the estimation of biophysical parameters in the remote sensing field, we focus on the definition of novel solutions to improve the regression performance when the quality or the quantity of ground truth samples is limited or when bias reference samples are available. Moreover, we aim to define general methodologies, thus applicable to heterogeneous data sources.

Clutter detection in radar sounder data

Another important topic in the context of environment monitoring is the analysis of icy terrains with active sensors data. Ice sheets dynamic is among the most relevant topic for glaciologist and climate change scientists. Indeed, radioglaciology uses remote sensing data to study sub-glacial conditions of rapidly changing ice sheets and their potential contribution to the rate of sea level rise. In particular, an effective way to study the ice subsurface on wide areas is exploiting the RS data (Bogorodsky et al. [1984], Fu and Daniels [2007]), called radargrams. RS is a nadir-looking active sensor that transmits a low-frequency electromagnetic wave towards the surface and measures the reflected power from both surface and subsurface with a penetration capability that depends on the electromagnetic characteristic of the conductive medium and on the radar design itself. Concerning EO, RS acquisitions are operated during dedicated airborne campaigns at the Earth polar regions. Some examples of RSs that acquire ice sheet information on Greenland and Antarctica are: Multifrequency Airborne Radar Sounder with Full-phase Assessment (MARFA) operated by the University of Texas Institute for Geophysics (UTIG) (Young et al. [2016]), Multichannel Coherent Radar Depth Sounder (MCoRDS) operated by Center for Remote Sensing of Ice Sheets (CReSIS) (Shi et al. [2010]) and POLARIS an ESA' s Ice Sounding Radar (Dall et al. [2010b]). The unique results obtained on Earth polar region from these sensors and the success of the existing planetary RS instruments, under development as Radar for Icy Moon Exploration (RIME) on the ESA JUICE mission (Bruzzone et al. [2015]), or the already operating mars missions as the Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS) and the SHallow RADar (SHARAD) (Jordan et al. [2009], Croci et al. [2011]) have encouraged dedicated studies for defining an EO RS missions on satellite platform that can enrich time repetition and space coverage of airborne surveys.

All the RS systems mentioned above need an efficient way to tackle the clutter problem, which intrinsically depends on the acquisition geometry of RSs. Due to the non-ideal directivity of the antenna beam, the received signals can contain components coming from off-nadir directions at the same time with the subsurface nadir returns. In the radargram, the clutter returns can overlap the returns coming from the subsurface, leading to partial or complete masking of the subsurface features. According to the solutions presented in

the state of the art, the clutter can be discriminated from the nadir return using two different approaches: i) designing multi-channel RSs that permit to use very directive antenna or to exploit the cross-channel phase, ii) exploit modeling using precise information about the surface topography. However, these approaches have drawbacks. Multichannel systems require high complexity hardware system design and large data volume resources to independently record the echoes received from the channels. These problems are particularly relevant in the design of satellite RSs for which resources (e.g., mass, volume, power and data volume allocation) are limited. The second class of approaches involves the use of a digital elevation model (DEM) of the surface to simulate synthetically radargrams and predict the position of clutter coming from the surface. However, in some cases DEMs, having sufficient spatial resolution to properly represent features in the radargram are not available. Moreover, these techniques can only identify surface echoes, without providing discrimination of clutter generated from a structure buried in the ice subsurface. This clutter problem can have an important impact on the science return that a RS system can gather, preventing the capability to analyze part of the subsurface features. For this reason, it is important to define effective and efficient solutions to detect the clutter contribution in the radargram.

1.2 Objectives and novel contributions of the thesis

The main goal of this thesis is to present novel signal processing methods for the analysis of remote sensing data. As mentioned above, both the estimation of biophysical terrain parameters and the analysis of icy terrains are fundamental issues that can be addressed through the analysis of remote sensing data acquired by active sensors.

Concerning the development of regression methods for the estimation of biophysical parameters, we aim to overcome one of the limitations of the current state-of-the-art methods. We address parameter estimation problems when a small quantity of reference samples are available, or when the reference samples set is bias. In this context, we propose two novel regression methods for the estimation biophysical parameters:

1. A novel hybrid method for the correction of the theoretical model inversion in bio/geophysical parameter estimation.
2. A novel semisupervised method for the estimation of biophysical parameters based on Support Vector Regression.

In the context of the analysis of icy terrains with radar, we propose a novel technique based on interferometry for clutter detection in two-channel RS data. The use of a two-

channel RS system allows the reduction of the hardware complexity with respect to the multi-channels systems currently used. The proposed method exploits a single-pass interferometric approach to discriminate the surface clutter from nadir subsurface return without the need of ancillary information about the surface topography (e.g., DEM).

In the following sections, we describe in greater detail the main novel contributions of the thesis. The former is devoted to the estimation of biophysical parameters, while the latter refers to the detection of clutter in RS data.

1.3 Biophysical parameter estimation

The main novel contributions of the thesis related to the analysis of remote sensing data for biophysical parameter estimation can be summarized as follows.

A novel hybrid method for the correction of the theoretical model inversion in bio/geophysical parameter estimation

The first contribution is a novel method for the estimation of biophysical parameters that integrates theoretical models with empirical observations associated to a small number of field reference samples. To this aim, we developed a novel hybrid approach that models and corrects deviations from correct target values when theoretical electromagnetic models are used for the inversion process. This is achieved in two steps. In the first step, deviations between estimations obtained by a theoretical model and empirical observations are initially computed. Then, deviations associated to unlabeled samples (for which reference measures are not existing) are characterized based on two different strategies: i) the global deviation bias (GDB) strategy (which assumes that the deviations of samples are constant within the input space); and ii) the local deviation bias (LDB) strategy (which assumes that the deviations of samples are variable within different portions of the input space). In the second step theoretical model estimates of unlabeled samples are corrected based on the estimated deviations. The experimental analysis carried out in the context of soil moisture content retrieval from microwave remotely sensed data confirm the effectiveness of the proposed hybrid estimation approach.

A novel semisupervised method for the estimation of biophysical parameters based on Support Vector Regression

The second contribution is a novel semisupervised learning (SSL) technique defined in the context of ϵ -insensitive SVR for parameter estimation. The proposed SSL method

aims to mitigate the problems of small-sized biased training sets without collecting any additional sample with reference measures. This is achieved on the basis of two consecutive steps. The first step is devoted to inject additional prior information in the learning phase of SVR in order to adapt the importance of each training sample to the distribution of the unlabeled samples. To this end, a weight is initially associated to each training sample based on a strategy that defines higher weights for the samples located in the high density regions of the feature space, while gives reduced weights to those that fall into low density regions of the feature space. Then, in order to exploit different weights for training samples in the learning phase of the SVR, we consider the weighted SVR (WSVR) algorithm. The second step is devoted to jointly exploit labeled and informative unlabeled samples for further improving the definition of the WSVR learning function. To this end, the most informative unlabeled samples that have an expected reliable target values are selected according to a novel strategy that relies on the distribution of the unlabeled samples in the feature space and on the WSVR function estimated at the first step. Then, we introduce a redefined WSVR algorithm that jointly uses labeled and unlabeled samples in the learning phase of the WSVR algorithm and tunes their importance by different values of regularization parameters. Experimental results obtained on two different datasets show the effectiveness of the proposed technique by comparing it with the standard SVR.

1.4 Clutter detection in radar sounder signals

The main novel contribution of the thesis related to the analysis of RS signals for clutter detection is summarized in the following.

A novel technique based on interferometry for clutter detection in two-channel RS data

As mentioned above the second part of the thesis is devoted to clutter detection in RS data. The capability of detect clutter is fundamental for the interpretation of sub-surface features, which are often of primary interest to geologists and planetary. Indeed, the clutter can prevent the capability of detect and analyze subsurface features in the radargram. The clutter is usually identified either by manual or automatic techniques that require ancillary information about the topography of the surface, or by using multi-channel RS systems with arrays of antennas. However, topographic information is not always available and multi-channel systems are generally too massive and costly to mount

on satellites for planetary exploration. In this contribution to the thesis, we propose a novel approach to clutter discrimination that is independent from ancillary information and limits the hardware complexity of the RS system. This approach uses a two channel RS and exploits cross-channel interferometric phase differences to discriminate the clutter. Our approach includes three main steps: i) manual feature extraction and theoretical phase difference estimation; ii) RS interferogram formation; iii) comparison of theoretical and real phase difference distributions. The proposed contribution was validated on RS data acquired in Greenland and provides a proof of concept for surface clutter discrimination using RS data. This contribution also aims to demonstrate that a clutter detection method based on interferometry can be implemented on two-channel RS data acquired from satellite platform. To this goal we assess the capability of the proposed method to discriminate clutter features from subsurface nadir returns using simulated data. Despite the analysis presented in this contribution can be generalized to a large number of radar sounders, we compare: i) an airborne RS having the same system design of MARFA with ii) a planetary RS operating on the Jupiter Icy Moons. Limiting the hardware complexity and thus the consumption of the most relevant resources for satellites sensors (e.g., mass, power, data rate, data volume), the proposed technique aims to provide a proof of concept for a planetary RS. This technique is a candidate to be considered for the implementation of the REASON RS for Europa mission (Moussessian et al. [2015]).

1.5 Structure of the Thesis

This document is divided into six main chapters. In Chapter 1, the introduction to this thesis dissertation has been provided. Chapter 2 illustrates the fundamentals and the background notions needed for the understanding of the thesis. The remaining four chapters are the novel contribution of the thesis. The first part, is devoted to the developed novel methods for the estimation of biophysical parameters. Chapter 3 presents a novel hybrid method for the correction of the theoretical model inversion in bio/geophysical parameter estimation. Chapter 4 illustrates a novel semisupervised method for the estimation of parameters in the framework of SVR. Chapter 5 describes the proposed interferometric technique for clutter detection using a two-channel radar sounder data. In each chapter an introduction to the specific topic is provided. Finally, in Chapter 6 the conclusions of the thesis are drawn along with proposal for future research and developments.

Chapter 2

Fundamentals and background

In this Chapter we review the fundamentals of microwave remote sensing instruments and we analyze the state of the art of regression methods for parameter estimation and clutter detection. First, we illustrate the basic principles of microwave systems relevant to this thesis. Afterwards, we focus on the estimation methods for parameter estimation. Finally, we discuss the state of the art regarding methods from clutter detection in radar sounder data.

2.1 Introduction to microwave remote sensing systems

This Section reviews the basic concept of the microwave sensor systems exploited in this thesis. First, the monostatic radar system and the general radar equation are defined. Then, the fundamentals about SAR and SAR interferometry are presented. Finally, we provide a description about RS systems.

2.1.1 Radar

Radar detection and ranging (Radar) refers to a technique as well as an instrument. The radar instrument emits electromagnetic pulses in the radio and microwave regime and detects the reflections of these pulses from objects in its line of sight. The radar uses the two-way travel time of the pulse to determine the range to the detected object and its backscatter intensity to infer physical quantities or the roughness of the surface. The most important design parameters of a radar systems are the central frequency f_c and the bandwidth B_w . f_c is typically in the range between few MHz and tens of GHz depending on the type of radar and the application. A monostatic radar receives the signal backscattered from a target using the same antenna used for pulse transmission (see Figure 2.1). This received power P_r can be computed with the radar equation that

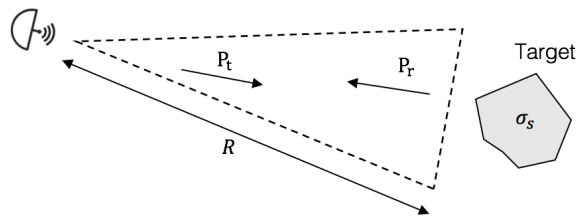


Figure 2.1: Monostatic radar configuration.

expresses the received power as a function of the transmitted power P_t , the wavelength λ which is associated to f_c , the gain of the antenna η_{ant} , the distance in range between the radar and the target R , the medium losses η_{loss} and the radar cross section RCS (Ulaby and Long [2014]). The radar equation can be written as:

$$P_r = \frac{P_t \lambda^2 \eta_{ant}^2}{(4\pi)^3 R^4 \eta_{loss}^2} RCS \quad (2.1)$$

This equation changes depending on the application and type of radar. The ratio between the received power P_r and the noise in which the radar is working is defined as SNR and represents an important factor for the assessment of the radar measurement. Signal-to-noise ratio (SNR) is calculated as follows:

$$SNR = \frac{P_r t_p}{k_B T_N} \quad (2.2)$$

where k_B is the Boltzmann constant, t_p is the pulse duration and T_N is the noise temperature that expresses the noise power introduced by the electronic components of the radar. In the radar equation the parameter that expresses the interaction between the wavelength and the target is the RCS. This parameter depends on the surface characteristics in terms of roughness or dielectric permittivity of the target. The backscattered signal received by the radar is composed of a mix of specular and diffuse components, depending on the roughness of the surface. Three different situations are envisaged: i) the wave is entirely reflected in the specular direction because the surface is smooth (i.e., coherent scattering). ii) the backscattering is mostly composed of a specular component but there is a small diffusive component because the surface is slightly rough and iii) the waves are scattered diffusely in all directions due to a very rough surface (i.e., incoherent scattering). The coherency or incoherency of the backscattered wave is driven by the relation between the radar wavelength and the roughness of the surface and it is critical in radar signal analysis.

This overview on radar systems is general to all types of radar systems. In the following we will focus on SAR systems, since they are exploited for biophysical parameter

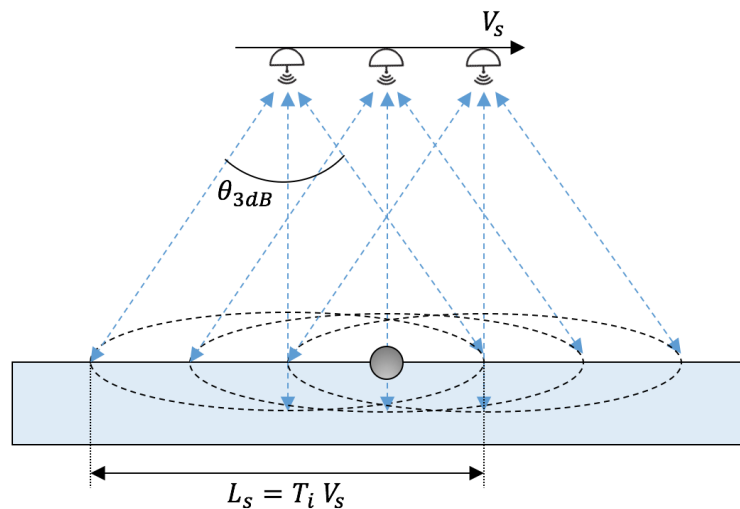


Figure 2.2: Principle of SAR.

estimation, which is one of the subject of the thesis. In this context, we provide also background information on radar interferometry. This concept has been extended to RS system for solving the problem of clutter detection on icy terrains. Finally, RS systems will be presented.

2.1.2 Synthetic aperture radar

Synthetic aperture processing or Doppler filtering is a well-known processing technique possible when there is relative motion between the radar and the target. The resolution of the system can be improved in the motion direction exploiting the analysis of the phase history of the target. Figure 2.2 illustrates the considered geometry in the case in which the radar is flying on the platform and the target is stationary. As the radar is moving along its path, an ideal point target on the ground is illuminated by the radar in a time interval t_i called integration time. The space covered by the radar during t_i corresponds to:

$$T_i = \frac{\theta_{3dB} R}{V_s} \quad (2.3)$$

$$L_s = V_s T_i \quad (2.4)$$

where L_s represents the length of the synthetic aperture, θ_{3dB} is the 3 dB aperture of the radar antenna in the along-track direction, and V_s is the velocity of the radar. During the integration time, the target response shows different Doppler shifts due to the relative motion of the spacecraft with respect to the target. Therefore, although different targets are present in the same antenna footprint, their returns have different Doppler shifts.

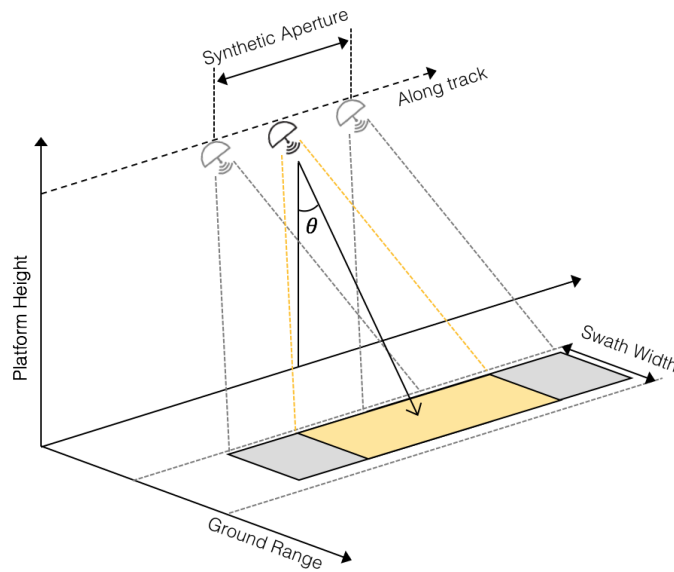


Figure 2.3: SAR acquisition geometry.

Coherent radars thus measure and record the phase history of the received signals. The phase information is then exploited to resolve the ground targets in the Doppler domain analyzing the phases of a series of consecutive echoes using a focusing algorithm. Thanks to this technique an antenna length longer than the physical one is synthesized.

SAR image formation

When the SAR processing is exploited on radars having side-Looking geometry it is capable to acquire images of the observed target. The main central frequencies used in SAR for imaging are L-band, C-band and X-band. These frequency bands allow the microwave signal to acquire SAR images independently from daylight, cloud coverage and weather conditions (Moreira et al. [2013]). SARs represent an evolution of the so-called real aperture radars (RAR). The main difference between RARs and SARs is the much finer along-track resolution of SARs, which is obtained by exploiting synthetic aperture processing. Airborne and satellite SAR systems illuminate the scene using the geometry shown in Figure 2.3. The antenna of the radar system is mounted on a flying platform moving in the along-track direction. The slant range is the direction perpendicular to the radar's flight path. The angle between nadir and the radar beam direction is called incidence angle. The swath width gives the ground-range extent of the radar scene, while its length depends on how long the radar is turned on. The transmitted signal commonly used is the so-called chirp, which is a frequency modulated pulsed waveform. The amplitude of the transmitted waveform is constant, while the instantaneous frequency

varies linearly over time t_p according to $f_i = k_r t$, where k_r is the chirp rate. The received echo signal data form a two-dimensional data matrix of complex samples, where each complex sample is given by its real and imaginary part, thus representing an amplitude and phase value. To obtain useful information on the scene the received signal should be processed according the following steps: i) a range compression of the chirp signal to a short pulse, exploiting a multiplication in the frequency domain (which aims to reduce the computational complexity) of each range line by the complex conjugate of the spectrum of the transmitted chirp; and ii) an azimuth compression that convolves the signal with its reference function, which is the complex conjugate of the response expected from a point target on the ground.

Geometric resolution

After the range compression, the slant-range resolution δ_{slt} and the ground-range resolution δ_{grd} resolution can be computed as:

$$\delta_{slt} = \frac{v_{light}}{2B_w} \quad (2.5)$$

$$\delta_{grd} = \frac{v_{light}}{2B_w \sin \theta} \quad (2.6)$$

where θ is the angle between nadir and the radar beam direction (i.e., incidence angle). As mentioned above, SAR systems can improve the azimuth resolution by processing the phase information of the complex signals. Under the assumption of fully focused processing, the along-track resolution can be described form the following equation:

$$\delta_{alt} = \frac{h_0 \lambda}{2L_s \cos \theta} \quad (2.7)$$

where the synthetic aperture corresponds to a synthetic antenna length L_s , which is the distance traveled by the sensor while illuminating a target with its beam, is equal to:

$$L_s = \frac{h_0 \lambda}{L_a \cos \theta} \quad (2.8)$$

From (2.7), one can observe that the azimuth resolution of a SAR sensor is theoretically only dependent on the length of the actual antenna, but not on the distance between sensor and target.

2.1.3 Radar interferometry

As mentioned before, the signal transmitted by radars is a waveform characterized by amplitude and phase. Its scattering changes both the amplitude and the phase of the signals. The use of the phase information leads to interferometry (Rosen et al. [2000]).

The idea behind interferometry is to combine together two SAR images, acquired either two different antennas or using repeated acquisitions. The use of interferometry enables the observation of relative distances as a fraction of the radar wavelength, and the distance in the sensor locations (i.e., Baseline) allows to measure angular differences. Phase measurements in interferometric SAR can be made at degree level accuracy, and with typical radar wavelengths in 3-80 cm range this corresponds to relative range measurements having millimeter accuracy. The SAR interferometry has been introduced in several application fields, such as: i) mapping and cartography (e.g., to produce digital elevation models (DEMs) from airborne platform (Madsen et al. [1993]) and land elevation maps at global scale using spaceborne interferometry (SRTM) (Rodriguez et al. [2005])), hazard alert and change measurements (e.g., displacement maps (Massonnet et al. [1994])), Earth science and climate change research (e.g., Glacier and ice sheet dynamics (Hoen and Zebker [2000])), geodesy (Eineder et al. [2011]) and atmospheric measurements (Ferretti et al. [2000])).

Different types of interferometric radar can be classified according to the geometric configuration of the baseline vector. On one side, in cross-track interferometric radars the two antennas are separated in the cross-track direction. This technique can be implemented: i) having a single antenna that performs repeat pass on the same area (see Figure 2.4a) or ii) with two antennas that fly on the same platform (see Figure 2.4c). Cross-track interferometric radars are usually exploited to compute topography and terrain deformations. On the other side, along-track interferometric radars have two antennas separated in the along-track direction and operate single pass (see Figure 2.4b) to exploit radial velocity measurements. In a radar image each pixel is a complex phaser representation of the coherent backscatter from the resolution element on the ground and the propagation phase delay. The signal received from each antenna is defined as in the follows:

$$s_1 = A_b e^{-j\phi_b} e^{-j\frac{4\pi}{\lambda}\rho_1} \quad (2.9)$$

$$s_2 = A_b e^{-j\phi_b} e^{-j\frac{4\pi}{\lambda}\rho_2} \quad (2.10)$$

where ρ_1 and ρ_2 , which are the path lengths of the received signals for left and right channel respectively, is the only parameter changing in the equations. This implies that, pixels in two radar images observed from two different antennas have nearly the same complex phaser representation of the coherent backscattering from a resolution element on the ground, but they have different propagation phase delay. The two SAR images can be combined to obtain the interferometric image exploiting the following equation:

$$s_{int} = s_1 s_2^* \quad (2.11)$$

Moreover, combining the two SAR images, the phase difference (i.e., interferometric phase)

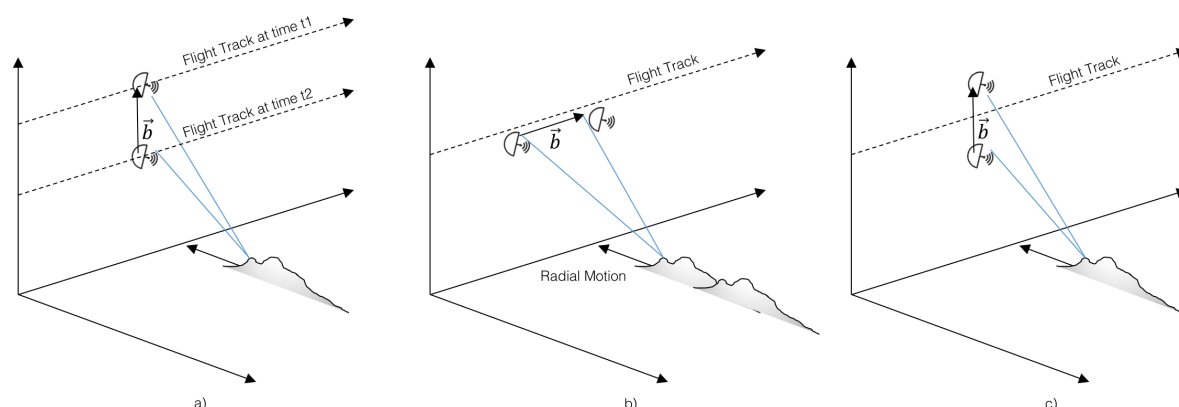


Figure 2.4: Interferometric radar configurations.

can be computed:

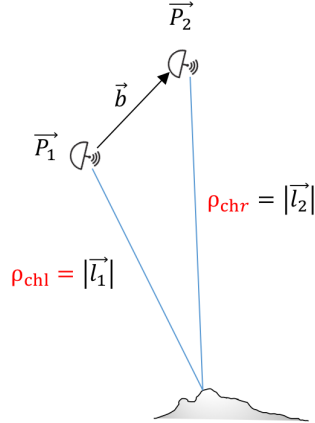
$$\Delta\phi = \frac{2\pi}{\lambda}(\rho_2 + \rho_1) - \frac{2\pi}{\lambda}(\rho_2 + \rho_1) = \frac{2\pi}{\lambda}(\rho_2 - \rho_1) \quad (2.12)$$

However, this is an approximation of the interferometric phase. The exact formula and its demonstration is given in Figure 2.5.

In the following, we focus our discussion on single-pass interferometry, which is characterized by one antenna that transmits and receives the echoes. The main advantage of single-pass interferometric systems is the reduction of temporal decorrelation between the two SAR images. This is fundamental to exploit topographic mapping applications. However, the baseline is limited by the platform dimension and constraints (e.g., mount points definition and aerodynamic considerations for airborne) and requires high accuracy in its length definition. The interferometric decorrelation is a critical factor that affects height and displacement accuracy (Zebker and Villasenor [1992]). The correlation γ is a complex measure that describes the similarity of the signals recorded at the two antennas, as given in the following equation:

$$\gamma = \frac{\langle s_1 s_2^* \rangle}{\sqrt{\langle s_1 s_1^* \rangle \langle s_2 s_2^* \rangle}} \quad (2.13)$$

Concerning a single-pass interferometric radar, the main cause of decorrelation is the thermal noise that affect the signal SNR, which is called SNR correlation γ_{SNR} and it can be measured on each pixel of the interferometric image. However, also the baseline decorrelation γ_B mainly caused by the change angular separation between the two antennas and the volumetric decorrelation γ_v can affect the interferometric system (i.e., $\gamma = \gamma_{SNR} + \gamma_B + \gamma_v$). The effect caused by the γ is called phase noise σ_θ and is computed



$$\Delta_{\phi t} = \frac{2\pi}{\lambda} (\rho_{chr} - \rho_{chl}) = \frac{2\pi}{\lambda} (|\vec{l}_1| - |\vec{l}_2|)$$

$$= \frac{4\pi}{\lambda} \left(\langle \vec{l}_2, \vec{l}_2 \rangle^{\frac{1}{2}} - \rho_{chl} \right)$$

$$= \frac{2\pi}{\lambda} \left(\langle \vec{l}_1 - \vec{b}, \vec{l}_1 - \vec{b} \rangle^{\frac{1}{2}} - \rho_{chl} \right)$$

$$= \frac{2\pi}{\lambda} \left((\rho_{chl}^2 - 2\langle \vec{l}_1, \vec{b} \rangle + b^2)^{\frac{1}{2}} - \rho_{chl} \right)$$

$$= \frac{2\pi}{\lambda} \rho_{chl} \left(\left(1 - \frac{2\langle \vec{l}_1, \vec{b} \rangle}{\rho_{chl}} + \left(\frac{b}{\rho_{chl}} \right)^2 \right)^{\frac{1}{2}} - 1 \right)$$

$$\vec{b} = \vec{P}_2 - \vec{P}_1 = \vec{l}_1 - \vec{l}_2$$

$$\vec{l}_2 = \vec{l}_1 - \vec{b}$$

$$b = |\vec{b}| = \langle \vec{b}, \vec{b} \rangle^{\frac{1}{2}}, \hat{l}_1 = \frac{\vec{l}_1}{|\vec{l}_1|} = \frac{\vec{l}_1}{\rho_{chl}}$$

Figure 2.5: Exact formula of the interferometric phase in the cross-track single-pass configuration.

as follows:

$$\sigma_{\phi} = \frac{1}{\sqrt{2N}} \frac{\sqrt{1 - \gamma^2}}{\gamma} \quad (2.14)$$

where N is the number of samples averaged by multilooking the SAR images.

This basics on SAR interferometry are relevant to the clutter detection method introduced in this thesis for the analysis of RS data.

2.1.4 Radar sounder

The RS is a nadir-looking radar working at low frequency (usually between few MHz and few hundreds of MHz). It is a particular type of SAR system that has the capability to investigate the structural and dielectric characteristic of a subsurface. RS systems are particularly suitable to sound deep ice, where the signal can deeply penetrate. The backscattered power can be received until few kilometers due to the dielectric transparency of this material (Bogorodsky et al. [1984]). In the following the main characteristics of RSs and the radargram formation are introduced.

Let us consider a RS instrument is mounted on a platform flying at an altitude h_0 from the ground. The platform altitude can vary between several hundreds of meters and several hundreds of kilometers, depending on the type of platform (i.e., airplane or satellite). The instruments emits a pulse having a duration t_p every pulse repetition

interval (PRI). The pulse travels from the sensor to the surface, where a part of the signal is reflected or scattered. A portion of it penetrates the subsurface and is backscattered by the subsurface targets and turns back through the ice and air to the sensor. In this scenario the targets are thermal or mechanical discontinuities that implies a dielectric contrast. While sending pulses the platform is moving in the so called along-track direction and the resulting data is called radargram. A radargram is a 2D image that represents the recorded echo power for a given range position in dept as a function of time (or distance), and as a function of the instrument position in the along-track direction. Figure 2.6 shows the acquisition geometry of a RS mounted on a airborne platform. As mentioned above, the radargram can be represented in both the time or the distance (i.e., depth) domain. This conversion depends on the dielectric properties of the medium in which the radar signal propagates. In particular, the velocity of propagation can be defined as:

$$v = \frac{1}{\sqrt{\mu_0 \mu_r \epsilon_0 \epsilon_r}} = \frac{v_{light}}{\sqrt{\epsilon_r}} \quad (2.15)$$

where μ_0 is the magnetic permeability of vacuum, μ_r is the material relative permeability, ϵ_0 is the dielectric permittivity of vacuum, ϵ_r is the dielectric permittivity of the material. From the formula, the propagation velocity of the waves through a real material is always smaller than the speed of the wave propagating in vacuum (i.e., light speed v_{light}). Considering the two-way path of the wave in the air and into the material, the depth is computed according to the following equation:

$$z = \frac{vt}{2} \quad (2.16)$$

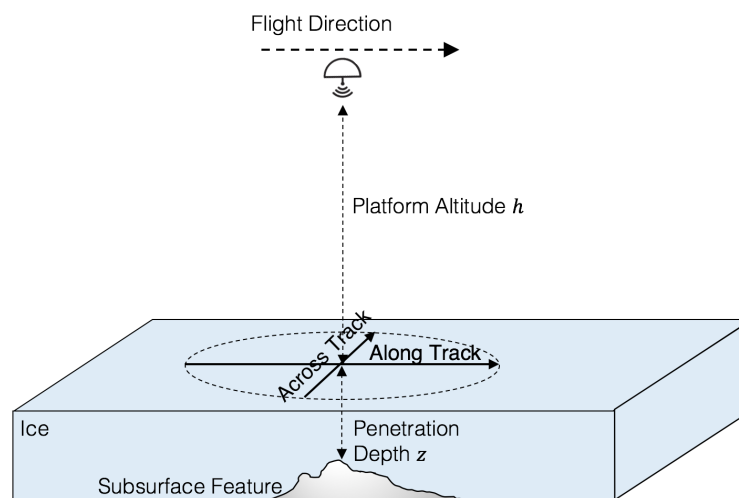


Figure 2.6: Nadir acquisition geometry of radar sounder systems.

where v , and consequently ϵ_r is the only parameter that relates the propagation depth z with the propagation time t . ϵ_r strongly affects also the penetration capability of the RS pulse. Indeed, it affects the main factors describing the signal propagation in the medium, which are: reflection, transmission and attenuation. The reflection and transmission effects can be described by the Snell's law under the assumption of flat surface. The reflection coefficient (i.e., scattering coefficient) between two interfaces (e.g., air and ice) is defined as follows:

$$\sigma_s = \left| \frac{\sqrt{\epsilon_{r,p}} - \sqrt{\epsilon_{r,p+1}}}{\sqrt{\epsilon_{r,p}} + \sqrt{\epsilon_{r,p+1}}} \right|^2 \quad (2.17)$$

where $\epsilon_{r,p}$ and $\epsilon_{r,p+1}$ are the relative dielectric constants of layer p and $p+1$, respectively. Assuming that no absorption loss is generated by the interface, the fraction of energy that is transmitted through the interface is given by the transmission coefficient $\tau_{p,p+1}$, as follows:

$$\tau_{p,p+1} = 1 - \sigma_s \quad (2.18)$$

An important factor that affects the signal propagation is the two-way power attenuation. The attenuation depends on the characteristics of the medium, in terms of dielectric properties (e.g., material, water content) and structure (e.g., porosity) (Fu and Daniels [2007]). It can be modeled considering an homogeneous layer according the following exponential equation:

$$\eta_{loss}^2 = e^{4\alpha_z z} \quad (2.19)$$

where α_z is the attenuation factor and z is the penetration depth in the subsurface. The attenuation factor α_z is given by:

$$\alpha_z = \omega \sqrt{\frac{\mu \epsilon'}{2}} \sqrt{\sqrt{1 + \left(\frac{\epsilon''}{\epsilon'}\right)^2} - 1} \quad (2.20)$$

where $\omega = 2\pi f$, μ is the magnetic permeability of the medium, and

$$\epsilon = \epsilon_0 \epsilon_r = \epsilon' - i\epsilon'' \quad (2.21)$$

where the term ϵ represents the dielectric permittivity of the material, which is given by the product of the vacuum permittivity ϵ_0 and the material relative permittivity ϵ_r . Considering this formulation, the attenuation is frequency dependent, it is higher at higher frequencies. This implies that at higher central frequency the penetration depth decreases.

Geometrical resolution Using a chirp signal, after the range compression, the vertical resolution of a RS mainly depends on the bandwidth Bw and is equal to:

$$\delta_z = \frac{v_{light}}{2Bw\sqrt{\epsilon_r}} \quad (2.22)$$

Thus, the effective resolution in the subsurface depends on the material in which the wave is traveling. Usually, weighting is applied during the range compression in order to reduce the sidelobes due to the signal processing. As a result, the effective range resolution worsens by a factor that depends on the applied weighting function (e.g., Hanning, Hamming). It is important to note that the bandwidth of the signal is a key factor also for the gain of the system. Indeed, radar systems using chirp signals can achieve a gain equal to the range compression factor, given by:

$$\eta_z = t_p B_w \quad (2.23)$$

In RS systems the along-track resolution after focused processing can be computed as

$$\delta_{alt} = \frac{V_s}{B_D} \quad (2.24)$$

where B_D is the Doppler bandwidth that is:

$$B_D = \frac{2V_s^2}{h_0\lambda} T_i \quad (2.25)$$

Equation (2.8) indicates the maximum ideal integration time. In real scenarios as for spaceborne RSs, it is commonly assumed that the coherent scattering from the ground is limited by the first Fresnel zone. The diameter of the Fresnel zone D_F is given by:

$$D_F = \sqrt{2\lambda h_0} \quad (2.26)$$

According to the Fresnel zone, the effective integration time becomes:

$$t_{i,eff} = \frac{D_F}{V_s} \quad (2.27)$$

Thus the along-track resolution calculated using the effective integration time is lower than the maximum value considering the ideal case. The number of echoes that should be processed to obtain the synthetic aperture depends on the effective integration time and the pulse repetition frequency (PRF). Such echoes are integrated to focus one resolution cell. As a consequence, the SNR of the focused signal increases by a factor that depends on the number of echoes (i.e., along-track compression factor η_a).

2.2 Overview on approaches for biophysical parameters estimation

The aim of this Section is to provide an overview on approaches for parameters estimation. First, we briefly introduce the problem of parameter estimation using remote sensing data. Afterwards, we describe the derivation of empirical data-driven relationships using empirical methods. Finally, we focus on approaches based on the inversion of physical models.

2.2.1 Biophysical parameters estimation problem

The accurate estimation of biophysical parameters from remote sensing images is a challenging task. From an analytical view point an estimation problem can be solved by designing the best mapping function possible that couples the parameter to be estimated (i.e., target) with a set of feature extracted from the remote sensing data. The mapping function must properly model the problem guarantying good generalization capability to minimize the overall estimation error. The main approaches presented in the remote sensing literature to solve parameter estimation problems are: i) derivation of empirical data-driven relationships using empirical methods and ii) inversion of physical based theoretical models. In the following sections these approaches are presented.

2.2.2 Derivation of empirical data-driven relationships using empirical methods

Empirical methods relies on the availability of a set of reference samples, i.e., couples of ground truth measurements of the desired target variable associated with the corresponding measurements of the remote sensor. These samples are exploited for deriving an empirical mapping. Simple function, such as linear, polynomial and logarithmic, can be used to describe this relationship among a set of variables and the corresponding measurement of the parameters of interest. However, they usually have limited capability to solve complex estimation problems. Due to the rising interest in the remote sensing community in this class of problems, more advanced empirical methods, relying on machine learning techniques, have been introduced in the field of biophysical variable estimation. In this context, supervised regression methods are an effective way to deal with the estimation of parameters. This advanced learning strategies can: i) approximate complex non-linear mapping problems exploiting the information contained in reference sample sets, ii) be independent from the distribution of the data and iii) integrate data coming from different sources. We recall the following theoretical frameworks: neural networks (Egmont-Petersen et al. [2002], Kwok and Yeung [1997]), Gaussian Process regression (Rasmussen [2006]) and Support Vector Regression (SVR) (Drucker et al. [1997], Smola and Schölkopf [2004]). Among them, due to their good intrinsic generalization ability and the robustness to noise, SVR methods demonstrated the capability to achieved promising results (Demir and Bruzzone [2014]). Bruzzone and Melgani [2005] introduced Support Vector Machines (SVM) in the remote sensing field to find a robust solution to regression problems. An improved approach was presented in Camps-Valls et al. [2006], where an epsiv-Huber cost function is proposed in the framework of SVR to improve the estimation results when few in situ measurements are available. A related work was proposed in

Durba et al. [2007], focused on the retrieval of leaf area index from multi-angle imaging spectroradiometer data. SVR strategies requires the user to set internal parameters. For this reason methods for the automatic selection of this parameters are presented in Moser and Serpico [2009] and Pasolli et al. [2012]. These works exploit a functional minimization correlated to regression errors and a multi-objective function based on a set of metrics (i.e., mean squared error and determination coefficient), respectively. More recently a multi-output SVR approach for the simultaneous estimation of different biophysical parameters has been presented (Tuia et al. [2011]). The use of SVR demonstrated its robustness when exploited to address the estimation of biophysical parameters from microwave sensors and in particular exploiting SAR images. As an example in Pasolli et al. [2011] and Pasolli et al. [2015], the authors present regression methods based on SVR for retrieval of the SMC from Radarsat-2 data in challenging scenarios as mountains areas are for SAR data. It worth noting that, when supervised regression approaches are exploited, the capability to obtain a reliable regression function strictly depends on the quality and quantity of the reference samples available to the learning phase of the algorithm. Small or biased training sample sets can prevent the capability of an empirical supervised estimation technique to model an accurate mapping function.

To solve this issue, semisupervised learning techniques are presented in the literature (Zhu [2010]). This approaches represent a solution to deal with low quality or small quantity training samples. Only few related works on SSL are presented in the remote sensing literature, while SSL is a deeply studied in the pattern recognition and machine learning communities. As an example, in Bazi et al. [2012] a semisupervised regression for biophysical parameter estimation using multi-objective optimization in the field of Gaussian processes is proposed. This work aim to iteratively involve high confidence unlabeled samples in the learning phase of the regression using an iterative algorithm. At each iteration, the labels of unlabeled samples are estimated and, among them, the unlabeled samples that can improve the regression performance are selected and added to the training set. Another SSL work is presented in Camps-Valls et al. [2009]. In this work the authors presents two semi-supervised methods for the estimation of biophysical parameters in the context of SVR. These methods rely on building a graph or hypergraph Laplacian using the available labeled and unlabeled samples that are used to deform the training kernel matrix.

2.2.3 Inversion of physical based theoretical models

The second class of approaches to address parameter estimation problems is based on the inversion of physical models. These approaches exploit electromagnetic models to define the mapping function. In this way, the function represents the interaction between

the electromagnetic radiation and the target variables. In the direct operational way, they simulate the response of a target object as function of: i) the target characteristics (i.e., structural, chemical and biophysical variables); and ii) the signal characteristics (i.e., wavelength, incidence angle, etc.). In the inverse operational way, the simulations can be used to represent the mapping between the measurements at the remote sensor and the target variable of interest. Due to the solid physical foundation and the wide range of applicability (in terms of both target properties and system characteristics), electromagnetic models can operate in more general scenarios also in cases where reference measures are not available. Thus, theoretical electromagnetic models with different levels of complexity and generality have been widely studied in the literature. When dealing with microwave emission and scattering, the most widely used model is the Integral Equation Model (IEM) that can estimate backscattering coefficients in the context of different configuration of active sensors, e.g., incidence angle, polarization and central frequency (Fung et al. [1992]). Furthermore, the IEM characterizes surface soil moisture in terms of dielectric constant and surface roughness, and thus it is appropriate for soil moisture estimation problems from radar measurements. The electromagnetic models may have high complexity and dependence on a huge number of input parameters. This issue makes the inversion process often analytically complex. To overcome this problem, several different inversion strategies have been presented in the literature based on: i) iterative search algorithms, such as the Nelder-Mead and the Netwon-Rapson methods (Meroni et al. [2004] and Paloscia et al. [2008]), which iteratively use different model parameter configurations to minimize a dissimilarity measure between simulated and measured electromagnetic response of a target object, ii) look-up table matching (Darvishzadeh et al. [2008]), which searches among a set of pre-computed simulated spectra the most similar to the remote measurement and iii) regression methods (Song et al. [2009] and Duveiller et al. [2011]), which exploit a set of simulated samples (i.e., target biophysical variables and simulated electromagnetic responses).

2.3 Overview on clutter in radar sounder data

The aim of this Section is to analyze the clutter problem associated to the acquisition of RS data. Moreover, we provide an overview of the state of the art regarding the clutter suppression techniques which use multi-channel RS systems. Finally, we present techniques for clutter detection that exploits information on the surface topography.

2.3.1 Clutter problem definition

Clutter is a well known problem that affects the acquisition of RS data both from airborne and satellite platforms. Among the other sources of noise (e.g., thermal noise, speckle and sidelobes) the clutter is the most difficult to be avoided. Figure 2.7 shows the geometry of this phenomenon. The beam pulse of the RS antenna usually defines a large footprint on the ground. If the ground swath contains off-nadir topographic elements, detectable at the RS resolution, the receiver records their backscattered signals. Appearing in the radargram under the level of the surface, these off-nadir contributions can be mistaken with subsurface returns. The presence of strong clutter may hamper the correct interpretation of the radargrams. The strength of clutter depends on the size of the detected off-nadir target and on its slope and roughness. Moreover, the power of the received clutter return is frequency dependent. In particular, it can monotonically increase with the frequency. This is one of the driving factor in the design of the current airborne sensors that usually have central frequency in the range between 60 and 150 MHz (Peters et al. [2007], Jezek et al. [2011b]). Since at higher frequencies the ice surface appears very rough compared to the wavelength, the use of RS working at higher central frequencies (e.g., 1 GHz) can lead to unwanted clutter returns of considerable strength (Bekaert et al. [2014]). According to the RS acquisition geometry reported in Figure 2.6, clutter returns can be received from both along-track and across-track directions. Doppler filtering and SAR processing (i.e., azimuth focusing) are successfully implemented to deal with along-track clutter. Azimuth focusing is able to reduce the area on the RS ground swath where the clutter is coming from. Figure 2.8 shows the reduction of the surface region that contributes to surface clutter in the along-track direction after the azimuth focusing. This reduces the clutter

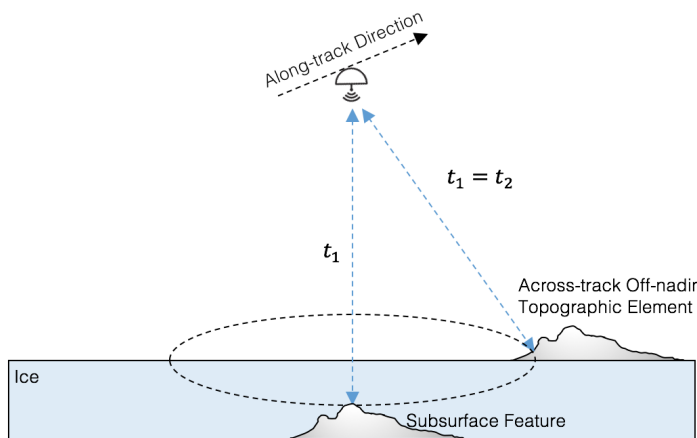


Figure 2.7: Geometry of the clutter problem.

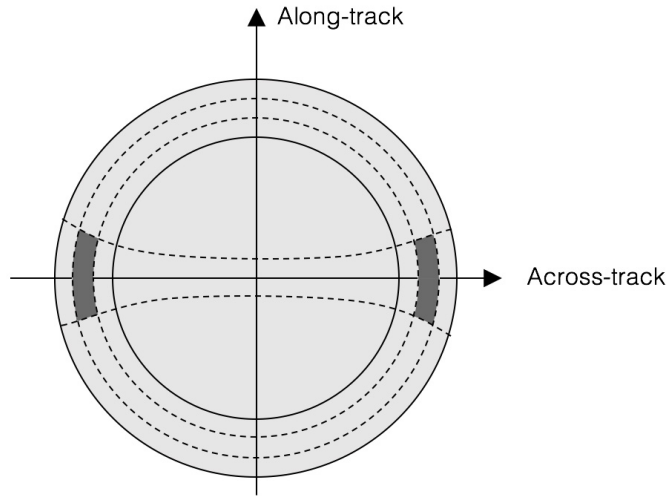


Figure 2.8: Surface clutter contribution before SAR processing (light gray locus), and after focusing in azimuth direction by Doppler processing and in range by range compression (dark gray cells). This example assumes flat topography in the absence of airplane roll.

problem to a one dimension problem. Thus, we mainly focus our interest on across-track clutter contribution because: i) SAR processing is a baseline for most of the RS systems to avoid along-track clutter, and ii) no synthetic aperture processing can be performed in the across-track direction (the illuminated points have zero Doppler shift).

The clutter issue makes the analysis of radargrams a challenging tasks. In the literature, two different approaches are presented: i) techniques for clutter suppression and ii) techniques for clutter detection. In the following, suppression and detection techniques are presented.

2.3.2 Clutter suppression techniques

Clutter suppression techniques aim to eliminate or mitigate the surface clutter contribution before the radargram formation. The capability to suppress the clutter, while the radar is acquiring the data, strictly depends on the system design. Radar systems capable to suppress the across-track clutter exploit sophisticated antenna with very directive beam pattern to illuminate a small portion of the surface or to exploit multi-channel systems, that, combining returns coming from different direction of arrivals, can underweight the antenna pattern in the expected clutter direction.

The P-Band multi-channels system presented in Bekaert et al. [2014], allows the development of both the Null Steering (NS) and the Minimum Variance Distortion-less

Response (MVDR) techniques. NS provides clutter reduction at the risk of a degraded SNR, while MVDR, which represents a combination of beam steering (i.e., assuming that the clutter angles are known the antenna pattern is steered to the angle of interest) and NS, is capable to jointly minimize the power of clutter and the white noise, however it depends on prior signal-to-clutter information. Wu et al. [2011] show that a tomographic method can reject the surface clutter as well as the beam steering does. Moreover, tomographic ice sounding provides good estimations of both the ice thickness and the bottom backscattering intensity. In Nielsen and Dall [2015], a multi-channel RS is used to analyze the direction of arrival (DOA) of ice sheet data collected on Antarctica. DOA estimation has been applied to data acquired with the four-channel POLARIS to improve the performance of surface clutter suppression techniques. The analysis of the DOA has recently recognized as a very interesting solution to address also ice thickness retrieval, icebed roughness and bed slope (Nielsen et al. [2017]). Jezek et al. [2011a] exploit a complex RS system composed of a two-way high-power transmitter and eight independent receiver channels to study the influence of surface clutter on the obscuration of the basal return at different elevations and frequencies. Moreover, using interferometric techniques the authors were able to estimate cross-track ice thicknesses and obtain strip maps of ice thickness along the flight track. Despite multi-channel RS systems have been mainly implemented for airborne platform, an attempt to suppress was envisaged on MARSIS mounting a secondary monopole antenna oriented along the nadir axis (Picardi et al. [2003]). This monopole was devoted to independently acquire the off-nadir returns. Unfortunately, antenna calibration issues made this technical solution difficult to be applied. Also polarimetry has been presented as a possible solution in Raney [2008]. Polarimetry is based on the transmission of a circular polarized signal. Under the assumption of specular reflection, the sense of polarization can be distinguished. Nadir returns present inverted polarization, while clutter returns result in the same polarization transmitted, because of the double bounce effect. In principle, this method requires only two Rx channels (H and V) for the detection of the polarization change. The state of the art reported above demonstrate a recent growth of interest in the development of multi-channel RS systems to tackle the clutter problem.

2.3.3 Detection techniques

Clutter detection techniques mainly involve algorithms for the clutter surface simulation that use as input a digital elevation model (DEM) of the surface. Different simulation algorithms have been presented in the literature to develop the cross-track mapping of the radar echoes onto the surface. These algorithms allow one to manually or automatically compare a radargram with a surface clutter simulation and to map the location and shape

of the clutter echoes in the radargram. Clutter detection is addressed in the literature for both airborne and satellite RSs. As an example, in Holt et al. [2006] the authors present an echo discrimination technique using the DEM to predict the location of off-nadir echoes and to compute the signal-to-clutter ratio on single-pass airborne data acquired by airborne on the Dry Valleys of Antarctica. In Wu et al. [2015] a radar signal simulator based on Kirchhoff approximation of rough surface scattering on multi-layer stratified media to improve the discrimination of ice bed from surface clutter in crevassed regions of Antarctica and Greenland. Detection techniques using topographic information are widely used also for the analysis of planetary RS data. As an example, in Ferro et al. [2013], the authors propose a clutter detection method made up three steps: first, simulate the clutter using an incoherent radar simulator; then, automatically coregisters radargrams and simulations; finally, extract surface clutter returns from the coregistered radargrams. Other simulators are presented in the literature to predict clutter on both SHARAD and MARSIS data. In this context, fast simulators to optimize the computational complexity (Iorio et al. [2008]) and precise coherent simulation algorithms based on the facets method (Nouvel et al. [2004]), are presented. Such studies prove a recent growth in interest in the development of techniques for clutter detection for the analysis of icy terrains using RS data. However, in some cases clutter detection techniques based on radar signal simulation are not applicable, because DEMs having sufficient spatial resolution to properly represent features in the radargram, are not available. This issue leads to the development of novel techniques capable to detect the clutter independently from topographic information about the surface.

Chapter 3

A novel hybrid method for the correction of the theoretical model inversion in bio/geophysical parameter estimation

This Chapter provides the first contribution of the thesis in the context of the estimation of biophysical parameters. In particular, it presents a novel hybrid method that aims at integrating theoretical analytical models with empirical observations. Exploiting two consecutive steps, the proposed method models and corrects deviations from reference target values when theoretical electromagnetic models are used for the inversion process. In order to estimate the deviation, in the first step two different strategies are presented: the global deviation biased (GDB) strategy, based on the assumption that the deviation is globally independent and identically distributed; and the local deviation biased (LDB) strategy, based on the assumption that the deviation is locally independent and identically distributed. Once the deviation is defined, the second step corrects the estimates obtained with the inversion of the theoretical model. The proposed hybrid method has been applied in the context of soil moisture content (SMC) estimation from microwave remotely sensed data.

3.1 Introduction

Accurate estimation of biophysical parameters (e.g., water chlorophyll concentration, leaf area index, soil moisture, tree stem diameter and volume) from remotely sensed data is an important research field. The biophysical parameters estimation (retrieval) can be

achieved by two different approaches: 1) derivation of empirical data-driven relationships; and 2) inversion of physical based theoretical electromagnetic models.

The first approach relies on the availability of labeled samples (for which reference measures are given) being used to derive an empirical mapping between remotely sensed data and target biophysical value. The empirical mapping is achieved through regression techniques, which estimate a functional relationship between a set of variables and corresponding target values (i.e., the measurement of the parameters of interest). Different machine learning methods have been used for addressing regression problems, e.g., neural networks, Gaussian Process regression, nearest neighborhood regression, kernel ridge regression, and Support Vector Regression (SVR) (Camps-Valls et al. [2006], Tuia et al. [2011], Pasolli et al. [2011]). Among several methods, the SVR has become popular due to its i) good generalization capability; ii) ability to handle problems with a small ratio between number of training samples and the size of the feature space; and iii) relatively limited computational load in the training phase (Demir and Bruzzone [2014]). The amount and the quality of the training samples are crucial to obtain accurate estimations by the use of regression algorithms. An inadequate number of training samples may cause a reduction on the estimation performance in terms of accuracy and generalization ability. However, collecting labeled samples is expensive (since generating reference measures may require significant time). Moreover, empirical observations are typically site and sensor dependent, as the reference samples are collected under specific operational conditions. This issue limits the possibility to extend the use of reference samples to different areas and different remote sensing systems since the observations remain valid only under the conditions in which reference samples have been collected (Meroni et al. [2004]).

Unlike the approach based on empirical observations, the second approach relies on theoretical models that require the definition of the desired mapping function associated to an analytical electromagnetic model. Such models are defined based on the physics of the mechanisms that model interaction between the electromagnetic radiation and the target variables in the direct operational way. The theoretical models simulate the response of a target variable as function of: 1) the target characteristics (i.e., structural, chemical and biophysical variables); and ii) the radiation characteristics (i.e., wavelength, incidence/reflection angle, etc.). In the inverse operational way, the simulations can be used to represent the mapping between the measurements at the remote sensor and the variable of interest. Due to the solid physical foundation and the wide range of applicability (in terms of both target properties and system characteristics), electromagnetic models can operate in more general scenarios also in cases where reference measures are not available. Thus, theoretical electromagnetic models with different levels of complexity and generality have been widely studied in the literature (Fung [1994], Karam et al.

[1992]). When dealing with microwave emission and scattering, the most widely used model is the Integral Equation Model (IEM) that can estimate backscattering coefficients in the context of different configuration of active sensors, e.g., incidence angle, polarization and central frequency (Fung et al. [1992]). Furthermore, the IEM characterizes surface soil moisture in terms of dielectric constant and surface roughness, and thus appropriate for soil moisture estimation problems from radar measurements. The IEM has been further improved to consider multiple scattering effects in the Advanced IEM model (Fung [1994]) and often coupled with layered structured or three dimensional models to handle complex target such as vegetated areas (Karam et al. [1992]). The electromagnetic models may have high complexity and dependence on a huge number of input parameters. This issue makes the inversion process often analytically complex. The electromagnetic models may have high complexity and dependence on a huge number of input parameters. This issue makes the inversion process often analytically complex. Moreover, the theoretical models have two main limitations: 1) they adopt simplifying assumptions in the modeling of real phenomena, which may result in deviations from correct estimations; and 2) they neglect sensor noise, calibration errors and local propagations that may produce variability of the relations between the target variables and the features in different regions of the feature space. The first issue can be mitigated by increasing the complexity of the model, which results in reduction on the generalization ability, while the second one can be addressed by applying a statistical component to the estimation obtained by the theoretical model. However, often the results are not satisfactory. Thus, it is crucial to predict the possible deviations of the estimations when theoretical models are considered.

In this chapter, we address the above-mentioned limitations of theoretical models by proposing a novel hybrid approach that aims at integrating theoretical analytical models with empirical observations (which are associated to few field reference samples). This is achieved on the basis of a two steps technique. In the first step, a theoretical model is initially constructed to derive a mapping between the input variables and the target biophysical values. Then, field reference samples are taken into account for modeling in the feature space the deviations between the theoretical model predictions and the true target values. This information is then used to characterize the deviations for unlabeled samples. This is accomplished by two different strategies: i) the global deviation bias (which assumes a constant behavior of the deviation in the whole feature space for compensating bias between theoretical models and empirical observations); and ii) the local deviation bias (which takes into account the more realistic condition in which the deviation is variable within different portions of the input space). In the second step, the theoretical model estimations of unlabeled samples are corrected based on the deviations estimated in the first step. Accordingly, the proposed approach mitigates the limitations (such as

the simplifying assumptions) of theoretical models while keeping their generalization ability and robustness due to the rigorous theoretical foundation. It is worth noting that the proposed hybrid approach requires the availability of only few labeled samples that leads to limited operational cost. In our experiments the proposed approach is applied to the estimation of soil moisture from microwave remotely sensed data. However, it is general and can be used in any other application where i) the above-mentioned limitations of the theoretical models are critical; and ii) the limited availability of field reference samples represent a strong limitation to the use of empirical observations only.

The remaining part of this Chapter is organized as follows. Section 3.2 presents the proposed hybrid method. Section 3.3 describes the data used in the experiments. Section 3.4 illustrates the experimental results. Finally, Section 3.5 draws the conclusion of this work.

3.2 Proposed hybrid method

Let us consider a remote sensing data $\Upsilon = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$ made up of a very large number M of samples, where $x_i \in \mathfrak{R}$ is the i -th sample defined as $[x_i^1, x_i^2, \dots, x_i^d]$, $i = 1, \dots, M$. $x_i^l, l = 1, \dots, d$, is the l -th variable of the i -th sample in Υ and d is the total number of variables (i.e., features). If ground reference samples do not exist, a possible approach to estimate the target values of each sample in Υ is to use theoretical models. Let us assume that G is a theoretical forward model, which simulates the behavior of the remotely sensed signals as a function of the target value under certain assumptions and simplifications intrinsic in its analytical formulation. The estimation problem can be defined by a functional relationship between a set of variables and corresponding estimated target values as:

$$y_i^t = G^{-1}(\mathbf{x}_i) \quad (3.1)$$

where $G^{-1}(\cdot)$ represents the desired inverse mapping between the variable \mathbf{x}_i and the estimated target value y_i^t . Due to the simplifications and approximations intrinsic within the theoretical model formulation, the estimation is usually affected by a deviation $\delta(\mathbf{x}_i)$ from the target value y_i , i.e.,

$$y_i = G^{-1}(\mathbf{x}_i) + \delta(\mathbf{x}_i) = y_i^t + \delta(\mathbf{x}_i) \quad (3.2)$$

In other words the real target value of each sample can be modeled by two terms: i) the estimation derived from the inverse theoretical model function G^{-1} ; and ii) a deviation $\delta(\mathbf{x}_i)$. The term $\delta(\mathbf{x}_i)$ represents a correction to the error due to the simplified analytical formulation of the model and the intrinsic noise in the data. The characterization of the

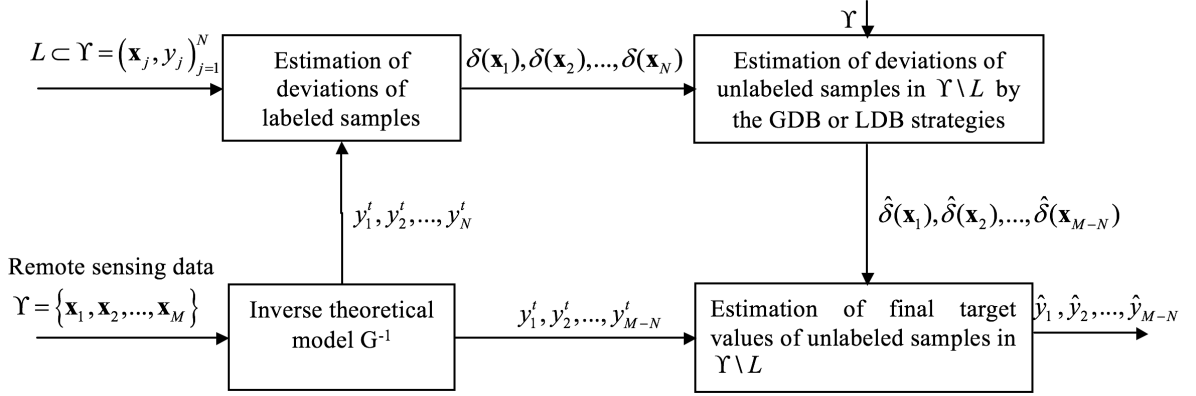


Figure 3.1: Block diagram of the proposed hybrid approach.

deviation of each unlabeled sample is crucial and depends on an availability of the prior information.

In this Chapter, we propose a hybrid approach that jointly uses the theoretical models and empirical observations (which can be related to a small number of ground reference samples) to estimate deviations associated to unlabeled samples and their final target values. Accordingly, we assume that a training set $L \subset \Upsilon$ made up of $N \ll M$ pairs $(\mathbf{x}_j, y_j)_{j=1}^N$ is initially available, where $\mathbf{x}_j \in \mathfrak{X}$ is the j -th training sample associated with the target value $y_j \in \mathfrak{Y}$. Note that the remaining samples in Υ (i.e., samples in Υ/L) are considered as unlabeled, i.e., the target measures of these samples are not available and should be estimated. The proposed hybrid approach consists of two consecutive steps: i) computing and modeling the deviations; and ii) estimating the final target values of unlabeled samples. Figure 3.1 shows the block scheme of the proposed hybrid approach. Each step of the approach is explained in detail in the following.

3.2.1 Computation and modeling of deviations

This first step is devoted to estimate the deviation of each unlabeled sample based on the available ground reference samples and their theoretical model estimates. To this end, initially the deviations of training samples are estimated as the differences between their correct target values (i.e., on the measurements of the parameters of interest) and the related theoretical model estimates, i.e.:

$$\delta(\mathbf{x}_j) = y_j - y_j^t, j = 1, 2, \dots, N \quad (3.3)$$

Then using the deviations of training samples, we aim at modeling the deviation values for the unlabeled samples for which the correct target values do not exist. Accordingly,

we propose two strategies that rely on two different assumptions associated to the distributions of deviations in the feature space: i) the global deviation bias (GDB), which assumes that the deviations are globally independent and identically distributed (i.i.d.); and ii) the local deviation bias (LDB), which assumes that the deviations are locally i.i.d. The detailed descriptions of the proposed strategies are given in the following.

Global deviation bias (GDB)

The GDB strategy assumes that the accuracy of a theoretical model is the same within all portions of the feature space, i.e., the distribution of the deviation δ between the output and the inverse model is globally i.i.d. On the basis of this assumption, the deviation associated with each unlabeled sample \mathbf{x}_i in $\Upsilon \setminus L$ is considered as uniform (i.e., the same deviation value for each unlabeled sample is considered) and characterized by the average deviation values of all training samples as follows:

$$\hat{\delta}(\mathbf{x}_i) = \frac{1}{|L|} \sum_{\forall \mathbf{x}_j \in L}^N \delta(\mathbf{x}_j), i = 1, 2, \dots, M - N \quad (3.4)$$

where $|L|$ is the total number of samples in L and $\hat{\delta}(\mathbf{x}_i)$ shows the estimated deviation for \mathbf{x}_i . The strategy is simple and has a low computational complexity. However, due to the simplifying assumptions intrinsic to theoretical models and the complexity of the inversion process (e.g., large number of variables and physical processes, and complex mathematical formulations), different accuracies of the theoretical model may be produced. Thus, assigning the same amount of deviation to all unlabeled samples may reduce the estimation accuracy of parameters. It is worth noting that using this strategy implicitly assumes that the distribution of the deviation δ between the output and the inverse model is globally independent and identically distributed.

Local deviation bias (LDB)

The LDB strategy, unlike the GDB strategy, assumes that the theoretical model can provide different accuracies in different portions of the feature space, i.e., the deviations are locally i.i.d. On the basis of this assumption, it aims at estimating a deviation for each unlabeled sample by assessing the deviations of training samples in the feature space that are located only within the same portion of the feature space with the unlabeled sample. This is achieved by adopting an estimation based on the k-nearest neighbor (k-nn) algorithm that is one of the simplest and most popular methods for the estimation of statistical variables. To this end, initially the distance between each unlabeled sample \mathbf{x}_i and the labeled sample $\mathbf{x}_j, j = 1, 2, \dots, N$ in the feature space is estimated. Then, the k-nearest labeled samples that have the smallest distance (i.e., the highest similarity) to

the unlabeled sample \mathbf{x}_i in the feature space are identified. Finally, the deviation value of each unlabeled sample \mathbf{x}_i in $\Upsilon \setminus L$ is defined as the average deviation values of the k -nearest labeled samples located in its neighborhood, i.e.,

$$\hat{\delta}(\mathbf{x}_i) = \frac{\sum_{\forall \mathbf{x}_j \in L} \delta(\mathbf{x}_j) W(\mathbf{x}_i, \mathbf{x}_j)}{\sum_{\forall \mathbf{x}_j \in L} W(\mathbf{x}_i, \mathbf{x}_j)}, i = 1, 2, \dots, M - N \quad (3.5)$$

where the function $W(\mathbf{x}_i, \mathbf{x}_j)$ outputs 0 or 1, depending on whether or not \mathbf{x}_j is among the k -nearest neighbors of the unlabeled sample \mathbf{x}_i , i.e., $W(\mathbf{x}_i, \mathbf{x}_j) = 1$ if \mathbf{x}_j is one of the k nearest neighbors of \mathbf{x}_i , and $W(\mathbf{x}_i, \mathbf{x}_j) = 0$ otherwise. Thus, $\sum_{\forall \mathbf{x}_j \in L} W(\mathbf{x}_i, \mathbf{x}_j) = k$.

An important issue in the LDB strategy is how to select an optimal value of the neighborhood parameter k to estimate reliable deviations. On the one hand, if the value of k is as high as the number N of the reference samples, the results of LDB strategy are the same of those of the GDB strategy. On the other hand, selecting a small k value may provide unreliable deviations affected by possible local noise or outliers. k is set according to a well-known rule of thumb (Mardia et al. [1982]), i.e., it is defined as the square root of the half of the total number of reference samples.

The LDB strategy avoids the unreliable estimations of deviations with the GDB strategy, by addressing a more realistic condition when compared to the GDB strategy. However, it requires higher number of reference samples than the GDB strategy to accurately model the deviations in different portions of the feature space. Thus, in the case of availability of a small number of reference samples, the simple GDB can be preferred.

3.2.2 Estimation of the final target values

After obtaining the deviation $\hat{\delta}(\mathbf{x}_i)$ for each unlabeled sample \mathbf{x}_i , the final target value of that sample is obtained by the sum of theoretical model estimate and its deviation estimated by using either the GDB or LDB strategy, as:

$$\hat{y}_i = y_i^t + \hat{\delta}(\mathbf{x}_i) \quad (3.6)$$

In other words, theoretical model estimates are corrected by the use of deviations. If the GDB is used, the same amount of correction is applied to the theoretical model estimates of all unlabeled samples. If the LDB is used, the deviations differ from each other, depending on the sample location in the feature space. It is important to note that the proposed hybrid method can be used for any biophysical parameter estimation problem. Varying the estimation problem requires to exploit the proper theoretical model (which should guarantee the considered i.i.d. assumptions), whereas it does not affect the formulation of proposed method. In addition, the proposed hybrid method is also independent from the adopted distance metric to identify the k -nn samples, and thus can be used with any of the metrics presented in the literature (Alpaydin [2009]).

3.3 Datasets description and experimental setup

In order to assess the effectiveness of the proposed approach, several experiments were carried out on three datasets, namely Scatt, SMEX and ActPass, in the context of the estimation of SMC from microwave remotely sensed data. The three datasets have been acquired from radar instruments at different central frequencies mostly on bare soil terrains. The considered datasets consist of backscattering radar coefficients together with field SMC surveys. The Scatt dataset is composed of 17 soil moisture reference samples acquired in two bare fields. Remote measurements were conducted with a ground C-band scatterometer using HH and VV polarizations at 23° and 40° incidence angles. The terrain was characterized by an extremely variable roughness, also within the same field due to plowing practice, and a medium/dry soil moisture condition. The samples were acquired at different times, from 1998 to 2004 on an agricultural field near Matera, Italy (Mattia et al. [2003]).

The SMEX dataset consists of 35 soil moisture reference samples coupled with dual-polarized backscattering coefficients (HH and VV) acquired in L-Band at 40° incidence angle. This dataset was acquired in Iowa, USA, from June 24th to July 12th, 2002, during a soil moisture experiment Notarnicola et al. [2006]. The main aim of the campaign was the validation of the soil moisture algorithm for the Japanese radiometer called Advanced Microwave Scanning Radiometer-Earth Observing System. The soil samples were acquired over four soybean fields characterized by a low biomass value. Concerning to the remote sensing backscattering measurements, the dual polarimetric radar data were acquired by an airborne Synthetic Aperture Radar (AirSAR).

The ActPass dataset consists of 29 reference samples acquired by a truck-mounted full-polarimetric radiometer-scatterometer operating at different frequencies (between 2.5 and 11 GHz) and incidence angles (Wegmuller et al. [1994]). The use of the same antenna for both active and passive measurements ensured that the two instruments observed the target under almost identical spatial conditions. The ground truth measurements of the volumetric soil moisture were acquired on three bare and smooth fields.

To obtain homogeneity with the simulated data that are used in the experiments for the inversion of the theoretical backscattering model, in this work the soil moisture measurements are converted into dielectric constant values (real part) through a well-known relationship (Hallikainen et al. [1985]). Table 3.1 summarizes the main characteristics of the three datasets.

In our experiments, we used the Integral Equation Model (IEM) (Fung et al. [1992]) as a theoretical model to simulate the microwave backscattering. This model is suitable to properly represent the microwave response when there is no contribution from the

vegetation in the backscattered signal, which is the case of a bare soil surface, or when the emitted wavelength is long enough, such as the L-Band, to make the biomass contribution negligible. In the results, we analyze both cases, i.e., Scatt and ActPass datasets are acquired only on bare soil field, whereas the SMEX is acquired on soybean fields (which are usually characterized by low biomass values). The heterogeneity of the experimental setup aims to demonstrate the effectiveness of the proposed hybrid model in different conditions. Concerning the inversion of the IEM theoretical model, the inverse function has been computed using the SVR technique due to the complexity of the IEM model inversion process. It is worth noting that SVR is used in the experiments for the inversion of the theoretical models due to its effectiveness proven in the literature (Pasolli et al. [2011]). In details, IEM model generated 300 simulated reference samples in the forward operational way Pasolli et al. [2012]. Such reference samples were generated under the same conditions (in terms of signal frequency, polarization, acquisition angle and soil characteristics) of the available real data. In the experiments, simulated backscattering values were corrupted by a Gaussian additive noisy contribution with mean and variance equal to 0 and 0.1, respectively. This was done in order to increase both the variability within the simulated data and the generalization ability of the estimator. The simulated samples have been used to train an SVR with Radial Basis Function (RBF) kernel. The values of SVR

Table 3.1: Main characteristics of the three datasets considered in the experimental analysis.

Dataset	Remote Sensing Sensor	# Field Samples	Study area	Surface Roughness [cm]	Dielectric Permittivity (ϵ)
Scatt	Field Scatterometer C-Band Dual-polarization (HH and VV)	17	Bare agricultural fields	$1.3 \leq s \leq 2.5$ $6.5 \leq l \leq 21.0$	$5 \leq \epsilon \leq 15$
SMEX	Airborne SAR L-Band Dual-polarization (HH and VV)	35	Vegetated soybean fields	$0.33 \leq s \leq 1.35$ $0.47 \leq l \leq 16.6$	$5 \leq \epsilon \leq 13$
ActPass	Truch-mounted Radiometer-Scatterometer Multi-frequency Full-polarization	29	Bare and smooth fields	$0.5 \leq s \leq 1.2$ $2.0 \leq l \leq 5.0$	$8 \leq \epsilon \leq 20$

parameters and of the spread of the RBF kernel were obtained by a standard grid search procedure with five-fold cross validation (Smola and Schölkopf [2004]). The simulation of the synthetic backscattering and the training processes were applied independently for each dataset. Then, the inverse function computed by the SVR regression was used to obtain the theoretical estimation of the SMC values. After the SVR regression, the deviations between theoretical model estimates and field reference measurements were characterized according to the GDB and LDB strategies. The k-nn algorithm involved in the LDB strategy is implemented by setting the value of k according to the rule of thumb (Mardia et al. [1982]) and by considering the Euclidean distance. Moreover, due to the limited number of reference samples available for the experiments, we consider a v -fold cross-validation procedure to assess the performance of the proposed approach. To this end, field reference samples were randomly divided into v subsets. Then, $(v - 1)$ subsets were iteratively used for the characterization of the deviation, while the remaining subset was exploited as an independent test set for assessing the performance of the proposed hybrid approach. This procedure was repeated until all reference samples were used as test samples. In the experiments, we varied the number v of folds from 2 to $N - 1$, where N is the total number of reference samples.

The resulting dielectric constant values were compared with the measured reference values in terms of mean and standard deviation of mean absolute error (MAE) and mean squared error (MSE) together with slope and intercept of the linear trend line between estimated and reference target values. We also considered scatter plots of estimated versus measured dielectric constant values for performance evaluation. We compared the results of the proposed method with those obtained by considering only the inverse theoretical model estimates, neglecting the correction procedure based on estimated deviations. In our experiments, the inverse theoretical model is associated to the empirical inversion based on the input-output relationships of the theoretical model that is achieved by the SVR technique.

3.4 Experimental results

3.4.1 Results on the Scatt dataset

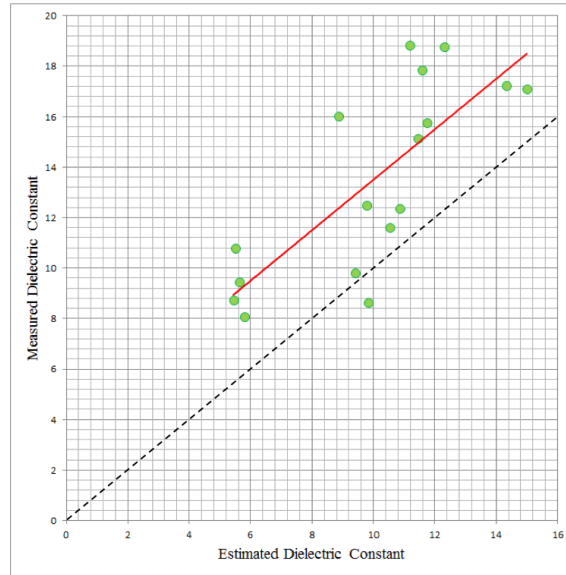
The results of the Scatt dataset has been obtained on two different configurations of features associated to the following polarization of the backscattering coefficient: i) HH acquired at 23° of incident angle (named afterword HH23) and ii) HH and VV both acquired at 23° (named afterword HH23VV23). Table 3.2 shows estimation performances in terms of mean and standard deviation of MAE and MSE together with slope and intercept obtained by the proposed hybrid method with the GDB and LDB strategies

and by the direct use of the inverse theoretical model. All the results in the table are obtained when $v=2$. By analyzing the table, one can observe that the proposed hybrid method with both strategies always yields the highest estimation performance in terms of MAE and MSE compared to the inverse theoretical model estimates under the two feature configurations. As an example, the proposed method with the GDB and the LDB yields MSEs of 5.92 and 4.76, respectively, while the theoretical model provides an MSE of 17.88 when the HH23 features are considered. When the HH23VV23 feature configuration is used, the GDB and LDB strategies reach MSEs of 8.99 and 6.32, respectively, whereas without any correction the theoretical model provides an MSE of 24.42. Moreover, the standard deviations of MSEs and MAEs show that the proposed LDB and GDB results are more stable than those of the standard inverse theoretical model estimations without correction. In greater detail, Table 3.2 shows that the HH23 feature provides in general better performance than the feature HH23VV23. This is due to the fact that HH and VV polarizations provide similar backscattering values. Adding VV polarization at the same incidence angle brings redundant information that can increase the level of noise in the input features (thus decreasing the performances of the regressor) (Notarnicola et al. [2008]).

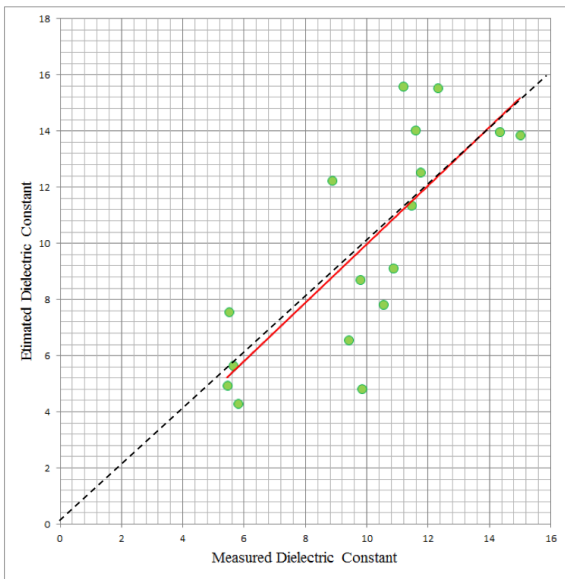
Figure 3.2 shows the scatterplots of measured dielectric constant versus estimated ones obtained by using the inverse theoretical model and the proposed hybrid method with the GDB and the LDB strategies when the HH23 feature is considered. From Figure 2 (related to HH23 feature) one can see that the theoretical model introduces a bias that can be corrected by using the GDB strategy, which is capable of improving the MAE, MSE and intercept metrics. In addition, the LDB strategy further improves the MAE and MSE with respect to the GDB strategy due to its capability to better model the

Table 3.2: Estimation performances obtained by the proposed hybrid approach with the GDB and LDB strategies and the standard inverse theoretical model (Scatt dataset).

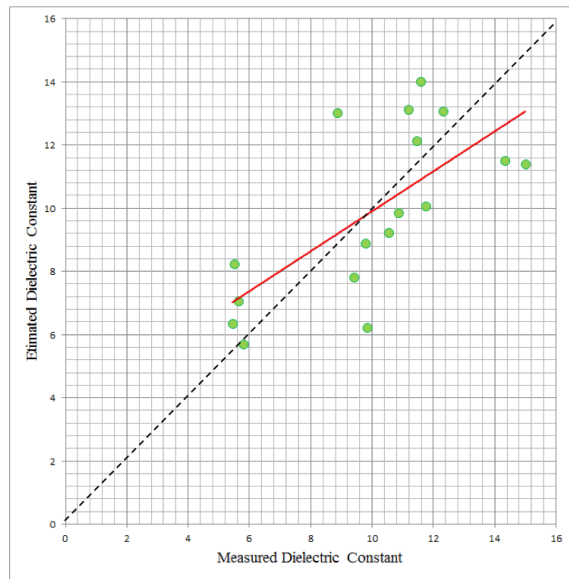
Features	Estimation Strategy		MAE		MSE		slope	intercept
			mean	std	mean	std		
HH23	Inverse Theoretical Model		3.63	0.20	17.88	2.74	0.99	3.49
	Proposed Approach	GDB	1.96	0.14	5.92	0.18	1.04	1.76
		LDB	1.86	0.20	4.76	0.16	0.63	3.35
HH23VV23	Inverse Theoretical Model		3.97	0.67	24.42	10.13	1.12	2.78
	Proposed Approach	GDB	2.42	0.04	8.99	2.56	1.14	-1.37
		LDB	2.07	0.16	6.32	1.03	0.69	3.32



a)



b)



c)

Figure 3.2: Scatterplots of measured versus estimated dielectric constant values obtained by (a) the inverse theoretical model, (b) the proposed hybrid approach with the GDB strategy, and (c) the proposed hybrid approach with the LDB strategy (Scatt dataset with HH23 features).

deviation under the assumption that the deviation is locally i.i.d.

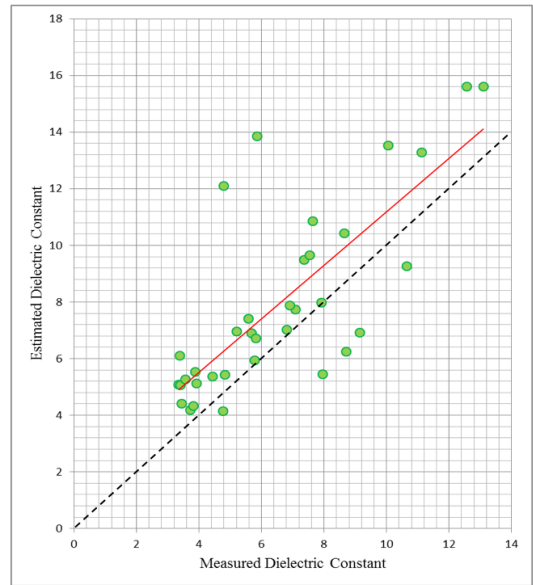
Table 3.3: Estimation performances obtained by the proposed hybrid approach with the GDB and LDB strategies and the standard inverse theoretical model (SMEX dataset).

Estimation Strategy		MAE		MSE		slope	intercept
		mean	std	mean	std		
Inverse Theoretical Model		1.90	0.74	6.42	5.26	0.94	1.76
Proposed Approach	GDB	1.56	0.58	5.11	3.52	0.96	0.23
	LDB	1.27	0.42	3.52	2.32	0.71	1.98

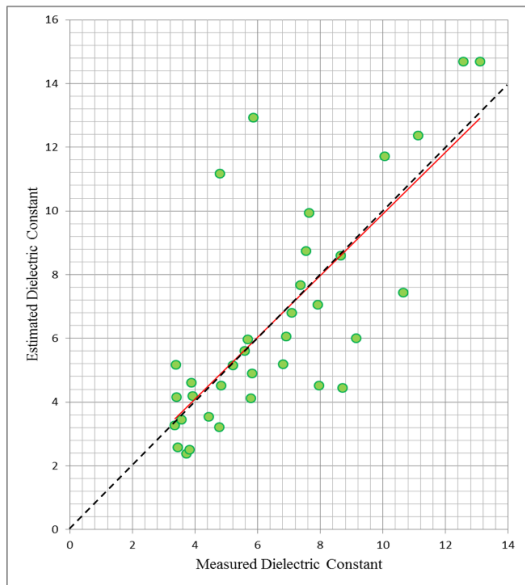
3.4.2 Results on the SMEX dataset

The experiments on the SMEX dataset are obtained by using a single feature configuration that is the dielectric permittivity from the data in HH and VV polarizations. Table 3.3 shows mean and standard deviation of MAE and MSE together with slope and intercept obtained by the proposed hybrid method with GDB and LDB strategies and the inverse theoretical model. All the results in the table are obtained when $v=2$. From the table one can observe that also in this case the proposed method with both strategies provides the best estimation performance in terms of MAE and MSE when compared to the theoretical model. As an example, the proposed method with the GDB and the LDB yields MAEs of 1.56 and 1.27, respectively, while the inverse theoretical model provides an MSE of 1.90. By analyzing Table 3.3, we can also observe that the MAEs and MSEs obtained with the proposed GDB and LDB methods are also more stable than those obtained by the standard inverse theoretical model without correction. A global analysis of the estimation performance points out that the proposed method achieves smallest MAE and MSE considering the LDB strategy (due to considering locally i.i.d. deviations), whereas the GDB strategy provides the best slope and intercept values (which are 0.96 and 0.26, respectively).

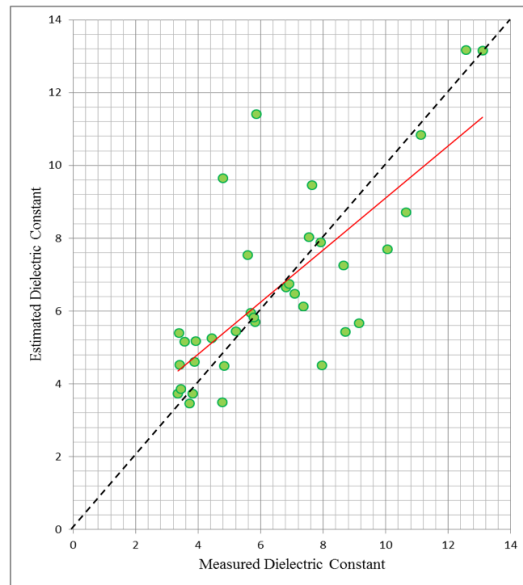
Figure 3.3 shows the scatterplots of estimated versus measured dielectric constant values. Similarly to the case given in Figure 3.2, the theoretical model results are affected by bias errors, while the LDB strategy leads to a strong improvement on the estimation accuracy and a significant reduction on the overestimation trend for high dielectric constant values. In the case of the GDB strategy, from the figure one can see that a slight overestimation of high dielectric constant values can be observed in the final output estimates. Similar behaviors are observed with different values of v . These results confirm the capability of the proposed hybrid method to improve estimation of performances of theoretical models.



a)



b)



c)

Figure 3.3: Scatterplots of measured versus estimated dielectric constant values obtained by (a) the inverse theoretical model, (b) the proposed hybrid approach with the GDB strategy, and (c) the proposed hybrid approach with the LDB strategy (SMEX dataset).

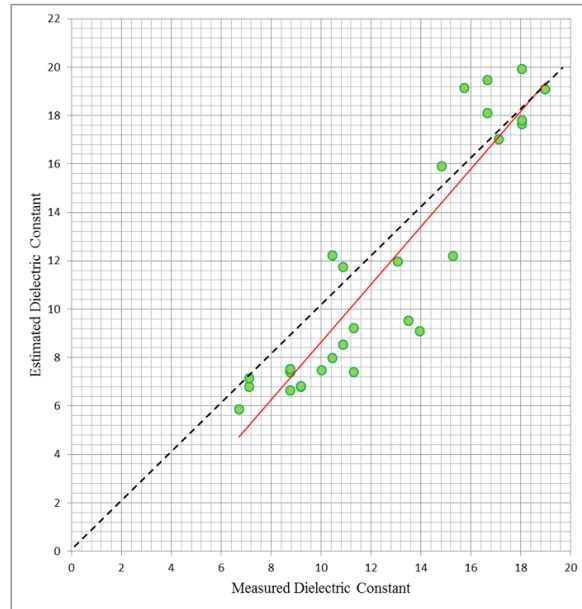
3.4.3 Results on the ActPass dataset

The experiments on the ActPass dataset are carried out by considering only the active components of the data acquired by the radiometer-scatterometer. This choice has been done according to the properties of the IEM theoretical model, which is specifically designed to simulate the backscattering coefficient for active systems. Among the scatterometer data, three different polarizations with different incidence angles are exploited in the experiments: i) data at HH polarization (named HH), ii) data at VV polarization (named VV) and iii) data at HH and VV polarizations (named HHVV). Table 3.4 reports the obtained mean and standard deviation of MAE and MSE together with slope and intercept obtained by the proposed hybrid method with GDB and LDB strategies and the inverse theoretical model. All the results in the table are obtained when $v=2$. By analyzing the table, one can observe that the proposed method with the GDB and LDB strategies, again, improves the results obtained by the theoretical model in terms of MAE and MSE. In addition, the LDB strategy leads to the smallest MAE and MSE. As an example, when the HHVV feature is considered, the LDB strategy yields an MAE of 1.36 and an MSE of 2.86, whereas the GDB and the theoretical model provide MAEs of 1.69 and 1.83, respectively, and MSEs of 4.18 and 4.93, respectively. One can see that the proposed LDB results in the smallest MAE and MSE among the difference feature configurations and also results more stable than the standard theoretical model without correction.

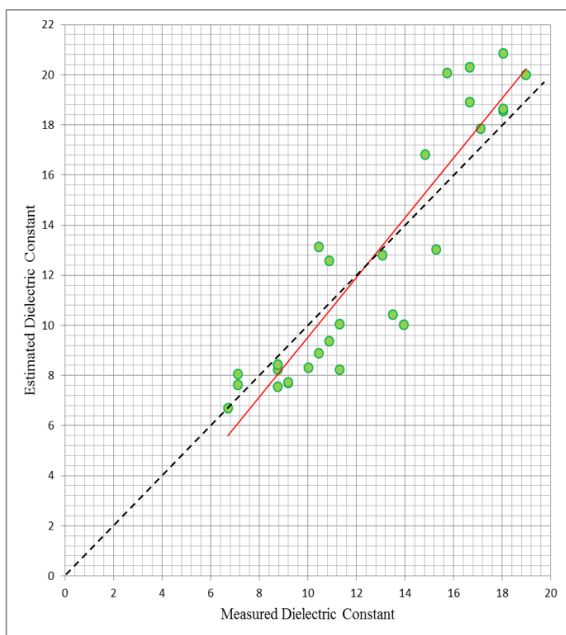
Figure 3.4 shows the scatterplots measured versus estimated dielectric constant values by the inverse theoretical model and the proposed method with GDB and LDB strategies.

Table 3.4: Estimation performances obtained by the proposed hybrid approach with the GDB and LDB strategies and the standard inverse theoretical model (ActPass dataset).

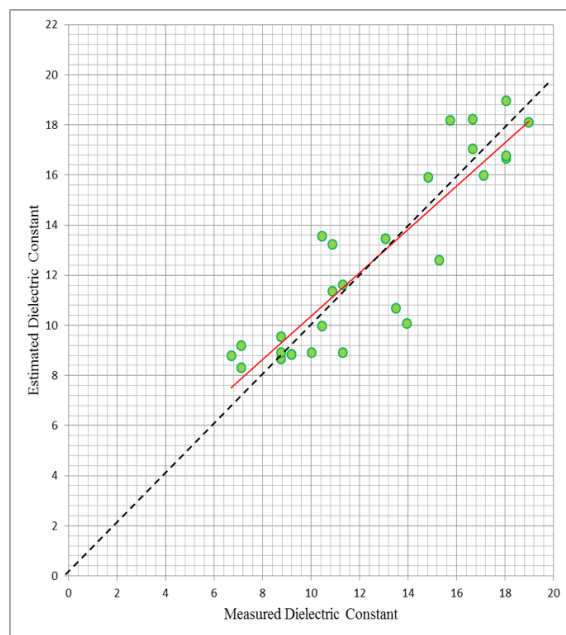
Features	Estimation Strategy		MAE		MSE		slope	intercept
			mean	std	mean	std		
HH	Inverse Theoretical Model		1.69	0.11	4.40	1.19	0.73	2.88
	Proposed Hybrid Approach	GDB	1.59	0.08	4.38	0.86	0.72	3.41
		LDB	1.54	0.08	4.01	0.59	0.78	2.78
VV	Inverse Theoretical Model		1.47	0.48	3.13	1.56	0.91	0.66
	Proposed Approach	GDB	1.41	0.47	2.86	1.56	0.91	1.04
		LDB	1.33	0.36	2.59	1.18	0.87	1.68
HHVV	Inverse Theoretical Model		1.83	0.26	4.93	1.35	1.19	-3.27
	Proposed Approach	GDB	1.69	0.26	4.18	1.36	1.19	-2.40
		LDB	1.36	0.15	2.86	0.85	0.87	1.70



a)



b)



c)

Figure 3.4: Scatterplots of measured versus estimated dielectric constant values obtained by (a) the inverse theoretical model, (b) the proposed hybrid approach with the GDB strategy, and (c) the proposed hybrid approach with the LDB strategy (ActPass dataset with the HHVV feature).

The scatterplots confirm the results on the quantitative analyses given in Table 3.4. Moreover, the figure shows that, when the theoretical backscattering model underestimates low values of dielectric permittivity and overestimates the high ones, the best performances are obtained with the LDB strategy. This is due to considering the deviation that is locally i.i.d. in a given position. Similar behaviors are observed with different values of v . These results further confirm that the proposed method significantly improves the performance of the theoretical model.

3.5 Conclusion

In this Chapter, a novel hybrid method that integrates theoretical analytical models and empirical observations is presented for an accurate estimation of biophysical parameters from remotely sensed data. The proposed method exploits the available field reference samples to characterize deviations between theoretical model based estimates (affected by biases and approximations intrinsic within the analytical formulation of theoretical models) and the measured true target values. To this end, we have presented two different strategies: i) global deviation bias (GDB) and ii) local deviation bias (LDB). The GDB strategy assumes that the deviation is globally i.i.d. in the feature space and thus estimates a constant for all samples to globally recover the bias of the theoretical model. The LDB strategy considers the more realistic assumption that the deviation is varying in different portions of the feature space whereas it is locally i.i.d. in a given position. Accordingly, the amount of deviation depends on the samples location in the feature space. However, the LDB strategy requires a higher number of reference samples than the GDB strategy to accurately model the different local deviations. Thus, in the case of availability of a small number of reference samples, the simple GDB can be preferred. After the deviations are computed, theoretical model estimates of samples are corrected based on these deviations.

The proposed hybrid method preserves the robustness and the generality of theoretical model estimates derived by physics foundations, while reducing the bias and imprecision occurred due to the simplifying assumptions of the models. It is worth noting that the proposed method requires a relatively small number of samples. Thus, the cost for collecting target measurements is affordable in many applications.

In the experimental analysis, we studied the effectiveness of the proposed hybrid method to estimate soil moisture from microwave remotely sensed data. From the results obtained on three different datasets we observed that the proposed hybrid method is promising as: i) it allows improved estimation performance by reducing biases and overcoming the effects of simplifications intrinsic within the analytical formulation of the-

oretical models; ii) it handles effectively the variability of the deviation function in the feature space; and iii) it is simple, easy to implement and fast during the processing. Although the experimental analysis has been carried out in a specific application domain, it is important to emphasize that the proposed hybrid method is general and can be also applied to other biophysical parameter estimation problems.

We would like to point out that the proposed hybrid method can be applied to different bio-geo physical parameter estimation problems by considering proper theoretical electromagnetic models (which should guarantee the considered i.i.d. assumption) associated to the considered problem. As a final remark, we would like to note that the proposed hybrid method achieves the estimation (and thus correction) of globally or locally uniform deviations by the use of the GDB and LDB strategies, respectively.

As a future development of this work, we plan to analyze the possible extension of our method to the use of other functional forms (polynomial, exponential, etc.). Moreover, we plan i) to study other techniques for the estimations of deviations based on different assumptions on the distributions of deviations in the feature space, and ii) to assess the effectiveness of the proposed hybrid method in other estimation problems (e.g. the retrieval of snow pack variables in mountain areas).

Chapter 4

A novel semisupervised method for the estimation of biophysical parameters based on Support Vector Regression

This Chapter provides the second contribution of the thesis in the context of the estimation of biophysical parameters. It presents a novel semisupervised method for the estimation of biophysical parameters based on SVR. The proposed method is made up of two steps: 1) the injection of priors information in the learning phase of SVR in order to adapt the importance of each training sample according to distribution of the unlabeled samples; 2) the jointly exploitation of labeled and informative unlabeled samples for further improving the definition of the WSVR learning function. In the first step each training sample is weighted on the basis of a novel strategy. Its goal is to define higher weights for the samples located in the high density regions of the feature space and smaller weights to those located in low density regions of the feature space. To exploit different weights in the learning phase of the SVR, the weighted SVR (WSVR) algorithm is considered. In the second step, the most informative unlabeled samples are selected. This is done using a novel strategy that relies on the distribution of the unlabeled samples in the feature space and on the WSVR function estimated at the first step. Then, the selected unlabeled samples and the label samples are included in the learning phase of the WSVR algorithm, which is redefined to tune the importance of the unlabeled samples using different values of regularization parameters. Experimental results obtained on two different datasets demonstrate the effectiveness of the proposed method to mitigate the problems of small-sized biased training sets.

4.1 Introduction

Accurate estimation of parameters from specific data, such as geo/biophysical parameters (e.g., soil moisture and forest parameters from remote sensing data) and biomedical parameters (e.g., glucose concentration in diabetic patients from biomedical data) is an important research field in pattern recognition. Parameter estimation can be achieved by supervised regression techniques, e.g., neural networks (Egmont-Petersen et al. [2002], Kwok and Yeung [1997]), Gaussian Process (GP) regression (Rasmussen [2006]) and Support Vector Regression (SVR) (Drucker et al. [1997], Smola and Schölkopf [2004], Collobert et al. [2001]). However, such techniques require the availability of training samples (i.e., samples for which reference measures are available) to be used in the learning phase of the supervised regression methods. The learning phase of a regression method aims at estimating a functional relationship between a set of variables and corresponding target values (i.e., the measurement of the parameters of interest) based on the training samples. Accordingly, the amount and the quality of these samples are crucial to obtain accurate estimations. However, collecting an adequate number of high quality training samples is often costly, time consuming (as producing reference measures may require significant time) and unfeasible in some cases. An insufficient number of training samples or the availability of biased samples may lead to poor estimation performance in terms of accuracy and generalization ability (Demir and Bruzzone [2014]).

To reduce the need and effort to collect training samples, semi-supervised learning (SSL) techniques have been presented in the literature Zhu [2010]. SSL methods jointly exploit training and unlabeled samples in the learning phase of the regression algorithm to describe a regression function that consider both the information present in the training set and the structure of all data in the feature space. Most of the previous studies in SSL have been developed in the context of classification problems (Mingmin and Bruzzone [2007], Cohen et al. [2004], Bennett and Demiriz [1998]). However, SSL has been marginally considered in regression problems. In Bazi et al. [2012] a semi-supervised regression framework that integrates multi-objective optimization and Gaussian processes is introduced for estimating chlorophyll concentrations in subsurface waters. The training phase of the GP regressor is achieved by exploiting the unlabeled samples are in conjunction with training samples. Estimation of the target values related to the unlabeled samples is achieved by the simultaneous optimization of two different criteria that are: 1) the empirical risk quantified in terms of the mean square error measure and 2) the marginal likelihood (ML). In Kang et al. [2016] a SSL method is presented in the framework of SVR problems with a specific application to virtual metrology in semiconductor manufacturing. The method evaluates the uncertainty of the target values of unlabeled samples based on

their distribution estimated with two probabilistic local reconstruction models. In Blum and Mitchell [1998] and Zhou and Li [2005], two different regression models are generated by exploiting sub-feature sets. In Zhou and Li [2005], a SSL method that exploits two k-nearest neighbor regressors with different distance metrics is presented. Each regressor assigns target values to the unlabeled samples for the other regressor, where the confidence of the target values of the unlabeled samples is estimated through the amount of reduction in mean squared error over the labeled samples located at neighborhood of that sample. A similar SSL approach is introduced in Wang et al. [2010] by exploiting two SVRs for the estimation of water quality parameters. In Wang et al. [2010] the authors propose an algorithm in the context of SVR framework, while in Zhou and Li [2007] a method called COREG which exploits the k-NN algorithm to compute the co-training regression models is presented. In Camps-Valls et al. [2009], two SSL methods for the estimation of biophysical parameters are introduced in the context of SVR. These methods are defined based on a graph or hypergraph Laplacian that is constructed by both the available labeled and unlabeled samples and used to deform the training kernel matrix. Then, the deformed kernel is exploited in the framework of SVR. Similarly to the semi-supervised method, also transductive regression methods (where the unlabeled test samples are known in the training phase of the regression) have been presented in the literature. For instance, in Cortes and Mohri [2006] a VapnikChervonenkis dimension error bounds for transductive regression that hold for all bounded loss functions is introduced, which has been inspired by the bounds defined in Vapnik [1995] for classification problems.

The effectiveness of the above-mentioned semi-supervised methods relies on the fact that the available training samples should not be too distant from the correct representation of the considered regression problem. It is crucial since the training samples may pose a bias on the part of the feature space. In the case that these samples are too distance from the real regression function, the risk is to find a correct solution from the theoretical viewpoint that does not model the property of the considered problem. Accordingly, when the available training set is incomplete and biased, these methods can provide unsatisfactory results.

In this Chapter, to deal with the above-mentioned critical problems, where small size biased training sets are available, we propose a novel SSL method. The proposed method is defined in the context of the ϵ -insensitive SVR, since it demonstrates good generalization capability and ability to handle high dimensional input spaces also when few training samples are existing (Demir and Bruzzone [2014]).

The proposed method is made up of two steps: i) injection of additional priors information in the initial learning of the SVR function; and ii) exploitation of labeled and unlabeled samples in the learning phase of the SVR. The first step aims at adapting the

importance of each training sample with respect to the distribution of the unlabeled samples to overcome the problems occurred due to biased and small-sized training sets. This is achieved by inserting additional priors information in the initial learning of the SVR that is described by using a novel strategy. According to this strategy, a weight value is initially defined associated to each training sample, where higher weights are given to the training samples located in the high-density regions of the feature space and vice versa. Then, to exploit the weight of each training sample in the learning phase of the SVR, we consider a weighted SVR (WSVR) algorithm previously presented in the literature Yang et al. [2007]. The WSVR gives more importance to the training samples with high weights and less importance with those of lower weights while modeling the SVR learning function. Thanks to this step, representativeness of the true underlying distribution in the case of initial biased training sets is better achieved, i.e., the structure of the unlabeled data is captured without any bias on the training samples while modeling the learning function of SVR. To further improve the definition of the SVR learning function, the second step aims to jointly use the labeled and unlabeled samples in the framework of the WSVR learning phase. This is achieved by adding a new term in the WSVR learning that exploits a subset of informative unlabeled samples (with their “pseudo” target values estimated at the first step). In the restructured WSVR, the labeled and unlabeled samples are associated to the different values of regularization parameters, i.e., a lower value of regularization parameter is assigned to unlabeled samples when compared to that of training samples. This choice is done to reduce the importance assigned to the unlabeled samples for which real target values are not existing. In the experiments, the performance of the proposed SSL method is demonstrated with respect to estimation of i) age of the abalones from the variables that model the physical characteristics of abalones and ii) soil moisture content from the variables that model the microwave remotely sensed data. It is worth noting that the most of the previous SSL studies in the framework of Support Vector Machine (SVM) are introduced for classification problems. However, these techniques cannot be directly used for SVR problems, because: i) the decision rule of SVM classification is different from that of SVR and ii) classification and regression problems are essentially different. Thus, in this Chapter we extend the use of SSL to the framework of SVR.

The remaining of the Chapter is organized into four sections. Section 4.2 presents the problem definition by introducing the ε -insensitive SVR. Section 4.3 introduces the proposed SSL method. Section 4.4 and 4.5 describe the considered dataset and illustrates the experimental results, respectively. Finally, Section 4.5 draws the conclusion of this work.

4.2 Problem definition

Let us assume that a small-sized and biased training set L made up of n pairs $\{(\mathbf{x}_i, y_i)\}_{i=1}^n$ is initially available, where $\mathbf{x}_i \in \mathfrak{R}^d$ is the i -th training sample, $y_i \in \mathfrak{R}$ is the corresponding target value and n is the number of training samples. The goal of the proposed SSL method is to include in the learning phase both samples in the training set L and those in a set U of u unlabeled samples $\{\mathbf{x}_j\}_{j=1}^n$ with $u \gg n$ to improve the estimation performance of the regression function. The proposed SSL method does not require that the real target values $\{y_j\}_{j=1}^n$ for unlabeled samples are known a priori. The proposed SSL method is defined in the framework of the ε -insensitive SVR that defines a function in the learning phase which is as flat as possible and has, at most, deviation ε from the real targets for all training samples. Thus, the original data is projected into a higher dimensionality space to find a linear regression function in this space as (Smola and Schölkopf [2004]):

$$f(\mathbf{x}) = \mathbf{w}\Phi(\mathbf{x}) + b \quad (4.1)$$

where $\mathbf{w} \in \mathfrak{R}^d$, $\Phi(\cdot)$ is the mapping function that projects the samples from the original variable space to a higher dimensional space, and b is the bias. The optimal function $f(\mathbf{x})$ in the higher dimensional space can be obtained based on the concept of soft margin by solving the following minimization problem:

$$\min \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + \sum_{i=1}^n C^L (\xi_i^L + \xi_i^{L*}) \right\} \quad (4.2)$$

$$\text{subject to } \begin{cases} y_i - (\mathbf{w} \Phi(\mathbf{x}_i) + b) \leq \varepsilon + \xi_i^L, \\ (\mathbf{w} \Phi(\mathbf{x}_i) + b) - y_i \leq \varepsilon + \xi_i^{L*}, & i = 1, 2, \dots, n \\ \xi_i^L, \xi_i^{L*} \geq 0, \end{cases} \quad (4.3)$$

where ξ_i^L and ξ_i^{L*} are slack variables that evaluate the distance of the training samples situated outside the ε -insensitive tube from the tube itself. $C^L > 0$ is a regularization (i.e., penalty) parameter that controls the tradeoff between the flatness of the function $f(x)$ and the tolerance to empirical errors. The minimization problem in (4.2) can be described in a dual formulation and solved by standard methods of quadratic programming. When a linear penalty function is considered, Lagrangian function is defined as:

$$\left\{ L(\alpha, \alpha^*) = \sum_{i=1}^n y_i(\alpha_i - \alpha_i^*) - \sum_{i=1}^n \varepsilon(\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(\mathbf{x}_i, \mathbf{x}_j) \right\} \quad (4.4)$$

where α_i and α_i^* , are the Lagrange multipliers, and $K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)\Phi(\mathbf{x}_j)$ is the kernel function that computes the similarity between samples in the transformed higher dimensionality space as a function of the samples in the original space without the explicit

definition of the mapping function $\Phi(\cdot)$. The minimization problem in the dual formulation becomes as follows:

$$\max \{L(\alpha, \alpha^*)\} \tag{4.5}$$

$$\text{subject to } \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*), \\ 0 \leq \alpha_i \leq C^L, & i = 1, 2, \dots, n \\ 0 \leq \alpha_i^* \leq C^L, \end{cases} \tag{4.6}$$

Finally, the estimation function is written as:

$$f(x) = \sum_{n=1}^n (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}_j) + b \tag{4.7}$$

where the training samples that are associated to the Lagrangian multipliers different from zero are known as support vectors (SVs) that are outside or on the boundary of the ε -insensitive tube from a geometrical viewpoint the SVs. The proposed SSL method redefines the ε -insensitive SVR to: 1) to adapt the importance of each training sample according to the distribution of the unlabeled samples to overcome the problems of small-sized biased training sets; and 2) to jointly exploit the labeled and unlabeled samples to further improve the regression function. The details of the proposed method are given in the following section.

4.3 Proposed SSL method for regression

We propose a novel SSL method in the context of the ε -insensitive SVR (SS-SVR) to overcome the problems of small sized-biased training sets without collecting additional samples with reference measures. The proposed method is defined based on two consecutive steps: i) injection of additional priors information in the initial learning of the SVR function; and ii) exploitation of labeled and unlabeled samples in the learning phase of the SVR to further improve the definition of the SVR learning function. The first step aims to exploit the additional prior information that is characterized by determining the importance of each training sample according to the distribution of the unlabeled and labeled samples in the feature space. To model the importance of training samples a novel clustering-based strategy is introduced, while to integrate the obtained weights within the learning phase of the SVR, a weighted SVM (WSVM) algorithm (Yang et al. [2007]) is adapted. The second step aims to further improve the definition of the SVR learning function by jointly using labeled and informative unlabeled samples. To this end, a novel strategy is developed to select the most informative unlabeled samples according to the SVR properties. Then, to jointly use these samples with label samples in the learning

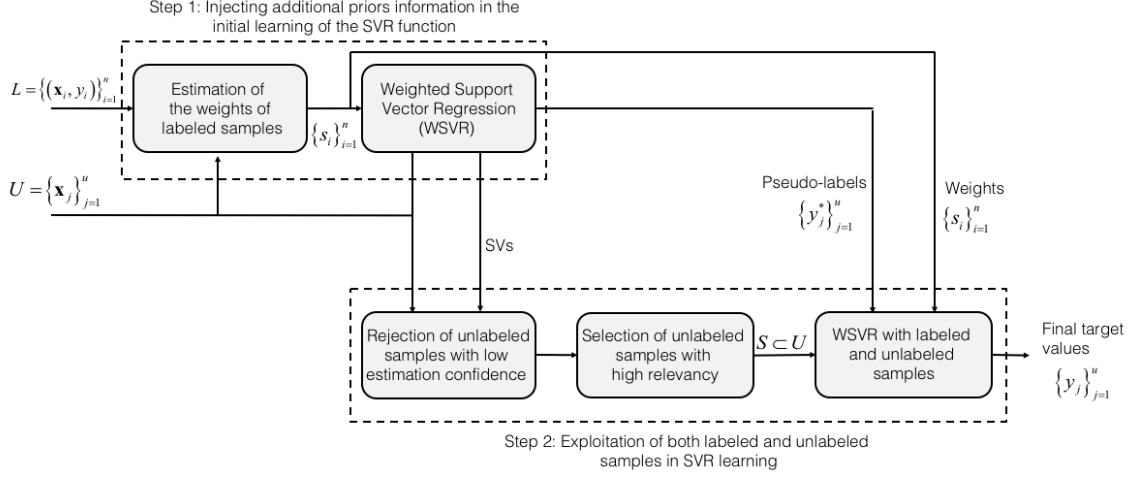


Figure 4.1: Block scheme of the proposed SSL method.

phase of the SVR, the WSVR algorithm is redefined. Figure 4.1 shows the block scheme of the proposed SSL method. In the following subsections, we present in further details each step of the proposed SSL method.

4.3.1 Injecting priors information in the initial learning of the SVR function

The first step is devoted to exploit additional prior information in the learning of the SVR function in order to overcome the problems occurred due to the small-sized biased training sets. To effectively characterize the priors we evaluate the importance of each training sample according to the distribution of the unlabeled samples in the feature space. Training samples that are located in high density regions of the feature space are more informative than those located in the low density regions. This is due to the fact that i) samples in the high density regions provide a better statistical representation of the underlying sample distribution; and ii) the estimation results on samples located in the high density regions of the feature space affect more the overall accuracy of the estimation process than those obtained on samples within low density regions. Accordingly, we consider that the importance of each training sample \mathbf{x}_i is associated the density d_i of the portion of the feature space where that sample is located. To estimate the density d_i of each training sample \mathbf{x}_i , the number of neighbor samples to the sample \mathbf{x}_i within a specified radius R in the feature space is computed. To define the value of R an iterative algorithm is adopted. This algorithm assumes that: 1) the portion of the feature space defined by the radius R should include only one training sample; and 2) each unlabeled sample should belong to only one portion this radius. To this end, each training sample

is initially considered as located at the center of the portion of the feature space and the initial value of R is defined as the average of the distances estimated between the training samples. The feature space is divided by the number of portion that is equal to the number of training samples. Then, the value of the radius R is reduced by a factor ζ . This process is iterated until each area include only one training sample, i.e., until there are no overlapping portions associated to more than one training sample. This choice guarantees that an unlabeled sample can belong to only one portion. It is worth noting that some of the unlabeled samples can remain outside the portions and thus they will not have any role in the importance estimation of the training samples. Once the radius R is set, the number $u_{i,R}$ of the unlabeled samples in i -th portion is calculated and then the density d_i is computed as $d_i = u_{i,R}/u_{i,R}^t$, where $u_{i,R}^t$ is the total number of unlabeled sample included in all areas defined by the radius R . Then, the density d_i of each sample \mathbf{x}_i is taken as a weight value s_i for that sample, i.e., $s_i = d_i$. In this way, training samples located high density regions of the feature space are associated to a high weight value and vice-versa. In order to exploit the weights of training samples in the SVR learning phase, we exploit the weighted SVR (WSVR) algorithm. The WSVR extends the SVR to the case in which different training samples give different contributions to the learning of the ε -insensitive tube. Accordingly, optimization problem given in (4.2) is redefined as:

$$\min \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + C^L \sum_{i=1}^n s_i (\xi_i^L + \xi_i^{L*}) \right\} \quad (4.8)$$

$$\text{subject to} \begin{cases} y_i - (\mathbf{w} \Phi(\mathbf{x}_i) + b) \leq \varepsilon + \xi_i^L, \\ (\mathbf{w} \Phi(\mathbf{x}_i) + b) - y_i \leq \varepsilon + \xi_i^{L*}, & i = 1, 2, \dots, n \\ \xi_i^L, \xi_i^{L*} \geq 0, \end{cases} \quad (4.9)$$

The main difference with respect to (4.2) is that the penalty value C^L of misestimation for each training sample \mathbf{x}_i has a different weighing term that is driven by s_i . Thus, the training samples with higher weights (which are associated with the portion of feature space with the highest density) are associated to high penalty values for their misestimations, whereas training samples with smaller weights are associated to small penalty values. In the WSVR, the Lagrangian function is the same as (4.4), whereas the constraints in the maximization problem are changed to:

$$\max \{L(\alpha, \alpha^*)\} \quad (4.10)$$

$$\text{subject to} \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*), \\ 0 \leq \alpha_i \leq C^L, & i = 1, 2, \dots, n \\ 0 \leq \alpha_i^* \leq C^L, \end{cases} \quad (4.11)$$

And finally the WSVR estimation function is defined as:

$$f(x) = \sum_{n=1}^n (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}_j) + b \quad (4.12)$$

The WSVR regression function obtained in (4.12) is applied to all unlabeled samples and then estimated target values y_i^* , $j = 1, 2, \dots, u$ are obtained to be used in the second step of the proposed method. It is worth noting that unlike the standard ε -insensitive SVR (where all training samples have the equal penalty), this step of the proposed method can avoid the problems occurred due to the availability of the small-sized biased training sets injecting the priors in the initial learning of the SVR. If $s_i = 1$, $i = 1, 2, \dots, n$ the WSVR results to be the same as the standard SVR.

In Figure 4.2 a qualitative 1-D example is given, where Figure 4.2 (a) represents the distribution of labeled samples (circles in black) and unlabeled samples (squares in green) in the feature space and Figure 4.2 (b) shows the definition of R . Figure 4.2 (c) shows the ε -insensitive tube obtained by applying the proposed WSVR algorithm. As we can observe from the figure, the WSVR learning function provide a good characterization of

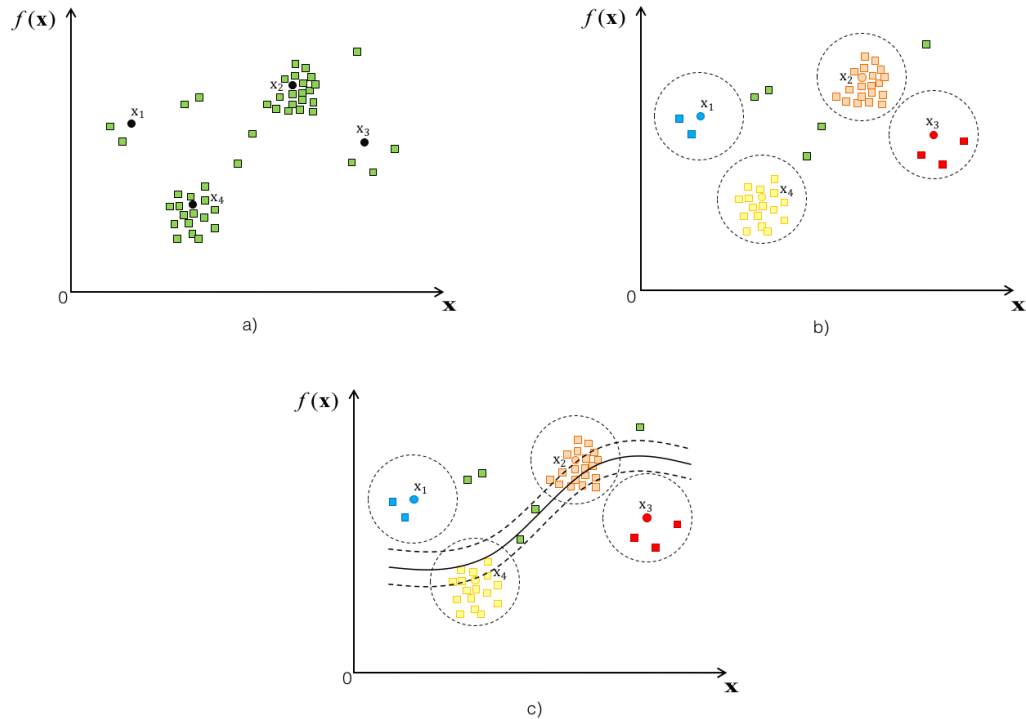


Figure 4.2: A qualitative example in a 1-D feature space: a) distribution of the biased data; the unlabeled samples are the squares in green and the labeled samples are the circles in black. b) Radius definition to estimate the weight of label samples. c) the WSVR learning function.

the position of the unlabeled samples in the feature space. This is done by studying the density of the region that each training sample exploiting the approach presented above. The first step of the proposed method can be summarized as follows:

Algorithm 1: Step 1 of the proposed SSL method

1: **Inputs:**

L (*Training samples set*)

U (*Unlabeled samples set*)

2: **Outputs:**

SVs (*Support Vectors SV*)

y_j^* , $j = 1, 2, \dots, u$ (*Estimated target values of the unlabeled samples*)

s_i (*Weights of training samples*)

3: **Begin**

4: Initialize the radius R as the mean distance among all the training samples \mathbf{x}_i

5: If there is any unlabeled sample \mathbf{x}_j that belongs more than one portion defined by the radius R , reduce the radius value of ζ . Repeat 4 until each unlabeled sample is associated to only one portion defined by R

6: Count the number of samples $u_{i,R}$ present in each portion

7: Estimate the density d_i as $d_i = u_{i,R}/u_{i,R}^t$. Then, the density d_i of each sample \mathbf{x}_i is taken as the weight value s_i , i.e., $s_i = d_i$

8: Train the WSVR with the training samples by considering their weights

9: Estimate pseudo target values to the unlabeled samples by using (4.12)

10: **End**

4.3.2 Exploitation of both labeled and unlabeled samples in the learning process of the regression function

The second step of the proposed SSL is devoted to better model the WSVR function obtained at the first step by exploiting the training samples with their weights and a subset $S \subset U$ of m informative unlabeled samples with their pseudo target values (which are estimated at the first step). The unlabeled samples are considered as informative if they have: 1) high certainty on their estimated (i.e., pseudo) target values (i.e., the WSVR has high confidence to correctly estimate the target values of these samples); 2) high relevancy, (i.e., they provide additional information to the definition of the WSVR function). In SVR problems, the samples inside the ε -tube do not have any role in determining the SVR function, while the samples outside or on the boundary of the ε -tube model the SVR function. In addition, the samples on or close to the boundary have

the largest confidence to be correctly estimated among the samples outside the ε -tube. Accordingly, we aim to select samples that are located as close as possible to the ε -tube and outside of the ε -tube. This is achieved by a novel strategy that includes two sub-steps: i) rejection of unlabeled samples with low confidence on their estimated target values, and ii) selection of unlabeled samples with high relevancy.

The unlabeled samples closest to SVs that are located on the boundary of the ε -tube can have high confidence on their estimated target values, while those close to the SVs that are outside the tube can be associated to the low confidence on their estimated target values. Thus, in order to analyze the confidence of an unlabeled samples, we initially identify the closest SV associated to each unlabeled sample. Then the unlabeled samples for which the closest SV is located outside the boundary of the ε -tube are considered as uncertain and thus rejected. To determine the locations of the SVs with respect to ε -tube, the information on the Lagrange multipliers of the SVs and the SVR regularization parameter value is considered. This is due to the fact that SVs that have Lagrange multipliers smaller than the SVR regularization parameter value are known as located on the ε -tube, while those that are associated to Lagrange multipliers larger than the SVR regularization parameter value is located outside the ε -tube. From the remaining unlabeled samples, we select the subset of m unlabeled samples, which have high relevancy. This is achieved by selecting most relevant $c \leq m$ samples associated to each SV located on the tube, where the number c is computed proportionally to the total number m of samples and the density of the related SV in the feature space. By this way, more samples that are associated with portions of the feature space having high density of samples are selected as relevant unlabeled samples.

Figure 4.3 shows a qualitative example of the unlabeled samples selection strategy. To exploit together the training samples with their associated weights and the set of informative unlabeled samples in the learning of the regression function, we introduce the restructured WSVR that redefines the optimization problem given in (4.8) as:

$$\min \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + C^L \sum_{i=1}^n s_i (\xi_i^L + \xi_i^{L*}) + C^S \sum_{j=1}^n (\xi_j^S + \xi_j^{S*}) \right\} \quad (4.13)$$

$$\text{subject to} \begin{cases} y_i - (\mathbf{w} \Phi(\mathbf{x}_i) + b) \leq \varepsilon + \xi_i^L, \\ (\mathbf{w} \Phi(\mathbf{x}_i) + b) - y_i \leq \varepsilon + \xi_i^{L*}, & i = 1, 2, \dots, n \\ \xi_i^L, \xi_i^{L*} \geq 0, \\ y_j - (\mathbf{w} \Phi(\mathbf{x}_j) + b) \leq \varepsilon + \xi_j^U, \\ (\mathbf{w} \Phi(\mathbf{x}_j) + b) - y_j \leq \varepsilon + \xi_j^{U*}, & j = 1, 2, \dots, m \\ \xi_j^U, \xi_j^{U*} \geq 0, \end{cases} \quad (4.14)$$

where, ξ_j^U and ξ_j^{U*} are the slack variables and C^U is the penalty parameter of unlabeled samples in S. Note that the set S of informative unlabeled samples is associated with the same regularization parameter C^U , i.e., the same amount of penalty is assigned to all the selected unlabeled samples. The minimization problem in (4.14) can be written in a dual formulation and solved according to standard methods of quadratic programming. In the case of a linear penalty function, Lagrangian function is obtained as:

$$\begin{cases} L(\alpha, \alpha^*, \beta, \beta^*) = \sum_{i=1}^n y_i(\alpha_i - \alpha_i^*) - \sum_{i=1}^n \varepsilon(\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_j^*(\beta_j - \beta_j^*) \\ - \sum_{i=1}^n \varepsilon(\beta_j + \beta_j^*) - \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(\mathbf{x}_i, \mathbf{x}_k) \\ + \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*)(\beta_j - \beta_j^*) K(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{2} \sum_{j=1}^n \sum_{z=1}^n (\beta_j - \beta_j^*)(\beta_z - \beta_z^*) K(\mathbf{x}_i, \mathbf{x}_z) \end{cases} \quad (4.15)$$

where β and β^* , are the Lagrange multipliers for the informative unlabeled samples. The minimization problem in the dual formulation becomes:

$$\max \{L(\alpha, \alpha^*, \beta, \beta^*)\} \quad (4.16)$$

$$\text{subject to} \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*), \\ 0 \leq \alpha_i \leq C^L, \quad i = 1, 2, \dots, n \\ 0 \leq \alpha_i^* \leq C^L, \\ \sum_{j=1}^m (\beta_j - \beta_j^*), \\ 0 \leq \beta_j \leq C^U, \quad j = 1, 2, \dots, m \\ 0 \leq \beta_j^* \leq C^U, \end{cases} \quad (4.17)$$

Finally, after the Lagrange multipliers are obtained, the final estimation function of the restructured WSVR is defined as:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + \sum_{j=1}^m (\beta_j - \beta_j^*) K(\mathbf{x}_j, \mathbf{x}) + b \quad (4.18)$$

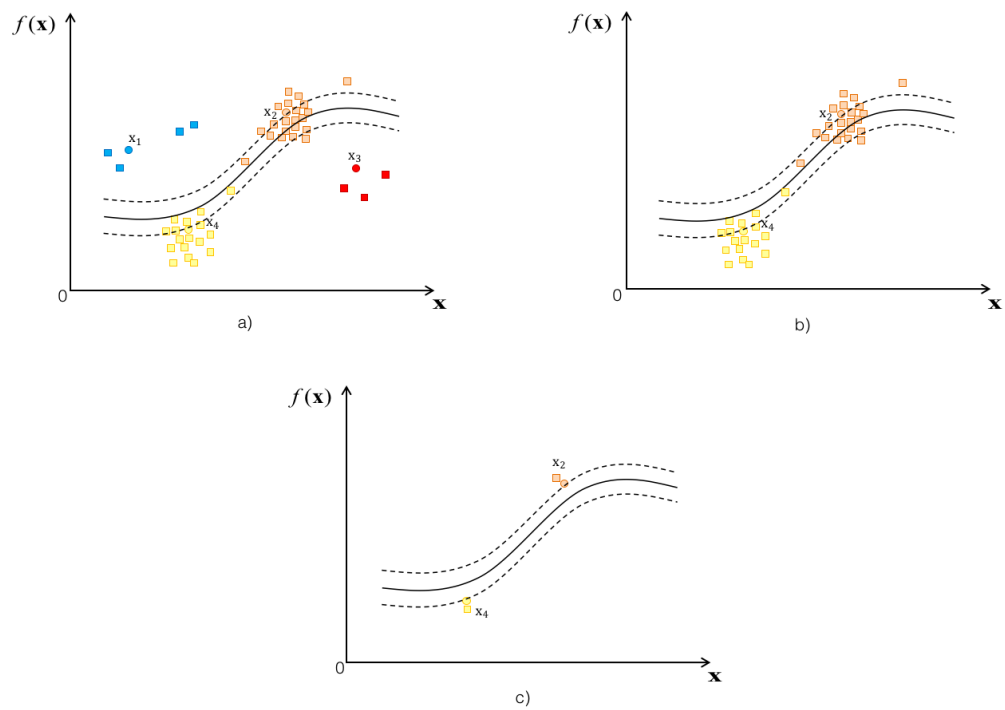


Figure 4.3: A qualitative example in a 1-D feature space. (a) the WSVR learning function obtained at the first step; (b) selection of SVs that lie on the ε -tube after elimination of samples out and far from the tube (c) selection of informative unlabeled samples having largest confidence and relevancy to be correctly estimated.

Algorithm 2: Step 2 of the proposed SSL method

- 1: **Inputs:**
 SVs (*Support vectors*)
 U (*Unlabeled samples set*)
 L (*Training samples set*)
 weights s_i of the training samples, where $i = 1, 2, \dots, n$
 - 2: **Outputs:**
Final target values
 - 3: **Begin**
 - 4: Find the closest SV associated to each unlabeled sample
 - 5: Eliminate the unlabeled samples associated to SVs located outside the ε -tube
 - 6: Compute the number of unlabeled samples to be associate to each SV, taking from each region the samples closest to the remaining SV
 - 7: Trained the re-structured WSVR by considering the labeled samples and most informative unlabeled samples
 - 8: Estimate the target values of the samples
 - 9: **End**
-

4.4 Dataset description and design of the experiments

In order to evaluate the performance of the proposed SSL method, we have conducted experiments on two different regression problems, aiming to estimate: 1) tree stem diameter from a set of features extracted by Light Detection And Ranging (LiDAR) data acquired on a forest, and 2) age of the abalones from the variables that model the physical characteristics of abalones. The first regression problem refers to a study area that is a forest site in the Southern Italian Alps. The LiDAR data located at Parco Naturale Paneveg-

Table 4.1: The list of variables used in the experiments and their physical meanings for the estimation of stem diameter.

Variable Name	Physical Meaning
1st maximum	Tree Height
mean 1st mean 3rd	Crown Internal Structure
1st minimum	Tree Height
4th range	Crown Internal Structure
1st intensity	Species

Table 4.2: The list of variables used in the experiments and their physical meanings for the estimation of abalones ages.

Variable Name	Physical Meaning
Sex	Sex of the abalone
Length [mm]	Longest shell measurement
Diameter [mm]	Measured perpendicular to the length
Height [mm]	Height of abalone with meat in the shell
Whole weight [g]	Weight of the whole abalone
Shucked weight [g]	Weight of the meat
Viscera weight [g]	Gut weight (after bleeding)
Shell weight [g]	Weight after being dried
Rings [g]	+1.5 gives the age in years

gio, Trento was acquired on July 16, 2008 by an Optech ALTM 3100EA sensor mounted onboard of a helicopter. The site is a typical Alpine forest with three main species (i.e., Norway spruce, silver fir, and European beech) and less common species (i.e., European larch and Scots pine). The LiDAR data have been interpolated and the resulting images have been segmented. The variables for describing each tree have been extracted from the segmented regions using the elevation and intensity channels of the four LiDAR returns (Dalponte et al. [2011]). Table 4.1 shows the detailed information on the variables used in the experiments and their expected physical meanings. For the Paneveggio dataset, the available ground reference samples (581 samples) were randomly divided in a test set of 229 samples, a biased labeled training set of 24 samples and an unlabeled sample set of 228 samples. It is worth noting that the training set is poor due to the small number of samples considered.

The second regression problem refers to a dataset for the estimation of the age of abalones that are included in the UCI Machine Learning Repository (Bache and Lichman [2013]). This dataset consists of 9 variables for which the detailed information together with their expected physical meanings are given in Table 4.2. The number of rings characterizes the age of the abalone, and therefore is the variable to predict. The available reference samples (4177 samples) were randomly divided to derive a test set of 2082 samples, a biased training set of 45 samples and an unlabeled sample set of 2050 samples. For the experiments, we adopted an SVR with Radial Basis Function (RBF) kernel. The values of the regularization parameters and, of the parameter ε and of the spread of the RBF kernel were obtained by a five-fold cross validation. The regularization parameter

C^L of SVR was tested between [10-1000] with a step size increment of 10. In the second step of the proposed method, the regularization parameter C^U of the WSVR (which is associated to the informative unlabeled samples) is set to $0.5 C^U$. This is done to limit the importance assigned to the unlabeled patterns for which real target values are not existing and give more importance to training samples on which we know the true target values. The ε parameter of SVR and the spread of the RBF kernel were tested between [0.01-1] and [0.01-2], respectively, with a step size of 10. The results referred to average results obtained on 10 trials associated to 10 initial randomly selected biased training set. Results are provided in terms of: i) mean and standard deviation (std) of the mean absolute error (MAE), ii) mean and standard deviation of the coefficient of determination R^2 . In the experiments, we have analyzed the effectiveness of the proposed SSL method by comparing the results with those obtained by the standard SVR method (Smola and Schölkopf [2004]). Moreover, we analyze the effectiveness of the steps of the proposed method by evaluating the performance when only the first and second steps are executed independently from each other. In addition, to evaluate the effectiveness of the proposed automatic radius selection strategy introduced in the first step of the proposed method, we compared the results with those obtained considering fix radius values. These values are computed in terms of Euclidean distance and for all the 10 trails they are fixed independently from the distribution of the unlabeled samples in the feature space. The number m of informative unlabeled samples being considered in the second step is selected equal to the number of SVs (obtained at step 1). All the experiments are implemented via MATLAB on a standard PC with Intel core i7 CPU 2.93 GHz.

4.5 Experimental results

4.5.1 Paneveggio dataset: estimation of stem diameter

The first set of experiments aims to assess the effectiveness of the proposed SS-SVR method for the estimation of the tree stem diameter from the Paneveggio dataset. Table 4.3 shows the results of the average (on 10 trials) MAE and R^2 obtained by the standard SVR and the proposed SS-SVR. From the table one can see that the proposed SS-SVR results in the best MAE and R^2 values and is also more stable (with respect to its lowest standard deviation) than the standard SVR. As an example, the SVR method reaches an MAE of 1.07, whereas the proposed method obtains an MAE of 0.88 and it is more stable with lower standard deviation over the ten trials.

Table 4.4 shows the results of the average (on 10 trials) MAE and R^2 obtained when the first step of the proposed method is applied by using: i) the proposed automatic radius selection strategy; and 2) random selection of radius. In the experiments, the

Table 4.3: Mean and standard deviation (std) of MAE and R^2 obtained by the standard SVR and proposed SS-SVR.

Method	MAE		R^2	
	mean	std	mean	std
SVR	1.07	0.15	0.53	0.1
Proposed SS-SVR	0.88	0.07	0.57	0.05

Table 4.4: Results of the first step of proposed method obtained when the value of radius is estimated either by the proposed strategy or by randomly fixing its value.

Method		MAE		R^2	
		mean	std	mean	std
Automatic Radius Estimation		1.05	0.15	0.53	0.1
Random Selection of the Radius	R = 0.03	1.18	0.22	0.50	0.1
	R = 0.05	1.13	0.20	0.50	0.1
	R = 0.07	1.27	0.26	0.50	0.1
	R = 0.08	1.30	0.28	0.49	0.1

radius is varied in range of [0.03 to 0.08] with a step size of 0.01. From the Table one can see that the proposed strategy to estimate the value of radius provides the best results when compared to the random selection strategy. From the results one can observe that a proper selection of R is important. Selecting the radius value randomly may lead to high MAE and low R^2 . For example, when R=0.08 the MAE is 1.30, while with the proposed strategy leads to a MAE of 1.05. In addition, the standard deviation of both MAE and R^2 demonstrate that the proposed step is stable. The reason is that, an arbitrary selection of the radius dimension can generate overlaps among different portions in the feature space making the density computation ambiguous and, consequently, the weighting training samples will not improve the regression function.

Table 4.5 shows the performance of the proposed method by exploiting: 1) only the first step; 2) only the second step; and 3) two steps together. By analyzing the table one can see that the proposed method with both steps provides the best MAEs and R^2 and results more stable (due to its lowest standard deviation) with respect to the cases where either first step or the second step is considered. In particular, using both steps yields a MAE of 0.88 and an R^2 of 0.57, that of first step and the second step provide MAEs of 1.05 and 1.11, and R^2 of 0.53 and 0.51, respectively. In addition, using both steps leads to more stable results with lower standard deviation over the ten trials. In details,

Table 4.5: Mean and standard deviation (std) of MAE and R^2 obtained by executing the proposed SS-SVR with and without the 2nd step.

Method		MAE		R^2	
		mean	std	mean	std
Proposed SS-SVR	Only step 1	1.05	0.15	0.53	0.1
	Only step 2	1.11	0.19	0.51	0.12
	Both steps	0.88	0.07	0.57	0.05

on the one hand if the first step is neglected (i.e., the second step is applied directly on the biased training set), the lowest performance is obtained with the highest MAE and lowest R^2 . This result demonstrates the importance of the priors’ injection phase of the proposed method. It is worth noting that if the priors are not injected in the case of biased training sets, the weights of the training samples cannot be accurately defined and thus the proposed method can fail to effectively model the regression function. Moreover, if only the first step is used, the performance is better than the case of using only the second step. This shows the importance of including the information on unlabeled samples while characterizing the regression function. The average computational time required for the first step is 20.5 seconds, while that for the second step is 1.8 seconds.

4.5.2 Abalone dataset: age of abalone

Similar experiments have been conducted on a second dataset, called Abalone, to evaluate the effectiveness of the proposed SSL method in the context of the estimation of age of abalones. Table 4.6 shows the average results (on the 10 trials) MAE and R^2 obtained by the standard SVR and the proposed SS-SVR. As one can see from the table, the best result is achieved again by the proposed SS-SVR method in terms of both MAE and R^2 . In particular, the MAE obtained with the standard SVR is 3.21, while with the proposed SS-SVR method is 2.69 and, the stability increases with a reduction of the standard deviation from both the MAE and the R^2 .

Similarly to Table 4.4 shown in the previous experiment, Table 4.7 shows the results of the average MAE and R^2 obtained on the 10 trials when the first step is applied, using: i) the proposed automatic radius selection strategy; and 2) a random selection of radius (fixed for all the 10 trials in the range [0.03 to 0.08]). From the table, one can see that the proposed automatic radius estimation strategy leads to the best results when compared to the random selection strategy. This result underlines the importance of a proper selection of the radius R . As an example, when $R = 0.08$ the MAE is 3.48, while

Table 4.6: Mean and standard deviation (std) of MAE and R^2 obtained by the standard SVR and proposed SS-SVR.

Method	MAE		R^2	
	mean	std	mean	std
SVR	3.21	0.62	0.43	0.04
Proposed SS-SVR	2.69	0.41	0.49	0.03

Table 4.7: Comparison of the results of the first step of proposed method obtained when the value of radius is estimated by either the proposed strategy or by randomly fixing its value.

Method		MAE		R^2	
		mean	std	mean	std
Automatic Radius Estimation		2.75	0.43	0.43	0.03
Random Selection of the Radius	R = 0.03	3.03	0.51	0.41	0.04
	R = 0.05	3.25	0.54	0.41	0.06
	R = 0.07	3.33	0.55	0.40	0.06
	R = 0.08	3.48	0.56	0.49	0.07

with the proposed strategy we obtain a MAE of 2.75. Moreover, the standard deviation of both MAE and R^2 demonstrate that the proposed automatic approach is more stable than the random selection strategy. This result is due to the fact that, when a large R is defined, the unlabeled samples can be part of more than one portion of the feature space while determining the density, which leads to misleading density estimation. This implies a reduction of the capability to model a reliable regression function.

Table 4.8 presents the performance obtained by applying: i) only the first step; 2) only the second step; and 3) the two steps together. As we can see in the table, the best results

Table 4.8: Mean and standard deviation (std) of MAE and R^2 obtained by executing the proposed SS-SVR with and without the 2nd step.

Method		MAE		R^2	
		mean	std	mean	std
Proposed SS-SVR	Only step 1	3.21	0.62	0.43	0.04
	Only step 2	2.75	0.43	0.43	0.03
	Both steps	2.69	0.41	0.49	0.03

are obtained by exploiting together step 1 and step 2 of the proposed method. The use of both steps leads to a MAE of 2.69, while using separately the first and the second, MAEs of 2.75 and 3.11, are obtained respectively. Also the stability of the result is increased, by jointly using both steps a lower standard deviation over the ten trials. In accordance to the performance obtained on the Paneveggio dataset, this result confirms the importance of exploiting the prior information in the learning phase when the training set is biased. In addition, this confirms that the exploitation of highly informative and highly relevant unlabeled samples together with the weighted biased training samples, can further improve the characterization of the regression function. The average computational time required for the first step is 40.1, while that for the second step is 2.8 seconds. It is worth noting that, the number of the both the labeled and the unlabeled samples in the dataset affects the computational time.

4.6 Conclusion

In this Chapter, we have presented a novel SSL method for the estimation of biophysical parameters from remotely sensed images in the framework of ε -insensitive support vector regression (SVR). The proposed method is made up two steps. The first step aims at injecting additional priors information in the initial learning of the SVR function in order to avoid the problems occurred due to the availability of the small-sized biased training sets. This is achieved by assigning different weights to training samples that are obtained by a novel strategy that relies on the density of samples in the feature space. According to the developed strategy higher weights are given to the training samples located in the higher density locations of the feature space, and vice versa. Then, a weighted SVR algorithm is exploited in order to adopt different weights to training samples. Using the prior information associated to the density, without including additional labels, allows one to better model the SVR learning function when compared to the standard ε -insensitive SVR. It worth noting that when the weight is 1, the result obtain is equal to the one obtained with the standard SVR algorithm. Thanks to the first step, the proposed method mitigates problems of small-sized biased sets of training samples for a proper modeling of the SVR learning function. The second step aims at jointly using labeled samples and a set of informative unlabeled samples in order to better characterize the SVR learning function. This is achieved by introducing a new term in the WSVR presented in the first step. In order to select the informative unlabeled samples that are located as closest as possible to the ε -tube and outside of the ε -tube a novel strategy is presented. Thanks to the second step of the proposed method, the performance of the parameter estimation is further improved. In the experiments, we studied the effectiveness of the proposed

SSL method for the estimation of: 1) tree parameters from LiDAR data and 2) ages of abalones. We compare the proposed method with standard SVR. Experimental results show that the proposed strategy allows one to improve the estimation performance with zero cost of the collection of reference samples. As future development, we plan to test the effectiveness of the proposed SSL method on other estimation problems (e.g. the retrieval of SMC from satellite data).

Chapter 5

A novel technique based on interferometry for clutter detection in two-channel radar sounder data

This Chapter provides the third contribution of the thesis in the context of clutter detection from RS data. We address the clutter issue by proposing a novel technique based on interferometry for clutter detection in two-channel radar sounder data. The proposed method includes three main steps: i) the manual feature extraction and theoretical phase difference estimation; ii) the RS interferogram formation; and iii) the comparison of theoretical and real phase difference distributions. Experimental results on RS data acquired in Greenland validate the proposed approach for surface clutter detection using two-channel RS data.

5.1 Introduction

In the Earth observation, radargrams allow the direct investigation of the bed topography and morphology, making available to glaciologist fundamental information for understanding the dynamics of terrestrial ice sheets and glaciers (Schroeder et al. [2014]). Several airborne RS instruments are currently operating in Antarctica and in Greenland (Gogineni et al. [2014], Peters et al. [2007], Dall et al. [2010a]). RSs have the unique capability to actively probe the subsurface of planetary bodies to many kilometers of depth (depending on the electromagnetic characteristics of the sounded terrain) from orbiting satellites. A fundamental research goal for RS observations of planetary ice masses and icy moons is the detection of liquid subsurface water. Two RSs currently operating for planetary exploration at Mars are MARSIS (Jordan et al. [2009]) and SHARAD (Crocini et al. [2011]). The science return from MARSIS and SHARAD observations have increased the interest

in sounder payloads for future missions. Two new instrument concepts for the study of the icy moons of Jupiter are the Radar for Icy Moons Exploration (RIME) (Bruzzone et al. [2015]), which will be carried on board of the European Space Agency (ESA) mission named Jupiter Icy Moons Explorer (JUICE), and the Radar for Europa Assessment and Sounding: Ocean to Near-surface (REASON), recently selected by National Aeronautics and Space Administration (NASA) for a mission to Europa (Grima et al. [2015]). The nadir acquisition geometry and wide beam patterns of RS systems can inhibit the interpretation of subsurface signals obscured by or confused with returns from lateral surface clutter. This clutter signal can prevent the definitive interpretation of subsurface features when lateral and nadir echoes are received with the same delay. This interference strictly depends on the surface topography and scattering, most significantly affecting the shallow regions of the subsurface due to the directivity of the beam pattern. Because of the RS acquisition geometry (see Figure 5.1), the acquired data can be affected by both along-track and cross-track clutter.

Different solutions to deal with the clutter problem have been presented in the literature. For along-track clutter, Doppler filtering and SAR processing are well-known techniques for clutter suppression and they are implemented as the baseline in most RS systems (Heliere et al. [2007]). The main approaches to deal with cross track clutter, presented also in Section 2.3 of this thesis, can be summarized in the two following approaches: i) clutter suppression by designing multi-channel RSs; and ii) clutter detection using precise information about the surface topography. In the first class of approaches, the cross-track clutter can be suppressed by using sophisticated antennas with very directive beams or by exploiting the cross-channel phase of multichannel systems. A multichannel system allows one to estimate the direction of arrival of the received signal (Gogineni et al. [2014], Jezek et al. [2011b], Dall et al. [2010a]) and to produce swath measurements by exploiting multiple phase centers (Wu et al. [2011]). A major drawback of this approach is that it requires high complexity in hardware system design and large data volume resources to independently record the echoes received from the channels. These problems are particularly relevant in the design of planetary RSs for which resources (e.g., mass, volume, power and data volume allocation) are very limited compared to airborne RSs. Concerning multichannel systems, MARSIS envisaged the use of a second antenna, which was a monopole oriented along the nadir axis aiming to receive the off-nadir surface returns and then subtract the clutter returns by the primary antenna composite signal. However, this method requires additional antenna pattern calibration which made the technique difficult to apply Picardi et al. [2003]. The second class of approaches involves algorithms for the clutter surface simulation that use as an input a DEM of the surface (Ferro et al. [2013], Nouvel et al. [2004], Iorio et al. [2008], Russo

et al. [2008]). Different simulation algorithms are presented in the literature to develop the cross-track mapping of the radar echoes onto the surface. These algorithms allow one to manually or automatically compare a radargram with a surface clutter simulation and to map the location and shape of the clutter echoes in the radargram. However, in some cases DEMs, which have sufficient spatial resolution to properly represent features in the radargram, are not available. Moreover, these techniques can only identify surface echoes, without providing discrimination of clutter generated from a structure buried in the ice subsurface.

In this Chapter, we address these problems by proposing a novel technique for the discrimination of surface clutter in radar sounder data by means of the analysis of two-channel RS system data. The proposed technique computes the statistical distribution of the cross-channel phase difference for both cross-track surface clutter and subsurface features at nadir. We apply an interferogram formation technique to extract the measured phase difference between the two channels. The technique proposed in this Chapter exploits a single pass interferometric approach with a cross-track baseline to discriminate the surface clutter from nadir subsurface returns. Interferometry has been widely applied in InSAR (Interferometric Synthetic Aperture Radar) (Rosen et al. [2000]). InSAR employs observations of the same area using two SAR complex signals acquired from different observation positions. InSAR is widely used in remote sensing for geophysical monitoring of natural hazards (e.g., earthquakes modeling in Massonnet et al. [1994]) or topographic map generation (Rosen et al. [2000]). Until now, interferometry was only used in a limited capacity with radar sounding data (Jezek et al. [2011a]) and never for clutter discrimination using a two-channel system. Although the cross-channel phase difference signal has been exploited for clutter detection and suppression few multi-channel radar systems, we propose a technique that limits the hardware complexity to two channels for the discrimination of surface clutter. Limiting the hardware complexity and thus the consumption of the most relevant resources for satellites sensors (e.g., mass, power, data rate, data volume), the proposed technique aims to provide a proof of concept for a planetary radar sounding approach being able to discriminate the surface clutter features in RS data. This technique is a candidate to be considered for the implementation of the REASON RS for Europa mission.

We demonstrate the proposed method through application to airborne sounding data acquired by the University of Texas Institute for Geophysics (UTIG) Multifrequency Airborne Radar Sounder with Full-phase Assessment (MARFA) radar in northern Greenland. However, the presented processing chain is generic and can also work with other RS sensors having two-channels. The proposed technique is robust to roll motion and allows the generation of phase distribution histograms, radargram interferograms, and interferomet-

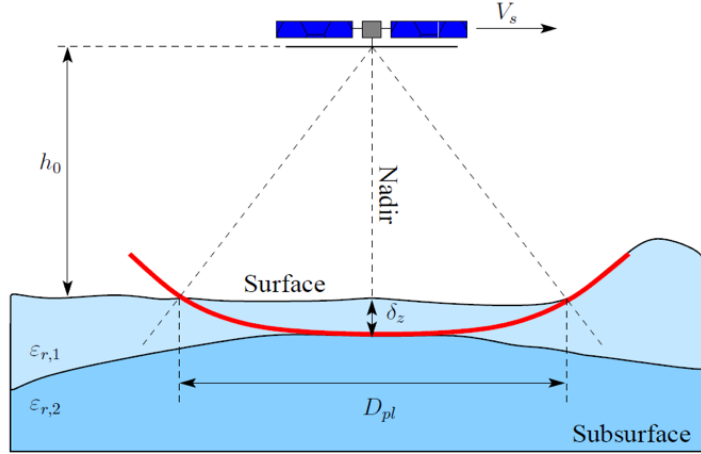


Figure 5.1: 2-D representation of the acquisition geometry of a RS instrument.

ric correlation maps that can be used for clutter detection.

The remaining Sections of this Chapter are organized as follows. Section 5.2 introduces the proposed clutter discrimination method. Section 5.3 presents the dataset used in the experiments, the main characteristics of the MARFA two-channel radar system and the results of the experiments. Then, Section 5.4 is devoted to discuss the use of the proposed approach on planetary RS data. Finally, Section 5.5 draws the conclusion of this work.

5.2 Proposed technique for cross track clutter detection

We assume to have a two-channel RS, which acquires data by two antennas with a cross-track baseline B . The two-channel acquisition results in two radargrams, which are 2-D complex matrices with n and m pixels in the along-track and range directions, respectively. The proposed technique exploits both the amplitude and the phase of the two channels and computes the cross-channel phase difference. The block scheme of the proposed technique is shown in Figure 5.2 and consists of three main steps: 1) feature extraction and theoretical phase difference estimation; 2) interferogram formation for the exploitation of the cross-channel phase difference; 3) comparison of theoretical and real phase difference distributions. Both the amplitude M , and the phase P , of the received signals are given as inputs to the technique. Step 1 uses the amplitude of the signal received by each channel and allows us to model the theoretical phase difference for two different hypotheses (nadir subsurface signal echo versus surface clutter echo). Step 2 generates an interferogram by standard interferometric processing. A unique feature of the RS interferogram is that it has the same nadir geometry as the RS data. This allows, in Step 3, the amplitude radargram and the interferometric data to be directly compared by matching the position of a

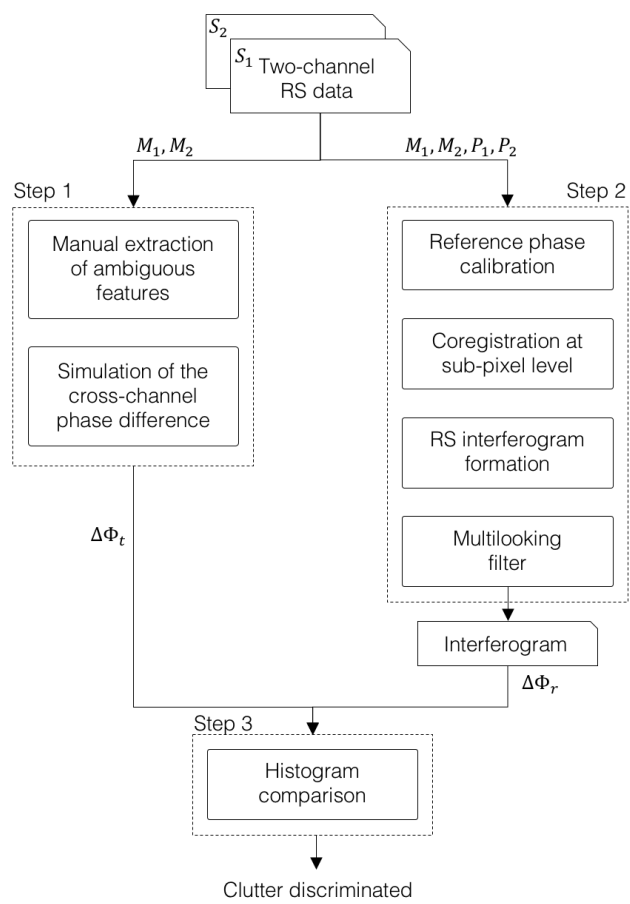


Figure 5.2: Block diagram of the proposed technique.

region of interest (containing an ambiguous feature) with the phase difference estimated from the two-channel RS data.

5.2.1 Manual feature extraction and theoretical phase difference estimation

The goal of this first step is to model the theoretical phase difference between signals received by the two RS channels. In preprocessed radargrams Peters et al. [2007], n features f_i , with $i = 1, 2, \dots, n$, of ambiguous origin (either nadir subsurface or off-nadir cross-track surface returns) can be manually selected. The feature extraction procedure selects the position of high-power pixels that belong to the feature region (see Figure 5.3). The delay due to these features with respect to the surface return can be extracted from the radargram, allowing the projection of the features onto the surface in the cross-track direction and the computation of the theoretical cross-channel phase difference Δ_ϕ for the feature f_i . The proposed simulation approach aims to estimate Δ_{ϕ_t} (i.e., theoretical cross-channel phase difference) according to the geometry shown in Figure 5.4, where a platform (e.g., airborne) mounts a RS sensor connected to two antennas collocated in the across-track direction and separated by a baseline B . Each antenna receives independently the echo coming from different locations on the terrain according to the surface geometry and the beamwidth of the antenna pattern. The expected phase difference can be computed

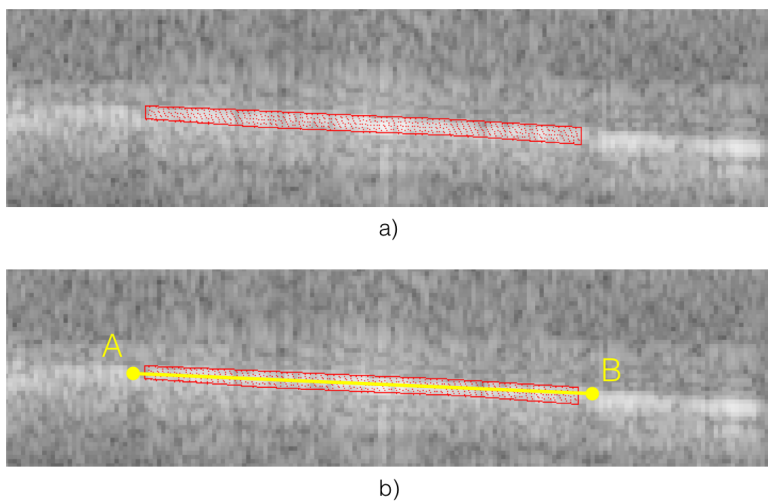


Figure 5.3: Example of manual feature selection. a) The red area corresponds to the set of pixels that belong to a feature of interest. b) The yellow line is the simple feature-model used to project the feature onto the surface in the cross-track direction.

as:

$$\Delta_\phi = \frac{2\pi}{\lambda}(\rho_{chr} - \rho_{chl}) = \frac{2\pi}{\lambda}\rho_{chl} \left(\sqrt{1 - \frac{2\langle \vec{b}, \hat{l}_1 \rangle}{\rho_{chl}} + \left(\frac{B}{\rho_{chl}}\right)^2} - 1 \right) \quad (5.1)$$

where ρ_{chl} and ρ_{chr} are the path lengths of the received signals for the left and the right channel respectively, and λ is the wavelength, $B = |\vec{b}|$ is the baseline vector, \hat{l}_1 is the unit line-of-sight vector from left to right. With the assumption of a flat surface and correction for aircraft roll, we compute the echo received from each antenna as the sum of two contributions (see Figure 5.2):

$$\rho_{chl} = H_j + D_j, \rho_{chr} = H_k + D_k \quad (5.2)$$

where H_j and D_j are the signal propagation paths in free space and into the subsurface (see Figure 5.4), received from the left channel, whereas H_k and D_k are the signal propagation paths in free space and into the subsurface recorded from the right channel. The collocation of P_j on the surface can be solved using Snell's law (Peters et al. [2007]), i.e.,

$$\sin \Theta_j = \sqrt{\epsilon_r} \cos \Theta'_j \quad (5.3)$$

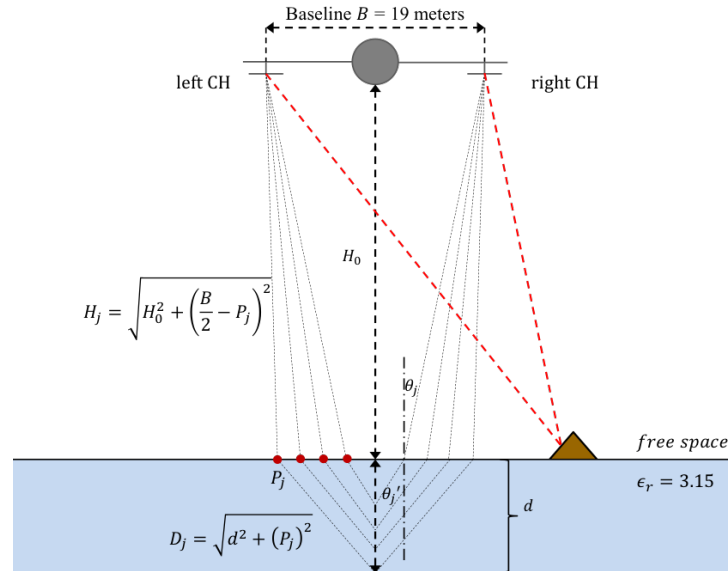


Figure 5.4: Two-channel sounding geometry. The red dots are reflection points on the surface. For echoes coming from nadir, the signal path length ρ_{chl} and ρ_{chr} are equivalent and $\Delta_\phi = 0$. The red dotted line represents an example of off-nadir signal path (for a clutter feature with the same delay), in which the difference between ρ_{chl} and ρ_{chr} makes $\Delta_\phi \neq 0$.

which can be written as

$$\frac{\frac{B}{2} - P_j}{H_j} = \sqrt{\epsilon_r} \left(\frac{P_j}{D_j} \right) \quad (5.4)$$

corresponding to the following fourth-order equation

$$(\epsilon_r - 1)P^4 - 2(\epsilon_r - 1)XP^3 + [(\epsilon_r - 1)X^2 + \epsilon_r H_0^2 - d^2]P^2 + 2d^2XP - d^2X^2 = 0 \quad (5.5)$$

which can be solved using Ferrari formula for quartic equations (Turnbull [1947]). This algorithm is more efficient than computing the eigenvalues of the companion for polynomials solution. From the four obtained solutions, we select the one that is compatible to the following geometrical constraints (as done in Heliere et al. [2007]):

$$X - P \geq 0 \quad \text{for } X \geq 0 \quad \text{or} \quad X - P \leq 0 \quad \text{for } X \leq 0 \quad \text{and} \quad P \in \mathfrak{R} \quad (5.6)$$

The same procedure can be iteratively applied to every depth and cross-track position. We compute this theoretical phase difference for each reflection point P_j that belongs to a feature for which we aim to determine its origin as either surface clutter or subsurface signal at nadir. In order to generate a realistic model for the phase distribution of the feature, we also model the interferometric phase noise. From interferometry theory, we compute the thermal noise correlation, γ , (Rodriguez and Martin [1992]) as

$$\gamma = \frac{1}{1 + SNR^{-1}} \quad (5.7)$$

where SNR is the Signal-to-Noise Ratio of the received signal, assuming identical signal and noise power for the two channels. Note that an increase in thermal noise decorrelation (i.e., $\delta = 1 - \gamma$) results in a higher phase noise variance, which is given from the Cramer-Rao bound Rodriguez and Martin [1992] as

$$\sigma_0 = \frac{1}{\sqrt{2N_l}} \frac{\sqrt{1 - \gamma^2}}{\gamma} \quad (5.8)$$

where N_l is the number of looks. This indicates that the number of looks strongly affects phase noise variance for small correlation values.

5.2.2 Interferogram formation for cross-channel phase difference

In order to combine the two channels and analyze their difference in phase, we use an interferometric approach to generate an interferogram with the same geometry as the radargram. The processing chain for generating the interferogram is made up of the following step: i) reference phase calibration, ii) co-registration at sub-pixel resolution, iii) interferogram formation, iv) multilooking. In addition to this, we applied a roll compensation approach. In the following, we explain in further detail each step of the processing

chain. The first step aims to correct the phase offset between the two channels due to the radar system itself. This deviation is commonly due to the length difference of the cables that connect antennas and receivers. This physical length difference affects the path length of the each received signal and, eventually, the phase difference between the two channels. For assessing this deviation, an internal reference signal for each channel can be recorded. The travel time of this calibration signal is not affected by external reflection. By comparing the two references, we can measure the time delay of the two received signals and, therefore, compute any internal phase deviation between R Ch (Right Channel) and L Ch (Left Channel). To do this, we independently apply the range compression to the two channels and for each range line we compute the difference in phase of the two channels. Then, this deviation is compensated as

$$C_R = C_R^* e^{-i\Delta\Phi_c} \quad (5.9)$$

where C_R^* is the complex radargram acquired by the R Ch and $\Delta\Phi_c$ is the phase difference between the two channels evaluated on the reference signal. It worth noting that this calibration system is suitable for assessing phase stability both during a single flight and over an entire acquisition campaign. On a single flight, the phase can be affected by fluctuations due to temperature variations. On an entire acquisition campaign, the reference sample is fundamental to correct for potential configuration changes in the radar system. Sub-pixel co-registration is required to generate interferometric radargrams. The purpose of co-registration is to align the cross-channel signals as precisely as possible. It is widely accepted in the literature that, in order to have a coregistration reliable for interferometry, a 1/10 pixel precision factor should be considered (Lin et al. [1992]). To achieve this, the two complex radargram C_R and C_L are resampled with an accuracy of a tenth of a pixel using a sinc interpolation kernel (Hanssen and Bamler [1999]). The two resampled radargrams are co-registered by maximizing the correlation estimator ρ_c , as:

$$\rho_c = \frac{E[C_R C_L^*]}{\sqrt{E[|C_R|^2] E[|C_L|^2]}} \quad (5.10)$$

To generate the interferogram, C_R is multiplied by the complex conjugation of C_L , as follows

$$\Psi = Arg(C_R C_L^*) \quad (5.11)$$

Finally, multilooking is applied to the interferogram in order to reduce the speckle noise

$$\Psi_m = Arg \left[\frac{1}{nm} \sum_{k=1}^n \sum_{h=1}^m S_1(k, h) S_1^*(k, h) \right] \quad (5.12)$$

We choose a rectangular windowing for multilooking (i.e., 2 by 6 pixels in vertical and horizontal direction, respectively), which takes into account the difference in range vs. along track resolution and feature aspect ratio. Concerning the roll motion correction, the proposed method takes advantage of the telemetry information that can be obtained from the platform. For example, sensors mounted on aircraft wings can measure the roll motion and the flight altitude. We mitigate the roll motion effect by exploiting a simple geometrical model of the radar acquisition that allows one to compute the correction angle. In the case of earth observation from aircraft platform, the roll motion correction algorithm can be calibrated on data acquired on the ocean surface being the water a completely flat reflector. The correction angle precision is equal to 1 degree.

5.2.3 Comparison of theoretical and real phase difference distributions

In the last step of the proposed technique, the distributions of the observed phase difference and the theoretical phase difference (for both nadir signal and cross-track clutter) are compared. The frequency histogram of the real phase difference $\Delta\Phi_r$ is extracted from the interferogram. The theoretical phase differences $\Delta\Phi_t$ are calculated in order to test two

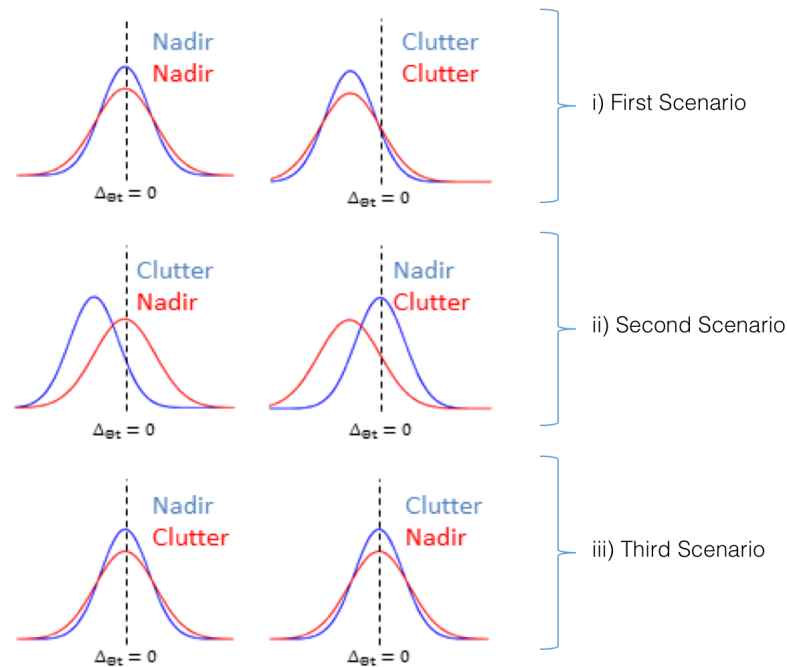


Figure 5.5: Qualitative illustration of the comparison between theoretical and real phase difference distributions. The three different scenarios are shown. In red and in blue the real and theoretical phase distributions respectively are presented.

hypotheses for the considered the feature: cross-track surface clutter vs. nadir subsurface signal. Therefore, two phase histograms associated with two features f_1 and f_2 are computed (i.e., $\Delta\Phi_{t1}$ and $\Delta\Phi_{t2}$). The simple Sum of Absolute Difference (SAD) is computed for each histogram pair ($[\Delta\Phi_r, \Delta\Phi_{t1}]$ and $[\Delta\Phi_r, \Delta\Phi_{t2}]$) to determine which theoretical distribution is a better fit for the observed distribution. SAD is the $L1$ – norm computed for a difference between two vectors and it varies between 0 and 2 Cha and Srihari [2002]. Moreover, other similarity measures could be considered. Our approach identifies three possible scenarios (see Figure 5.5): i) the observed phase distribution matches one of the theoretical distributions (either for clutter or subsurface), ii) the observed distribution matches none of the theoretical distributions (thus suggesting a subsurface off-nadir return); and iii) the observed distribution matches both the clutter and subsurface distributions (requiring ancillary data to resolve the ambiguity). In scenarios i) and ii), we are able to uniquely identify if the investigated feature is clutter or a subsurface return, because the off-nadir echo has a phase difference distribution with mean different from zero. On the other end, scenario iii) represents the case in which even if an echo comes from an off-nadir region both the real and the simulated phase differences are equal to zero, requiring additional information to discriminate a nadir feature from clutter. Scenario iii) is possible when the baseline is very large and/or when the flight altitude is low. However, when the proposed approach is applied to RS acquired from satellite, only scenarios i) and ii) are possible.

Two limitations of the proposed approach need to be discussed. i) The first one is related to the discrimination of overlapped features (i.e., when the subsurface and off-nadir returns are totally overlapped). Note that, this issue is also critical for every single-channel RS system. The proposed method could solve this problem by tracking longer the features in the along-track direction of the radargram. This is a reasonable solution for scientific interpretation of the radargram, because features of interest typically extend beyond the point of overlap with clutter contributions. ii) The second problem is the clutter discrimination ambiguity due to low flight altitude acquisitions (i.e., phase difference equal to zero for a clutter feature). As discussed above (scenario iii)), this issue occurs only for airborne observations, where the relatively low flight altitude makes the phased difference wrapping faster than in the satellite case (see simulated results presented and discussed in Section 5.4).

5.3 Dataset description and experimental results

The data used in this study were collected by UTIG in the Spring of 2014 over Bowdoin Glacier (Greenland), during a collaborative project with the Scott Polar Research

Institute at Cambridge University. The MARFA system used is an updated version of HiCARS radar (Peters et al. [2007]). The collected tracks offer the opportunity to analyze different kinds of subsurface and clutter features (see Figure 5.6). The first track, NAQLK/JKB2j/ZY1b, is collected in a region of Greenland far from sources of surface clutter (e.g. nunataks, crevasse fields). Tracks GOG/JKB2j/BWN01b and GOG/JKB2j/BWN01a are acquired in a region where the subsurface probed by the RS is surrounded by nunataks (i.e., an exposed peak, not covered with ice or snow), making the interpretation (using single-channel RS data) challenging or impossible due to the strong presence of clutter.

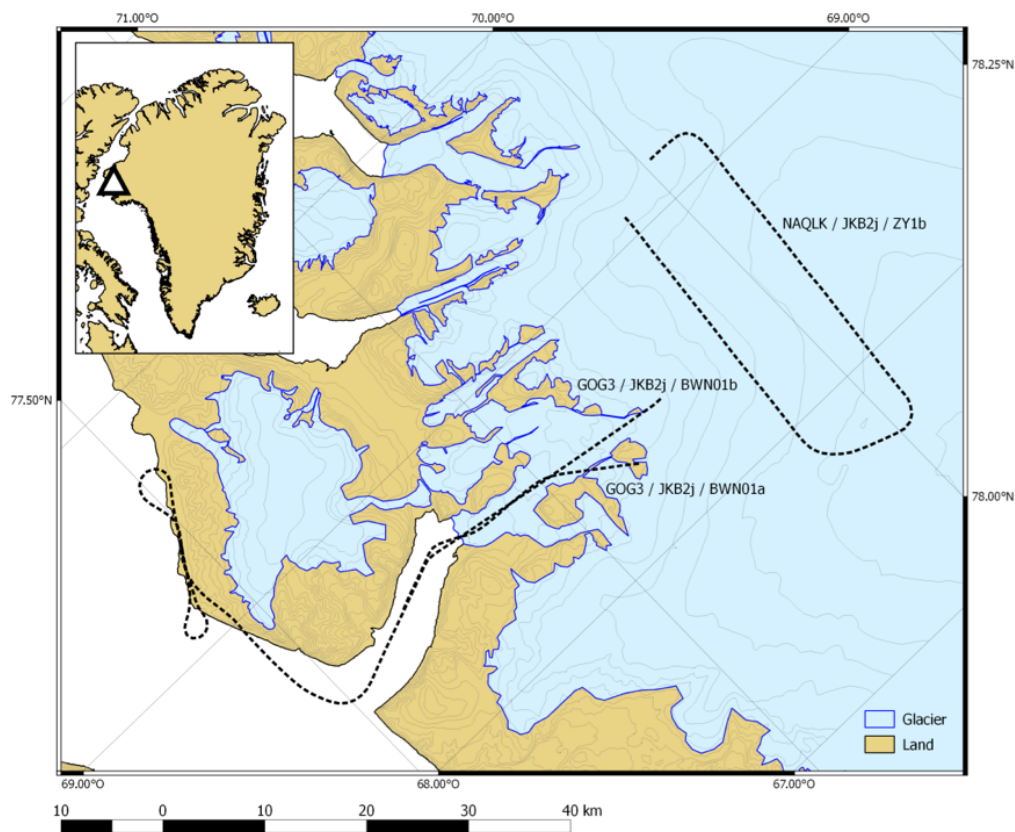


Figure 5.6: Location of the RS tracks collected by MARFA on the Greenland ice shell.

5.3.1 MARFA radar system description

The High Capability Airborne Radar Sounder (HiCARS) is an ice-penetrating radar operated by the UTIG Peters et al. [2007]. This instrument allows researchers to build profiles of ice sheets that show the ice surface, internal layers, and the bedrock in Antarctica and

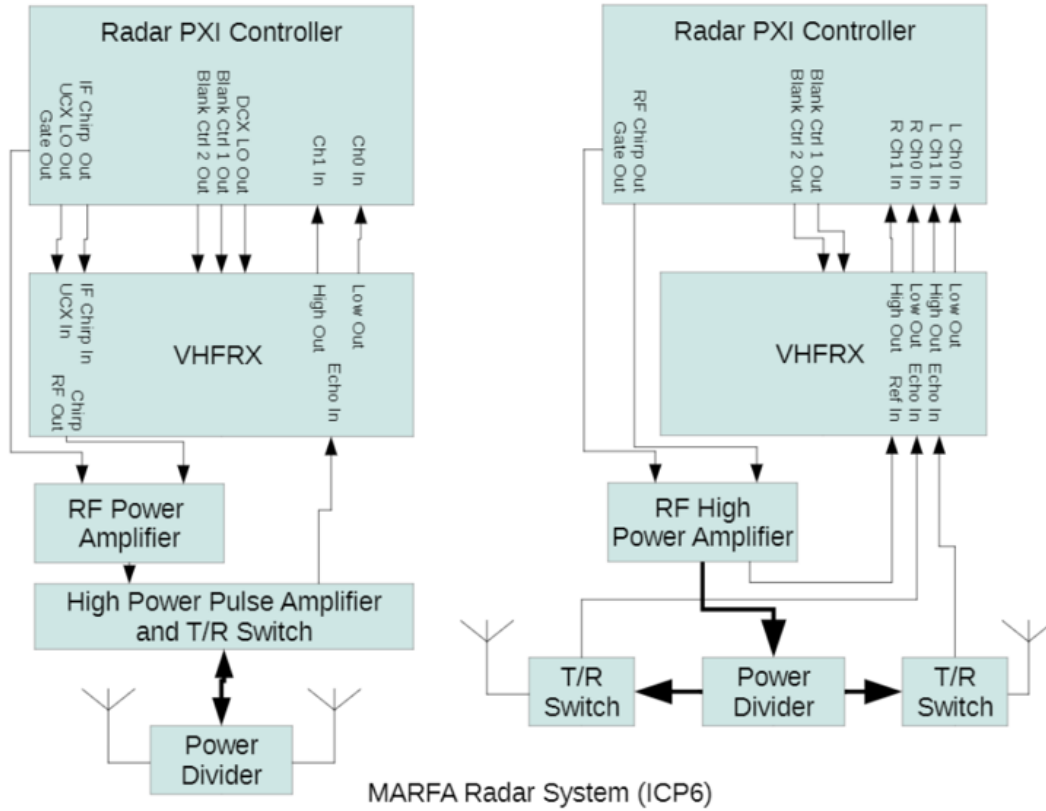


Figure 5.7: Block scheme of the MARFA radar system.

Greenland. HiCARS has been configured with an array of two dipole antennas installed on the wings of a DC-3 aircraft. The radar system has been recently updated by UTIG in

Table 5.1: Typical operating parameters for MARFA RS.

Main System Parameters	Parameter Values
Operating Frequency	52.5-67.5 MHz
Bandwidth	15 MHz
Wavelength	~5 meters
Number of Antenna Elements	2 (TX/RX)
Baseline	19 meters
Sampling Frequency	50 MHz
Pulse Repetition Frequency	6250 Hz
Pre-summing Factor	32

order to independently record the echoes received by each antenna. Figure 5.7 shows the updated configuration called MARFA. Two transmit/receive switches were used instead of one, the low noise amplifier system in VHF Receiver (VHFRX) was duplicated, and another digitizer module was added, allowing the echoes from each antenna to be recorded separately. To monitor the reference phase, during pulse transmission, a sample of the generated RF chirp was obtained from the RF power amplifier stage and multiplexed into the received radar trace. The system configuration used in most recent campaigns is shown in Table 5.1.

5.3.2 Dataset 1: deep subsurface

The first RS track is 50 line-km long and the resolutions in the along-track and range direction are 15 meters and 5 (in pure ice) meters, respectively.

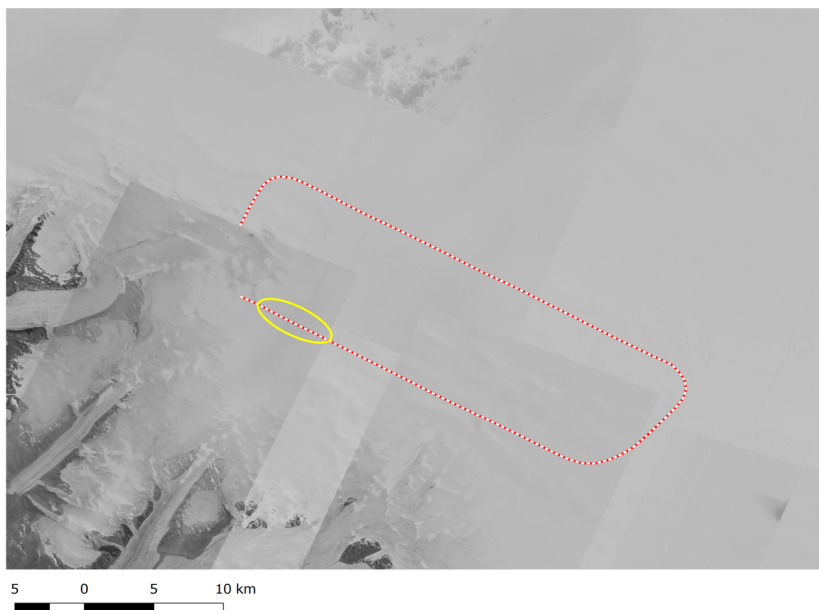


Figure 5.8: Deep subsurface dataset (track NAQLK/JKB2j/ZY1b); optical imagery acquired on Greenland (DigitalGlobe).

Figure 5.8 shows the NAQLK/JKB2j/ZY1b radar track acquired on Greenland. This track was acquired over the ice sheet interior, where we do not expect the presence of surface clutter. This line provides an initial demonstration of the capability to distinguish a subsurface nadir return without any prior knowledge about the surface characteristics of the terrain probed. Analyzing the radargram in Figure 5.9a alone would not enable determination of whether the return below the surface is clutter or a nadir subsurface return (e.g., bedrock). By contrast, Figure 5.9b shows that the echo can be identified as nadir

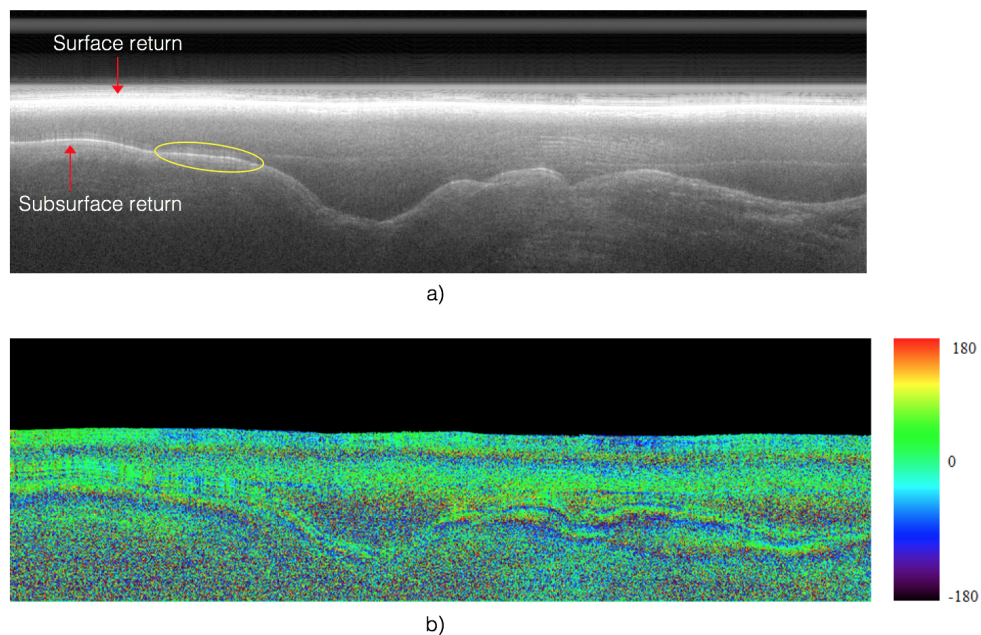


Figure 5.9: Deep subsurface dataset (track NAQLK/JKB2j/ZY1b): (a) Amplitude radargram obtained from a single channel (left) of MARFA radar (11.5 km long portion of the entire acquisition), (b) RS interferogram obtained with the proposed method.

subsurface return, because it has a cross-channel phase difference equal to zero. Following the proposed method, a feature was manually selected from the amplitude radargram and hypothesis (nadir return and clutter) histograms were simulated and compared to the observed phase histogram. Figure 5.10 shows these three phase histograms. They show that the phase difference distribution under nadir return hypothesis is more similar (particularly in terms of the average of the distribution) to the real phase difference for the selected feature. This result is also supported by the metrics reported in Table 5.2, where the SAD between real and simulated distributions is close to zero for the nadir return case (i.e. 0.29), and the mean of the distribution is also close to zero (i.e. -1.3). It is important to note that the expected value for the mean is zero when the selected feature is a subsurface feature. This result confirms the hypothesis that the selected feature is the bedrock of the icy region analyzed. An additional piece of supporting evidence can be obtained from the optical image (DigitalGlobe) in Figure 5.8, which shows that there are no off-nadir surface features at the same delay that MARFA could have acquired along the track ZY1b. The bedrock elevation map computed in Bamber et al. [2013] can also validate the result obtained with the proposed method. Although the resolution of this data is limited (i.e., the uncertainties in bed elevation exceeding ± 100 m), the results

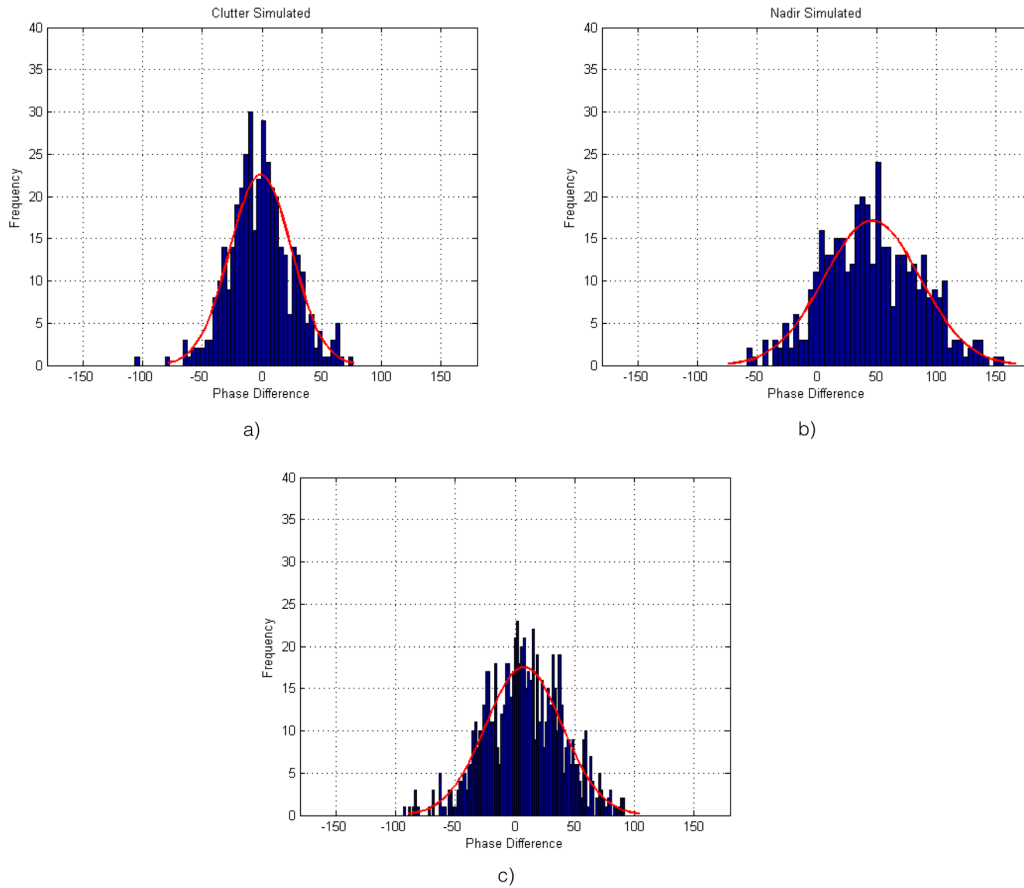


Figure 5.10: Feature shown in Figure 5.9a: (a) theoretical phase difference histogram for nadir hypothesis, (b) clutter hypothesis and (c) real phase difference histogram extracted from the real interferogram.

presented in this Section match the thickness map. This confirms the capacity of the proposed approach to identify a deep subsurface signal at nadir discriminating it from a clutter feature.

Table 5.2: SAD obtained between the real and the simulated (on both nadir and clutter hypothesis) histograms and mean of the simulated distributions.

Feature Type	SAD	mean
HP1: nadir	0.29	-1.30
HP2: clutter	0.74	43.73

5.3.3 Dataset 2: Clutter feature detection

The second RS track considered is 120 line-km long and the resolutions in the along-track and range direction are 15 meters and 5 (in pure ice) meters respectively. In



Figure 5.11: Clutter feature detection dataset (track GOG/JKB2j/BWN01b); optical image acquired on Greenland (DigitalGlobe).

track BWN01b, due to the challenging acquisition plan designed for this data, both the radargram and the interferogram are very complex. We focus our attention on the feature shown in Figure 5.12a and 5.13. After applying the proposed method, we obtained the histograms of this feature, which can be analyzed in details in Figure 5.14 and Table 5.3. Among the simulated histograms, Figure 5.14b (i.e., clutter hypothesis) has higher similarity to the observed phase difference histogram (Figure 5.14c). The SAD metric also supports the clutter-generated hypothesis. As shown in Table 5.3, the SAD value for the clutter hypothesis is smaller than the one for the nadir hypothesis. Moreover, the mean value of the real phase difference distribution matches only one of the two simulations (i.e. clutter). Figure 5.11 confirms that the proposed method is able to correctly identify the selected feature as surface clutter: both the range between the aircraft and the feature projected on the terrain and the shape of the feature itself correspond well between the optical data and the amplitude radargram. The independent detection of clutter is a relevant feature that demonstrates the ability of the proposed method to prevent the

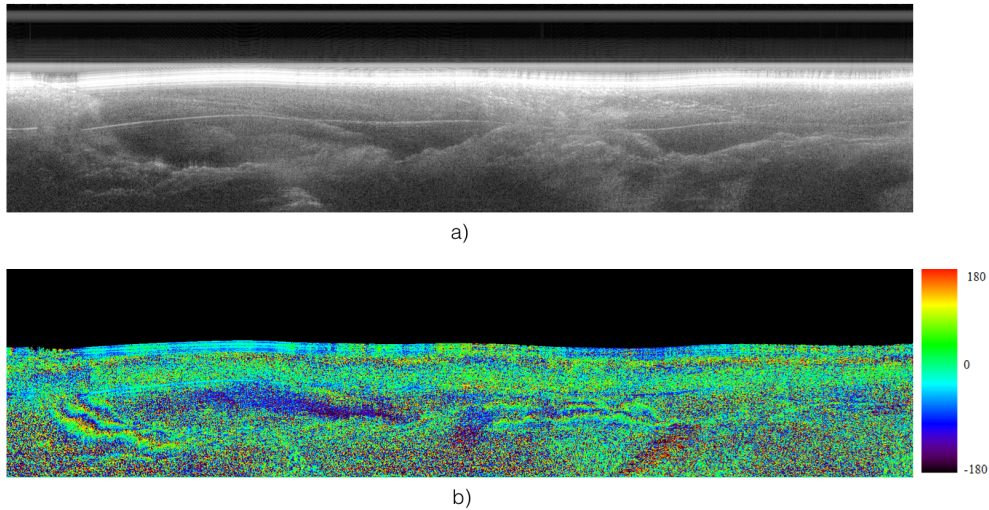


Figure 5.12: Clutter feature detection dataset (track GOG/JKB2j/BWN01b); (a) amplitude radargram obtained from a single channel (left) of MARFA radar (10 km long portion of the entire acquisition), (b) RS interferogram obtained with the proposed method.

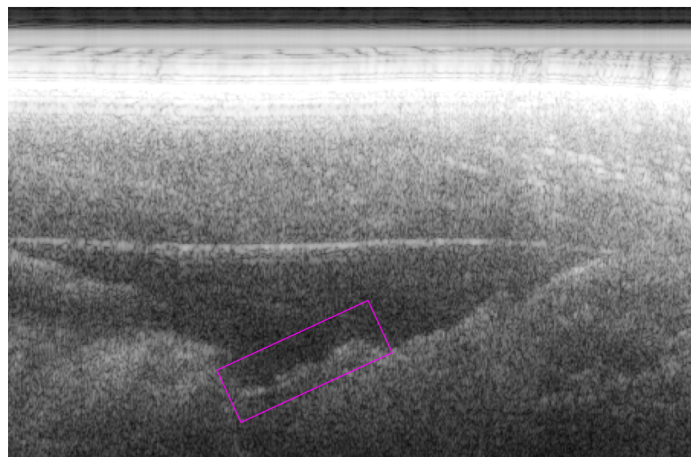


Figure 5.13: Clutter feature detection dataset (track GOG/JKB2j/BWN01b); zoomed portion of the radargram containing the selected feature.

misinterpretation of off-nadir features, by modeling and matching the phase difference of a two-channel RS system.

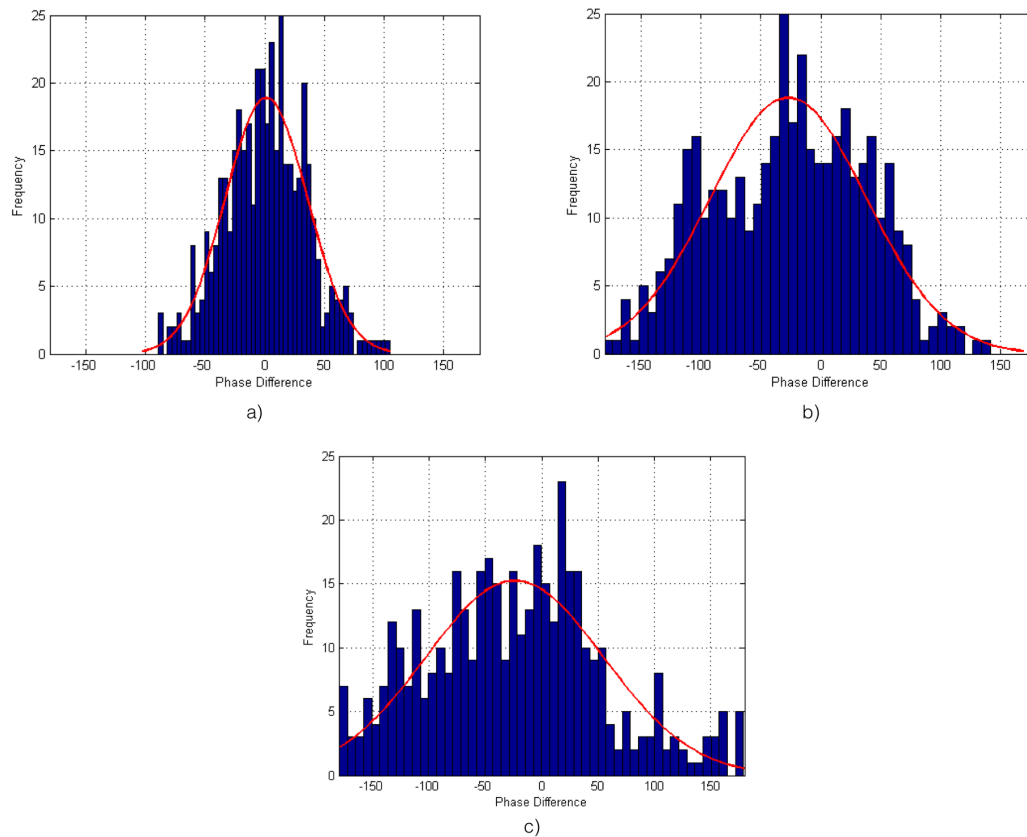


Figure 5.14: Feature shown in Figure 5.11: (a) theoretical phase difference histogram for nadir hypothesis, (b) clutter hypothesis and (c) real phase difference histogram extracted from the interferogram.

Table 5.3: SAD obtained on real and simulated (on both nadir and clutter hypothesis) histograms and mean of the simulated distributions for the selected feature.

Feature Type	SAD	mean
HP1: nadir	0.87	1.30
HP2: clutter	0.34	-27.03

5.3.4 Dataset 3: Clutter and subsurface discrimination

The last RS track is 78 line-km long and the resolutions in the along-track and range direction are 15 meters and 5 (in pure ice) meters, respectively.

From the histograms of the phase difference (Figure 5.18) one can determine that f_1 (see Figure 5.18a and Figure 5.18e) is a clutter feature. This is confirmed by projecting f_1 (according to its delay) in the across track direction onto the surface. The position of

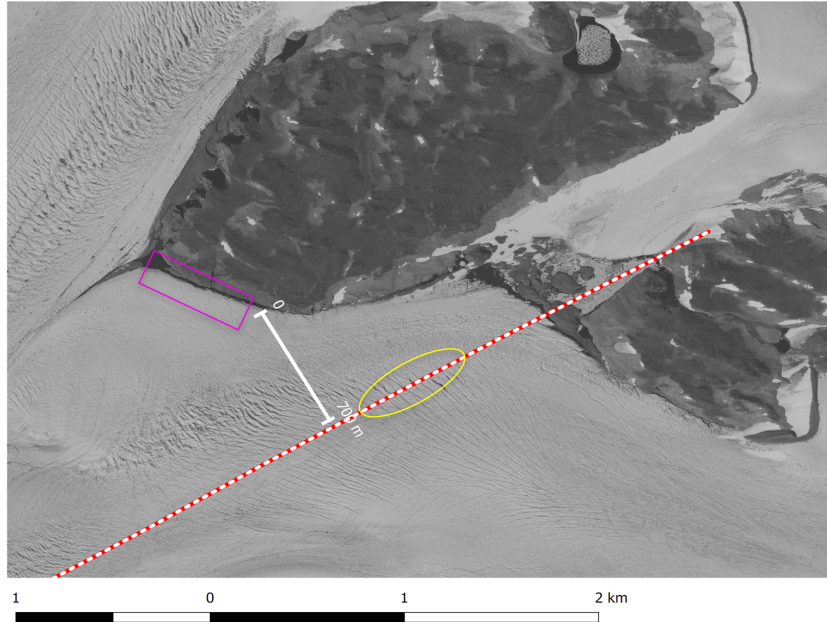


Figure 5.15: Clutter and subsurface discrimination dataset (track GOG/JKB2j/BWN01a); optical image acquired on Greenland (DigitalGlobe).

the projection matches the distance indicated in Figure 5.15 between the track and the nunatak, which is the source of the clutter feature. This result is also supported by the quantitative metrics in Table 5.4, where the SAD is 0.39 for the clutter hypothesis. This means that there is good similarity between the real and the simulated (clutter hypothesis) histograms. Conversely, a SAD factor of 1.18 expresses a lower similarity value between the real histogram and the nadir simulation. For feature f_2 , both the qualitative (Figure 5.18) and the quantitative (Table 5.5) results support that it is a nadir feature. Taking advantage of the available optical image (computing the projection of feature f_2 on the surface) we confirm that the cross track projection of the feature do not match any off nadir return. This result is confirmed by the quantitative metrics in Table 5.5, where the SAD is 1.59 for the clutter hypothesis. This means that there is low similarity between the real and the simulated histograms when clutter hypothesis is considered. On the other side, a SAD factor of 0.18 expresses a higher similarity value between the real histogram and the simulated histogram under nadir hypothesis (with a phase difference mean equal to 0.25).

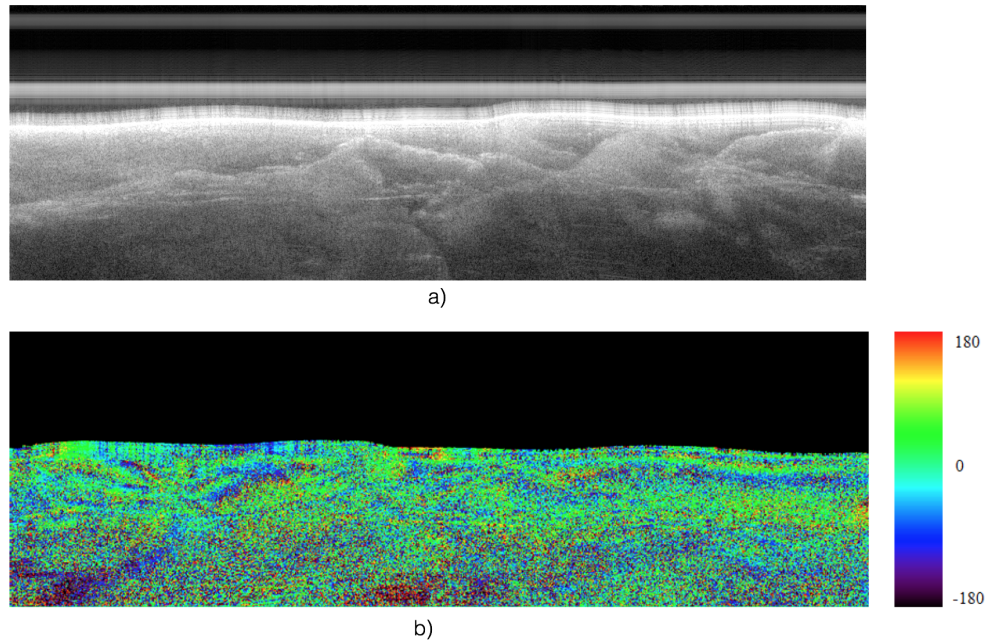


Figure 5.16: Clutter and subsurface discrimination dataset (track GOG/JKB2j/BWN01a); (a) amplitude radargram obtained from a single channel (left) of MARFA radar (8 km long portion of the entire acquisition), (b) RS interferogram obtained with the proposed method.

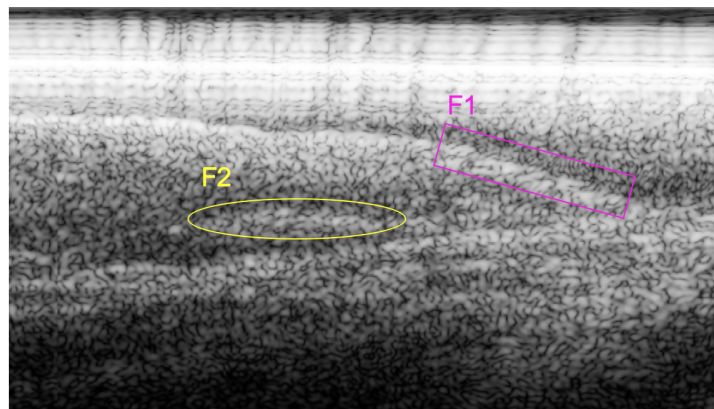


Figure 5.17: Clutter and subsurface discrimination dataset (track GOG/JKB2j/BWN01a); zoomed portion of the radargram containing the selected features.

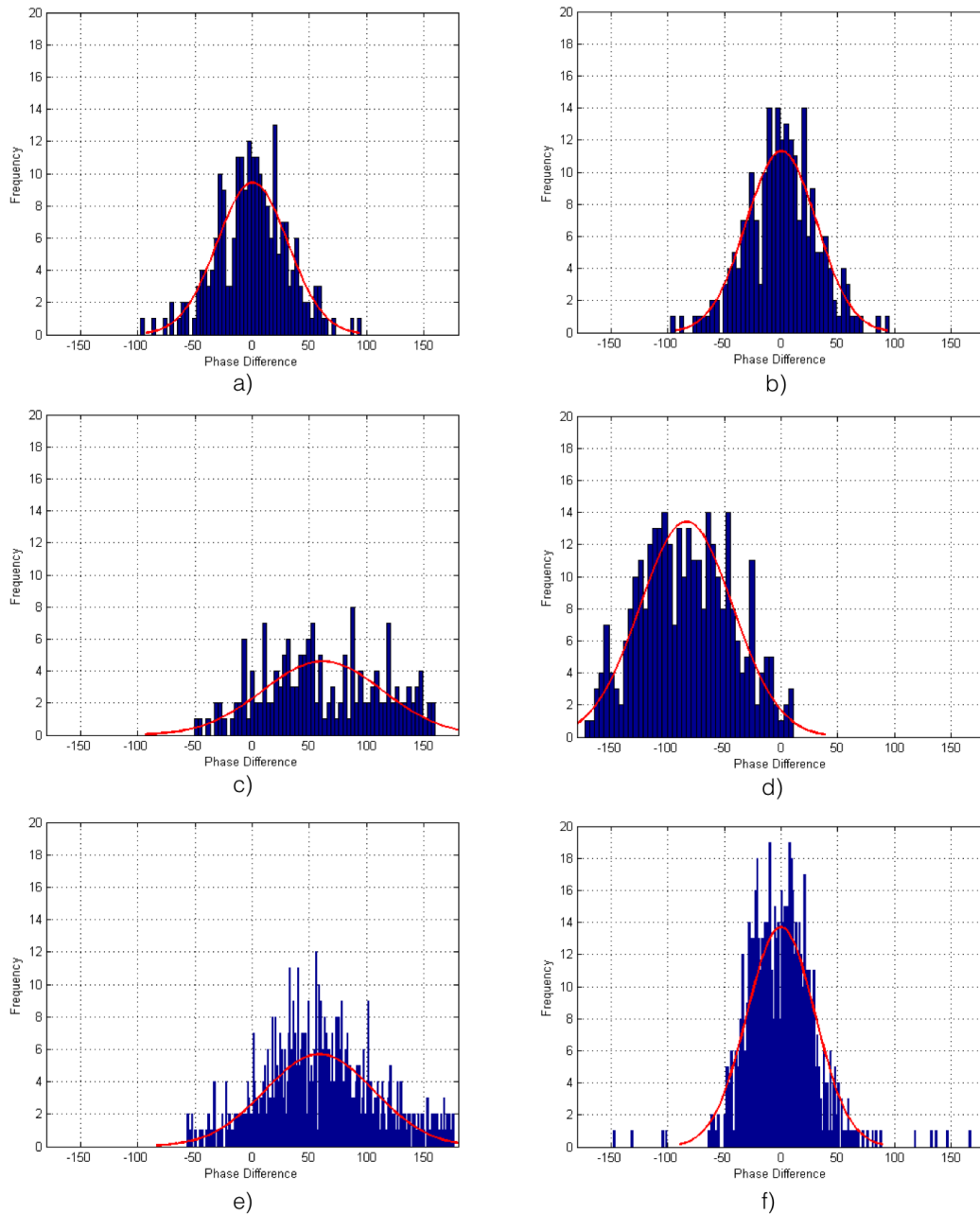


Figure 5.18: Feature f_1 shown in Figures 5.15 and 5.17: (a) theoretical phase difference histogram for nadir hypothesis and (c) clutter hypothesis and (e) real phase difference histogram. Feature f_2 shown in Figures 5.15 and 5.17: (b) theoretical phase difference histogram for nadir hypothesis and (d) clutter hypothesis and (f) real phase difference histogram.

Table 5.4: SAD values obtained between the real and the simulated (on both nadir and clutter hypothesis) histograms and the mean of the simulated distributions for feature f_1 .

Feature Type	SAD	mean
HP1: nadir	1.18	0.24
HP2: clutter	0.39	61.53

Table 5.5: SAD values obtained between the real and the simulated (on both nadir and clutter hypothesis) histograms and mean of the simulated distributions for feature f_2 .

Feature Type	SAD	mean
HP1: nadir	0.18	0.25
HP2: clutter	1.59	-83.4

5.4 Cross track clutter detection in planetary RS

In order to extend the study presented on airborne RS data to clutter detection using planetary RS, we compare the theoretical phase difference obtained for two observation scenarios: i) airborne sounding and ii) planetary sounding. This implies the use of two different platforms carrying the radar, i.e., aircraft and satellite respectively. It is worth noting that the airborne platform flies at a lower altitude than the satellite one and it is more affected by motion, roll in particular. Despite the analysis presented in this Section can be generalized to a large number of radar sounders, we compare: i) a system having the same design of MARFA with ii) a planetary radar operating on the Jupiter Moons. The planetary radar is supposed to be a two channels sounder having two main differences at system level with respect to MARFA: a lower central frequency and a slightly shorter baseline. In particular, the central frequency chosen is 9 MHz, which is the same of the RIME radar (Bruzzone et al. [2015]) currently under development for the JUICE mission, while the selected baseline is 10 meter.

The first part of this analysis is devoted to the comparison of the theoretical phase difference among the two channels for both the airborne and the satellite case. Figure 5.19 shows two examples of the phase difference behavior for a single frame. It is worth noting that the flight altitude and the length of the baseline strongly affect the phase difference. Thus, these variables have an impact on the effectiveness of the proposed method. In particular, for a given baseline in the data acquired by a low-altitude airborne system (see Figure 5.19a) the phase wraps rapidly in the cross-track, which can cause phase ambiguities for single pulses. For higher altitudes (see Figure 5.19b), the phase difference is more slowly varying, giving a unique phase signature in the across-track direction when the two channel RS system reaches the altitude of a satellite platform. Consequently, in the satellite case we expect to have a phase difference equal to zero only for a nadir subsurface feature. This can be understood by comparing Figure 5.19a and Figure 5.19b. Figure 5.19a shows that for a flight altitude of 1000 m with a baseline large 19 m we can estimate to have 7 phase difference periods in the radar footprint that can cause cross-track and nadir echo ambiguity. On the other side, Figure 5.19b shows that for a flight altitude of 400 km with a baseline large 10 m we can estimate to have only 1 phase difference period in the radar footprint that leads to have no cross-track and nadir echo ambiguity. The second part of the analysis aims to simulate the expected received radargram and the related interferogram in the same planetary case presented in 5.19b. Moreover, in the simulation we assess the capability of the proposed method to discriminate clutter features from subsurface nadir returns. The synthetic clutter feature corresponds to a depression of 300 meters depth on the Europa surface that could be compared to a crater in the real

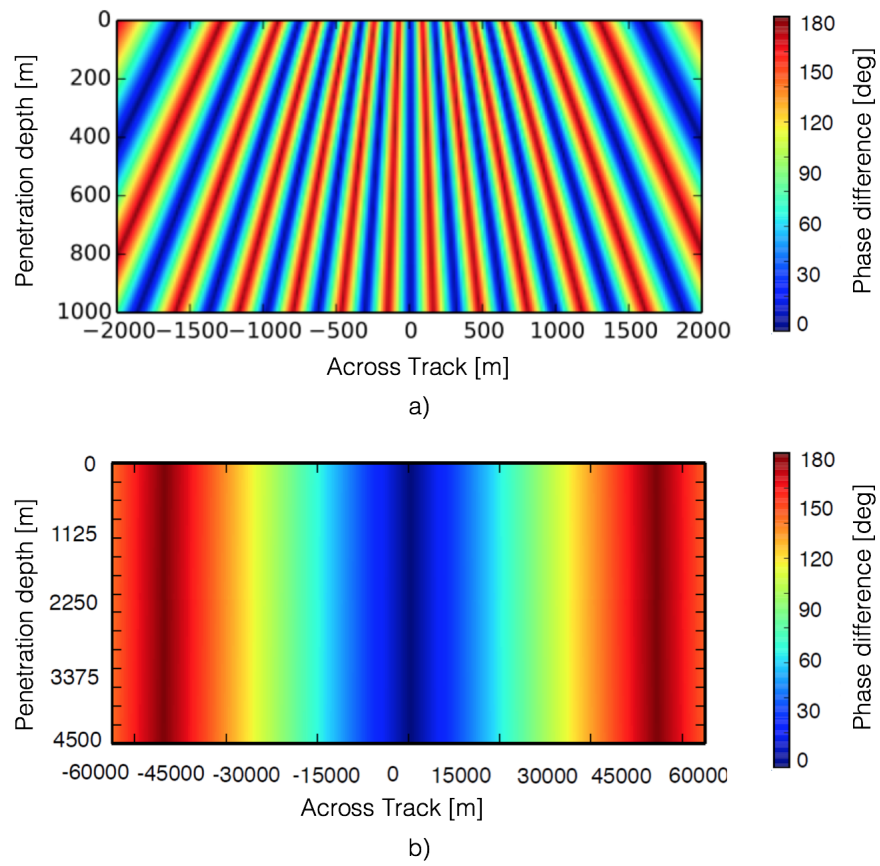


Figure 5.19: Theoretical phase difference behavior for a single range line in the case of a) phase difference expected at low altitude, with $H_0 = 1000 \text{ m}$ and $B = 19$ (MARFA airborne radar configuration); b) phase difference expected for satellite application, with $H_0 = 400 \text{ km}$ and $B = 10 \text{ m}$.

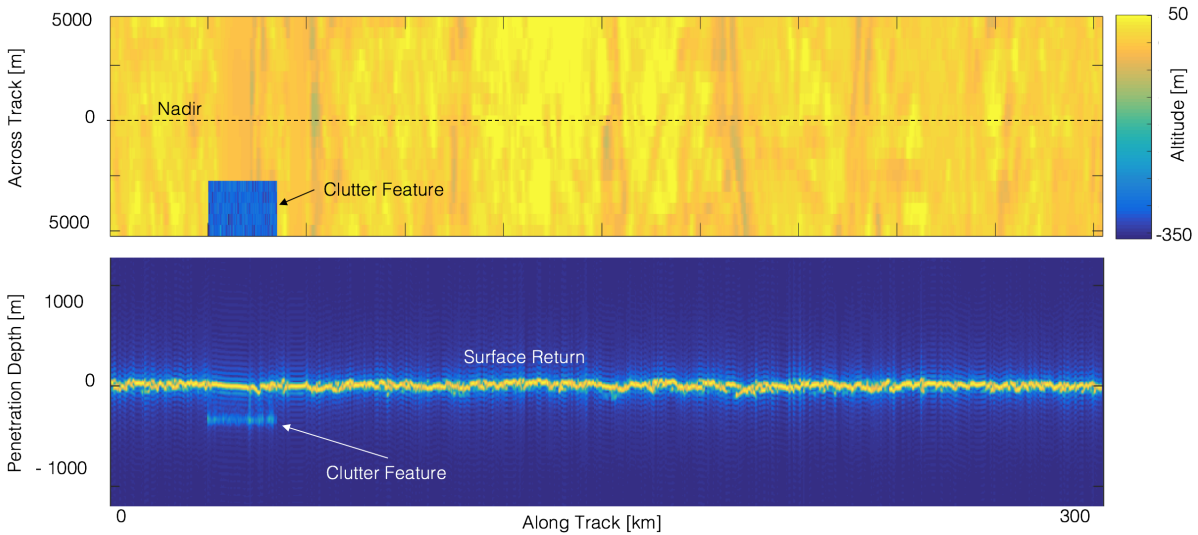


Figure 5.20: Configuration of the simulation for RS orbiter case: (a) digital elevation model of the Europa icy moon and synthetic clutter feature. (b) Simulated radargram.

scenario. This off-nadir surface feature is located 2.5 Km far from the nadir position and it is 2.5 Km large (see Figure 5.20). Figure 5.21 shows a radargram simulation (Iorio et al. [2008]) containing the surface and a clutter feature. In Figure 5.21b one can observe the simulated interferogram, where the clutter and the nadir surface returns have different values in terms of phase difference (i.e., $\Delta\Phi = 0$ for the nadir return and $\Delta\Phi \neq 0$ for the off-nadir clutter). The histograms shown in Figures 5.22a and 5.22b allows one to unequivocally discriminate the clutter from the nadir echo features. In Figure 5.22b one can notice that the phase difference at the surface is always zero. When the experimental airborne data is exploited (see Figure 5.10), the situation at the surface is different, i.e., the phase difference at the surface can assume non-zero values. This is mainly due to the effect of the surface topography at the MARFA central frequency (i.e., 60 MHz). We observed that for a flat surface, as in the case of data acquired on the ocean surface (used to calibrate the roll correction algorithm), the phase difference is perfectly equal to zero. When the surface is rough, higher is the frequency higher is the impact of the surface roughness on the phase stability.

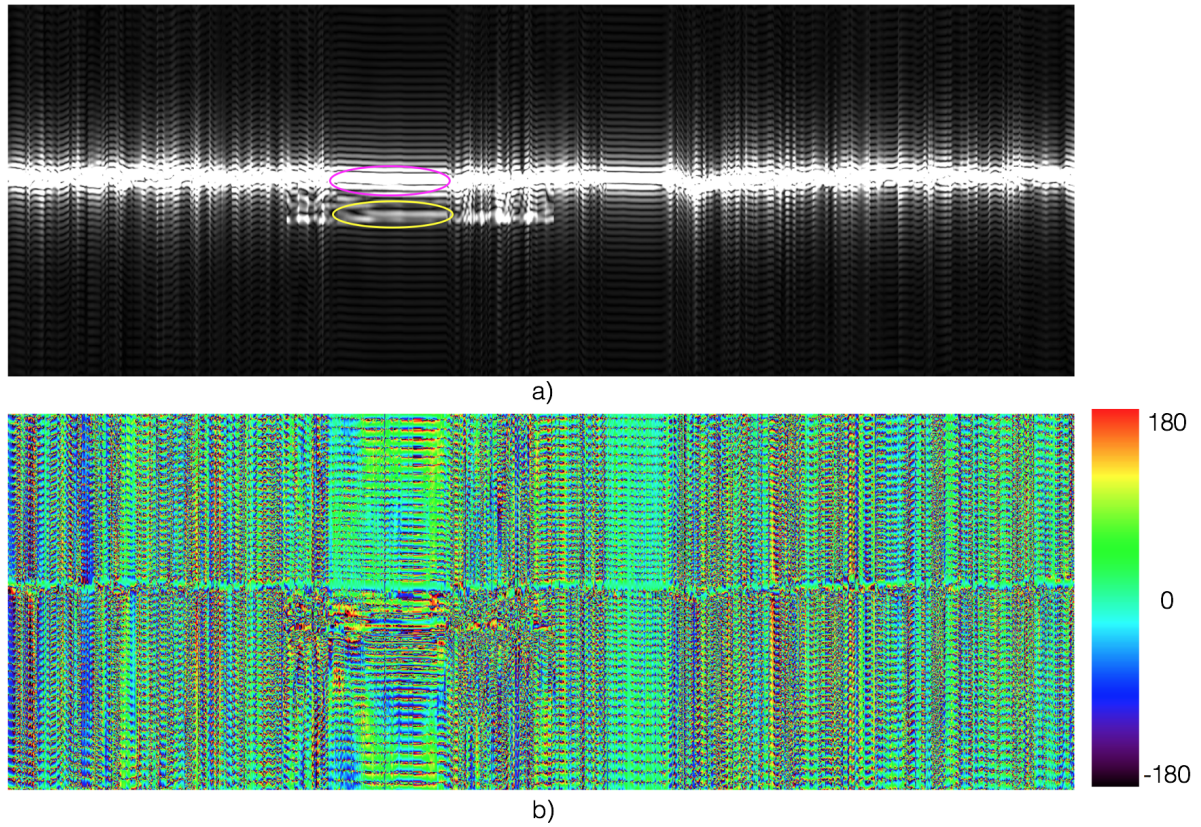


Figure 5.21: Planetary two-channel RS data simulation at Europa (10 meters baseline and 400 km orbit altitude): (a) radargram and (b) interferogram. (c) Phase difference histogram for the magenta feature (surface echo). (d) Phase difference histogram for the yellow feature (clutter).

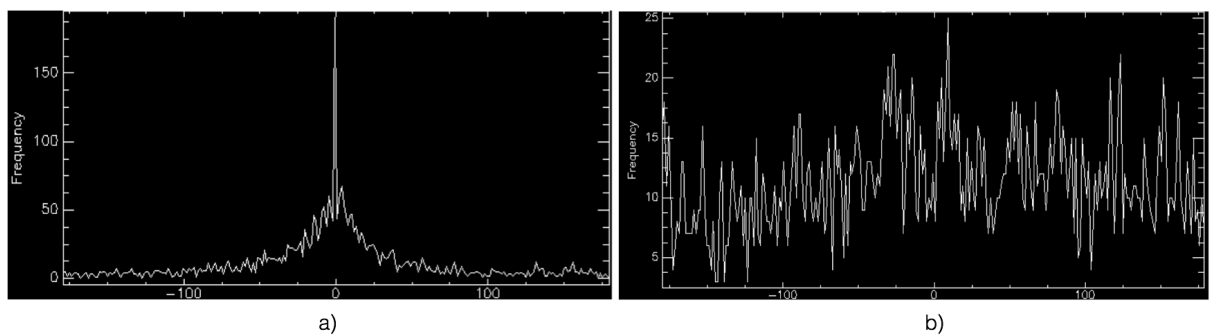


Figure 5.22: Planetary two-channel RS data simulation at Europa (10 meters baseline and 400 km orbit altitude): (a) radargram and (b) interferogram. (c) Phase difference histogram for the magenta feature (surface echo). (d) Phase difference histogram for the yellow feature (clutter).

5.5 Discussion and conclusion

In this Chapter, we have presented a proof of concept for a new radar sounding and analysis approach, which can discriminate between surface clutter and subsurface signal. The proposed method achieves clutter feature discrimination through an analysis of two-channel RS data. By studying the difference between estimated and observed phase-difference distributions, surface clutter can be discriminated from subsurface signal at nadir. This provides a solution to the problem of clutter discrimination in RS data interpretation, which is not possible using only a single-channel RS. Theoretical analyses and experimental results demonstrate that the proposed interferometric method is effective for: i) detection of clutter or nadir echoes separately and ii) discrimination of nadir from clutter echoes independently from ancillary data on surface topography. An additional feature of the proposed technique is the capability to discriminate a nadir subsurface from subsurface clutter features. However, we could not validate this feature due the lack of reference data on the subsurface. Limitations of the proposed method for low altitude acquisition have been discussed in Section 5.2. Despite these challenges, the experimental results obtained on three different datasets, (acquired from MARFA radar on Greenland), demonstrate the effectiveness of the interferometric approach. Simulations show that the performance of the proposed method dramatically improves when a satellite platform carries the radar on board, which is expected to be the primary target application for the proposed technique. This significantly enhances the scientific interpretation and the scientific return of planetary RS data. This enhancement is achieved through a method that requires the low-complexity hardware considering a relatively simple two-channel RS system. Therefore, it is particularly relevant in the design of planetary RSs for which resources are limited. For these reasons, the proposed interferometric technique is an interesting candidate to solve the clutter detection problem in the Europa mission Moussessian et al. [2015]. As future development, we plan to study a repeat pass interferometry approach for clutter detection in radargram acquired by orbital RS.

Chapter 6

Conclusions and future work

This Chapter concludes the thesis by recalling the context of the addressed research and by providing a summary of the novel contributions and a critical discussion of the related experimental results presented in the document. Finally, we provide ideas for future developments of the proposed methods.

6.1 Research context and summary of novel contributions

In this thesis we have presented novel signal processing methods for the estimation of biophysical parameters and for the analysis of icy terrains. The proposed methods represent valuable contributions for environmental monitoring analyzing remote sensing data acquired with active sensors. Such methods improve the quality of the analysis dealing with the scarcity of the reference samples. This allows the reduction of the operational cost associated to the collection of reference measurements, which is essential to support both practical application scenarios and scientific studies.

In the first part of this thesis, we introduced two main novel contributions on the development of regression techniques for the estimation of biophysical parameters. In particular, the proposed methods are: 1) A novel hybrid method for the correction of the theoretical model inversion in bio/geophysical parameter estimation; and 2) a novel semisupervised method for the estimation of biophysical parameters based on Support Vector Regression.

In the second part, we presented a novel technique based on interferometry for clutter detection in two-channel RS data.

Biophysical parameters estimation

In the first contribution of this thesis, we developed a novel hybrid approach to integrate theoretical models with empirical observations related to few reference samples to esti-

mate biophysical parameters. This novel method is made up of two main steps. In the first step, a theoretical method is initially inverted to estimate the target values. Then, the deviation between the target values estimated by the theoretical method inversion and the true target values is computed by exploiting field reference samples. To characterize the deviation we developed two different strategies. The first strategy, called *global deviation bias* (GDB), computes a constant for all the samples under the assumption that the deviation is globally independent and identically distributed in the feature space. The second is the *local deviation bias* (LDB) strategy that computes the deviation according to the position of the samples in the feature space under the hypothesis that the deviation changes in different portions of the feature space. We have shown the effectiveness of both the GDB and LDB strategies in estimating SMC from microwave sensor datasets acquired by airborne. The experimental results obtained on three different datasets demonstrate that the proposed method improves the estimation performance by reducing the impact of intrinsic errors due to the analytical formulation of theoretical models. In the experiments, we evaluated different feature combinations, such as different polarization and incident angles. For each dataset, the GDB strategy was able to reduce the estimation error (in terms of MAE and MSE) with respect to the inverse theoretical model. The LDB strategy demonstrated its capability to further improve the correction performance by modeling the deviation under more realistic assumptions. However, the LDB strategy requires a higher number of good quality reference samples than the GDB strategy to accurately model the deviation. For this reason, in real world application, when few reference samples are available, the GDB approach is preferred. We want to highlight that the proposed hybrid method can be applied to other estimation problems by considering proper theoretical models.

In the second contribution of this thesis, we have presented a novel SSL method for the estimation of biophysical parameters from remotely sensed images in the framework of ε -insensitive support vector regression (SVR). The proposed method is made up of two steps. The first step aims at injecting additional prior information in the initial learning of the SVR function in order to avoid the problems occurred due to the availability of the small-sized biased training sets. This is achieved by giving different weights to training samples that are obtained by a novel strategy that relies on the density of samples in the feature space. According to the developed strategy higher weights are given to the training samples located in the higher density locations of the feature space, and vice versa. Then, a weighted SVR algorithm is introduced in order to adopt different weights to training samples. The second step aims at jointly using labeled samples and a set of informative unlabeled samples in order to better characterize the SVR learning function.

This is achieved by introducing a new term in the WSVR presented in the first step. In order to select the informative unlabeled samples that are located as closest as possible to the ε -tube and outside of the ε -tube a novel strategy is presented. The proposed SSL has been applied to the variables extracted from two datasets related to: i) the estimation of the tree stem volume and ii) the estimation of the age of abalones. The experimental results show that the proposed strategy allows one to improve the estimation performance with zero cost of the collection of reference samples. The main advantages of the proposed method are: i) the properly modeling of the SVR learning function when small-sized biased sets of training samples are available and ii) the reduction of the operational cost associated to the collection of reference measurements, ii) the improvement in terms of performance of the parameter estimation by introducing unlabeled samples in the training phase of the learning.

Clutter detection method

In the third contribution of this thesis, we have presented a novel clutter detection method in the context of RS data analysis. In particular, we relies on the analysis of two-channel RS data by using a single-pass interferometric approach to study the phase-difference distribution. The proposed method is based on three steps. The first step allows the manual selection of a RS feature (clutter or subsurface) and the accurate modeling of the theoretical phase difference among the cross-track antennas of the RS. The second step is devoted to generate a radar sounder interferogram for extracting the real phase-difference distribution related to the feature selected in step 1. Finally, in step 3, the histograms of real and simulated phase difference are compared to define if the selected feature is a clutter or a subsurface echo. The experimental results have been obtained on three different datasets acquired from a two-channel RS system working at 60 MHz on Greenland and they demonstrate the effectiveness of the interferometric approach. The main advantages with respect to state of the art method for clutter suppression and detection are: i) the capability to detect the clutter independently from reference data on the surface topography, and ii) the possibility to limit hardware complexity of airborne RS systems. Moreover, the proposed clutter detection method can be a valuable solution to discriminate nadir subsurface from subsurface lateral echoes. Finally, the formation of an interferogram image using RS data is novel and is a relevant contribution to the field of RS data analysis. We identified two main limitations of the proposed interferometric method: i) completely overlapped subsurface and clutter features cannot be discriminated, and ii) low flight altitude acquisition can limit its effectiveness. The first one, that affects also single-channel RS systems, can be solved by analyzing longer features in the along-track direction of the radargram. This is reasonable because usually a subsurface return ex-

tends beyond the area overlapped with the clutter. The second problem occurs in airborne observations, where the relatively low flight altitude makes the phase difference changing fast. This decreased phase stability implies also an high sensitivity of the interferometric phase to the surface roughness. In this thesis, the analysis has been also extended to the satellite case. Simulated data shown that the performance of the proposed method dramatically improves when the RS is on board of a high altitude platform. In addition to this, the proposed method can use a low-complexity hardware system, which is particularly suitable in the design of satellite RSs. This can be fundamental in planetary mission aimed to explore remote celestial bodies. In this type of applications, the topographic reference information are usually limited, thus the proposed method allows the clutter discrimination without ancillary data. For this reason, a primary target application of the proposed clutter detection method is the analysis of terrains using two-channel RSs carried on satellite platform.

6.2 Concluding remarks and future developments

The studies, the methods and the results presented in this thesis regard the analysis of microwave remote sensing data. Such methodologies are efficient instruments to optimize the research on climate change and ice sheets dynamics. In this context, microwave remote sensing data represents a unique non-intrusive mean to obtain large scale information. The goal of the presented contributions is to correctly interpret this data with general techniques applicable in different fields of research and easily scalable to a wide range of applications.

Future developments regarding the regression methods for the estimation of biophysical parameters focus on their full generalization in order to make them suited for the analysis of a wider set of data. In particular, future works should address the following points:

- Extend the proposed hybrid approach to the use of other functional forms (polynomial, exponential, etc.) to model the mapping function.
- Test the effectiveness of the proposed hybrid approach in other estimation problems (e.g. the retrieval of snow pack variables in mountain areas).
- Study new techniques to compute the deviation between the inverse theoretical model and the empirical reference data.
- Test the effectiveness of the proposed SSL method on other estimation problems (e.g. the retrieval of SMC from satellite data).

Concerning the proposed method for clutter detection using radar sounder data, the future development should be devote to the:

- Development of a repeat pass procedure to apply the interferometric approach for clutter detection to radargrams acquired by orbital RS.
- Extension of the proposed method to a completely automatic processing chain.
- Test of the proposed clutter detection method on different real datasets acquired at different frequencies or at different altitude on the polar region of Earth.

Chapter 7

List of Publications

7.1 Journal Papers

1. D. Castelletti, L. Pasolli, L. Bruzzone, C. Notarnicola and B. Demir, "A Novel Hybrid Method for the Correction of the Theoretical Model Inversion in Bio/Geophysical Parameter Estimation," inIEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4764-4774, Aug. 2016.
2. D. Castelletti, D. M. Schroeder, S. Hensley, C. Grima, G. Ng, D. Young, Y. Gim, L. Bruzzone, A. Moussessian and D. D. Blankenship, "An Interferometric Approach to Cross-Track Clutter Detection in Two Channel VHF Radar Sounders," submitted to IEEE Transactions on Geoscience and Remote Sensing.
3. F. Di Paolo, S. E. Lauro, D. Castelletti, G. Mitri, F. Bovolo; B. Cosciotti, E. Mattei, R. Orosei, C. Notarnicola, L. Bruzzone, E. Pettinelli, "Radar Signal Penetration and Horizons Detection on Europa Through Numerical Simulations," inIEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol.PP, no.99, pp.1-12.

7.2 Conferences

1. D. Castelletti, B. Demir, L. Bruzzone, "A semisupervised support vector regression method to estimate biophysical parameters from remotely sensed images," Proc. SPIE 9244, Image and Signal Processing for Remote Sensing, Oct. 2014.
2. D. Castelletti, D. M. Schroeder, S. Hensley, C. Grima, G. Ng, D. Young, Y. Gim, L. Bruzzone, A. Moussessian and D. D. Blankenship, "Clutter detection using two-

- channel radar sounder data,” 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, 2015, pp. 1052-1055.
3. C. Notarnicola, L. Pasolli, G. Cuzzo, F. Greifeneder, G. Bertoldi, S. Della Chiesa, G. Niedrist, D. Castelletti, U. Tappeiner, L. Bruzzone, M. Zebisch, ”Temporal and spatial soil moisture dynamics in mountain meadows by integrating Radarsat 2 images and ground data,” 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, 2014, pp. 1222-1225.
 4. L. Bruzzone, J. J. Plaut, G. Alberti, D. D. Blankenship, F. Bovolo, B. A. Campbell, D. Castelletti, Y. Gim, A. M. Ilisei, W. Kofman, G. Komatsu, W. McKinnon, G. Mitri, A. Moussessian, C. Notarnicola, R. Orosei, E. Pettinelli D. Plettemeier, ”Jupiter ICY moon explorer (JUICE): Advances in the design of the radar for Icy Moons (RIME),” 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, 2015, pp. 1257-1260.

Bibliography

- Alpaydin, Ethem. *Introduction to Machine Learning*. MIT Press, 2009. ISBN 9780262315548.
- Bache, K. and Lichman, M. Uci machine learning repository. *University of California, School of Information and Computer Science, Irvine, CA*, 2013. doi: <http://archive.ics.uci.edu/ml>.
- Bamber, J. L.; Griggs, J. A.; Hurkmans, R. T. W. L.; Dowdeswell, J. A.; Gogineni, S. P.; Howat, I.; Mouginot, J.; Paden, J.; Palmer, S.; Rignot, E., and Steinhage, D. A new bed elevation dataset for greenland. *The Cryosphere*, 7(2):499–510, 2013. doi: 10.5194/tc-7-499-2013. URL <http://www.the-cryosphere.net/7/499/2013/>.
- Bazi, Y.; Alajlan, N., and Melgani, F. Improved estimation of water chlorophyll concentration with semisupervised gaussian process regression. *IEEE Transactions on Geoscience and Remote Sensing*, 50(7):2733–2743, July 2012. ISSN 0196-2892. doi: 10.1109/TGRS.2011.2174246.
- Bekaert, D.; Gebert, N.; Lin, C.; Heliere, F.; Dall, J.; Kusk, A., and Kristensen, S. Savstrup. Multichannel surface clutter suppression: East antarctica p-band sar ice sounding in the presence of grating lobes. *Annals of Glaciology*, 55(67), 2014. doi: 10.3189/2014AoG67A100.
- Bennett, K. and Demiriz, A. Semi-supervised support vector machines. *NIPS 11*, page 368374, 1998.
- Blum, Avrim and Mitchell, Tom. Combining labeled and unlabeled data with co-training. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory, COLT' 98*, pages 92–100, New York, NY, USA, 1998. ACM. ISBN 1-58113-057-0. doi: 10.1145/279943.279962. URL <http://doi.acm.org/10.1145/279943.279962>.
- Bogorodsky, V.; Bentley, C.; , and Gudmandsen, P. *Radioglaciology*. University of Michigan Press, d. h. d. reidel publishing company edition, 1984.
- Bolten, J. D.; Crow, W. T.; Zhan, X.; Jackson, T. J., and Reynolds, C. A. Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(1):57–66, March 2010. ISSN 1939-1404. doi: 10.1109/JSTARS.2009.2037163.
- Bruzzone, L. and Melgani, F. Robust multiple estimator systems for the analysis of biophysical parameters from remotely sensed data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(1):159–174, Jan 2005. ISSN 0196-2892. doi: 10.1109/TGRS.2004.839818.
- Bruzzone, L.; Plaut, J. J.; Alberti, G.; Blankenship, D. D.; Bovolo, F.; Campbell, B. A.; Castelletti, D.; Gim, Y.; Ilisei, A. M.; Kofman, W.; Komatsu, G.; McKinnon, W.; Mitri, G.; Moussessian, A.; Notarnicola, C.; Orosei, R.; Patterson, G. W.; Pettinelli, E., and Plettemeier, D. Jupiter icy moon explorer (juice): Advances in the design of the radar for icy moons (rime). In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1257–1260, July 2015. doi: 10.1109/IGARSS.2015.7326002.

- Camps-Valls, G.; Bruzzone, L.; Rojo-Alvarez, J. L., and Melgani, F. Robust support vector regression for biophysical variable estimation from remotely sensed images. *IEEE Geoscience and Remote Sensing Letters*, 3(3):339–343, July 2006. ISSN 1545-598X. doi: 10.1109/LGRS.2006.871748.
- Camps-Valls, G.; Munoz-Mari, J.; Gomez-Chova, L.; Richter, K., and Calpe-Maravilla, J. Biophysical parameter estimation with a semisupervised support vector machine. *IEEE Geoscience and Remote Sensing Letters*, 6(2):248–252, April 2009. ISSN 1545-598X. doi: 10.1109/LGRS.2008.2009077.
- Cha, Sung-Hyuk and Srihari, Sargur N. On measuring the distance between histograms. *Pattern Recognition*, 35(6):1355 – 1370, 2002. ISSN 0031-3203. doi: [http://dx.doi.org/10.1016/S0031-3203\(01\)00118-2](http://dx.doi.org/10.1016/S0031-3203(01)00118-2). URL <http://www.sciencedirect.com/science/article/pii/S0031320301001182>.
- Cohen, I.; Cozman, F. G.; Sebe, N.; Cirelo, M. C., and Huang, T. S. Semisupervised learning of classifiers: theory, algorithms, and their application to human-computer interaction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(12):1553–1566, Dec 2004. ISSN 0162-8828. doi: 10.1109/TPAMI.2004.127.
- Collobert, Ronan; Bengio, Samy, and Williamson, C. Svmtorch: Support vector machines for large-scale regression problems. *Journal of Machine Learning Research*, 1:143–160, 2001.
- Cortes, Corinna and Mohri, Mehryar. On transductive regression. In *Proceedings of the 19th International Conference on Neural Information Processing Systems, NIPS’06*, pages 305–312, Cambridge, MA, USA, 2006. MIT Press. URL <http://dl.acm.org/citation.cfm?id=2976456>.2976495.
- Croci, R.; Seu, R.; Flamini, E., and Russo, E. The shallow radar (sharad) onboard the nasa mro mission. *Proceedings of the IEEE*, 99(5):794–807, May 2011. ISSN 0018-9219. doi: 10.1109/JPROC.2010.2104130.
- Dall, J.; Kristensen, S. S.; Krozer, V.; Hernandez, C. C.; Vidkjær, J.; Kusk, A.; Balling, J.; Skou, N.; Sobjrg, S. S., and Christensen, E. L. Esa’s polarimetric airborne radar ice sounder (polaris): design and first results. *IET Radar, Sonar Navigation*, 4(3):488–496, June 2010a. ISSN 1751-8784. doi: 10.1049/iet-rsn.2009.0035.
- Dall, J.; Kristensen, S. S.; Krozer, V.; Hernandez, C. C.; Vidkjær, J.; Kusk, A.; Balling, J.; Skou, N.; Sobjrg, S. S., and Christensen, E. L. Esa’s polarimetric airborne radar ice sounder (polaris): design and first results. *IET Radar, Sonar Navigation*, 4(3):488–496, June 2010b. ISSN 1751-8784. doi: 10.1049/iet-rsn.2009.0035.
- Dalponte, M.; Bruzzone, L., and Gianelle, D. A system for the estimation of single-tree stem diameter and volume using multireturn lidar data. *IEEE Transactions on Geoscience and Remote Sensing*, 49(7):2479–2490, July 2011. ISSN 0196-2892. doi: 10.1109/TGRS.2011.2107744.
- Darvishzadeh, R.; Skidmore, A.; Schlerf, M., and Atzberger, C. Inversion of a radiative transfer model for estimating vegetation {LAI} and chlorophyll in a heterogeneous grassland. *Remote Sensing of Environment*, 112(5):2592 – 2604, 2008. ISSN 0034-4257. doi: <http://dx.doi.org/10.1016/j.rse.2007.12.003>. URL <http://www.sciencedirect.com/science/article/pii/S0034425707004968>. Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue.
- Demir, B. and Bruzzone, L. A multiple criteria active learning method for support vector regression. *Pattern Recognition*, 47(7):2558 – 2567, 2014. ISSN 0031-3203. doi: <http://dx.doi.org/10.1016/j.patcog.2014.02.001>. URL <http://www.sciencedirect.com/science/article/pii/S0031320314000375>.
- Drucker, Harris; Burges, Christopher J. C.; Kaufman, Linda; Smola, Alex J., and Vapnik, Vladimir. Support vector regression machines. In Mozer, M. C.; Jordan, M. I., and Petsche, T., editors, *Advances in Neural Information Processing Systems 9*, pages 155–161. MIT Press, 1997. URL <http://papers.nips.cc/paper/1238-support-vector-regression-machines.pdf>.

- Durba, A. A.; King, R. L., and Younan, N. H. Support vector regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment*, 107:348 – 361, 2007.
- Duveiller, Grgory; Weiss, Marie; Baret, Frdric, and Defourny, Pierre. Retrieving wheat green area index during the growing season from optical time series measurements based on neural network radiative transfer inversion. *Remote Sensing of Environment*, 115(3):887 – 896, 2011. ISSN 0034-4257. doi: <http://dx.doi.org/10.1016/j.rse.2010.11.016>. URL <http://www.sciencedirect.com/science/article/pii/S0034425710003354>.
- Egmont-Petersen, M.; de Ridder, D., and Handels, H. Image processing with neural networks a review. *Pattern Reconition Journal*, 25(10):2279–2301, Oct 2002. doi: [http://dx.doi.org/10.1016/S0031-3203\(01\)00178-9](http://dx.doi.org/10.1016/S0031-3203(01)00178-9).
- Eineder, M.; Minet, C.; Steigenberger, P.; Cong, X., and Fritz, T. Imaging geodesy toward centimeter-level ranging accuracy with terrasar-x. *IEEE Transactions on Geoscience and Remote Sensing*, 49(2):661–671, Feb 2011. ISSN 0196-2892. doi: 10.1109/TGRS.2010.2060264.
- Engman, E.T. and Gurney, R.J. *Remote Sensing in Hydrology*. Chapman and Hall, London, 1991.
- 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.; Kimball, J.; Piepmeier, J. R.; Koster, R. D.; Martin, N.; McDonald, K. C.; Moghaddam, M.; Moran, S.; Reichle, R.; Shi, J. C.; Spencer, M. W.; Thurman, S. W.; Tsang, L., and Zyl, J. Van. The soil moisture active passive (smap) mission. *Proceedings of the IEEE*, 98(5):704–716, May 2010. ISSN 0018-9219. doi: 10.1109/JPROC.2010.2043918.
- Ferretti, A.; Prati, C., and Rocca, F. Nonlinear subsidence rate estimation using permanent scatterers in differential sar interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5):2202–2212, Sep 2000. ISSN 0196-2892. doi: 10.1109/36.868878.
- Ferro, A.; Pascal, A., and Bruzzone, L. A novel technique for the automatic detection of surface clutter returns in radar sounder data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5):3037–3055, May 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2219315.
- Fu, A.K. and Daniels, D.J. *Ground Penetrating Radar*. U.K.: Inst. Eng. Technol., 2nd ed. london edition, 2007.
- Fung, A. K. *Microwave Scattering and Emission Models and Their Applications*. Artech house edition, 1994.
- Fung, A. K.; Li, Z., and Chen, K. S. Backscattering from a randomly rough dielectric surface. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2):356–369, Mar 1992. ISSN 0196-2892. doi: 10.1109/36.134085.
- Gogineni, S.; Yan, J.-B.; Paden, J.; Leuschen, C.; Li, J.; Rodriguez-Morales, F.; Braaten, D.; Purdon, K.; Wang, Z.; Liu, W., and Gauch, J. Bed topography of jakobshavn isbrae, greenland, and byrd glacier, antarctica. *Journal of Glaciology*, 60(223), 2014.
- Grima, Cyril; Blankenship, Donald D., and Schroeder, Dustin M. Radar signal propagation through the ionosphere of europa. *Planetary and Space Science*, 117:421 – 428, 2015. ISSN 0032-0633. doi: <http://dx.doi.org/10.1016/j.pss.2015.08.017>. URL <http://www.sciencedirect.com/science/article/pii/S0032063315002470>.
- Hallikainen, M. T.; Ulaby, F. T.; Dobson, M. C.; El-rayes, M. A., and k. Wu, L. Microwave dielectric behavior of wet soil-part 1: Empirical models and experimental observations. *IEEE Transactions on Geoscience and Remote Sensing*, GE-23(1):25–34, Jan 1985. ISSN 0196-2892. doi: 10.1109/TGRS.1985.289497.
- Hanssen, R. and Bamler, R. Evaluation of interpolation kernels for sar interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1):318–321, Jan 1999. ISSN 0196-2892. doi: 10.1109/36.739168.

- Heathman, Gary C; Starks, Patrick J; Ahuja, Lajpat R, and Jackson, Thomas J. Assimilation of surface soil moisture to estimate profile soil water content. *Journal of Hydrology*, 279(14):1 – 17, 2003. ISSN 0022-1694. doi: [http://dx.doi.org/10.1016/S0022-1694\(03\)00088-X](http://dx.doi.org/10.1016/S0022-1694(03)00088-X). URL <http://www.sciencedirect.com/science/article/pii/S002216940300088X>.
- Heliere, F.; Lin, C. C.; Corr, H., and Vaughan, D. Radio echo sounding of pine island glacier, west antarctica: Aperture synthesis processing and analysis of feasibility from space. *IEEE Transactions on Geoscience and Remote Sensing*, 45(8):2573–2582, Aug 2007. ISSN 0196-2892. doi: 10.1109/TGRS.2007.897433.
- Hoen, E. Weber and Zebker, H. A. Penetration depths inferred from interferometric volume decorrelation observed over the greenland ice sheet. *IEEE Transactions on Geoscience and Remote Sensing*, 38(6):2571–2583, Nov 2000. ISSN 0196-2892. doi: 10.1109/36.885204.
- Holt, John W.; Peters, Matthew E.; Kempf, Scott D.; Morse, David L., and Blankenship, Donald D. Echo source discrimination in single-pass airborne radar sounding data from the dry valleys, antarctica: Implications for orbital sounding of mars. *Journal of Geophysical Research: Planets*, 111(E6):n/a–n/a, 2006. ISSN 2156-2202. doi: 10.1029/2005JE002525. URL <http://dx.doi.org/10.1029/2005JE002525>. E06S24.
- Iorio, M.; Fois, F.; Mecozzi, R.; Catallo, C.; Picardi, G.; Seu, R., and Flamini, E. Fast radar signal simulator for sar ground penetrating applications. In *2008 IEEE Radar Conference*, pages 1–4, May 2008. doi: 10.1109/RADAR.2008.4720762.
- Jezeq, K. C.; Gogineni, S.; Wu, X.; Rodriguez, E.; Rodriguez-Morales, F.; Hoch, A.; Freeman, A., and Sonntag, J. G. Two-frequency radar experiments for sounding glacier ice and mapping the topography of the glacier bed. *IEEE Transactions on Geoscience and Remote Sensing*, 49(3):920–929, March 2011a. ISSN 0196-2892. doi: 10.1109/TGRS.2010.2071387.
- Jezeq, K. C.; Gogineni, S.; Wu, X.; Rodriguez, E.; Rodriguez-Morales, F.; Hoch, A.; Freeman, A., and Sonntag, J. G. Two-frequency radar experiments for sounding glacier ice and mapping the topography of the glacier bed. *IEEE Transactions on Geoscience and Remote Sensing*, 49(3):920–929, March 2011b. ISSN 0196-2892. doi: 10.1109/TGRS.2010.2071387.
- Jordan, R.; Picardi, G.; Plaut, J.; Wheeler, K.; Kirchner, D.; Safaeinili, A.; Johnson, W.; Seu, R.; Calabrese, D.; Zampolini, E.; Cicchetti, A.; Huff, R.; Gurnett, D.; Ivanov, A.; Kofman, W.; Orosei, R.; Thompson, T.; Edenhofer, P., and Bombaci, O. The mars express {MARSIS} sounder instrument. *Planetary and Space Science*, 57(1415):1975 – 1986, 2009. ISSN 0032-0633. doi: <http://dx.doi.org/10.1016/j.pss.2009.09.016>. URL <http://www.sciencedirect.com/science/article/pii/S0032063309002803>.
- Kang, Pilsung; Kim, Dongil, and Cho, Sungzoon. Semi-supervised support vector regression based on self-training with label uncertainty: An application to virtual metrology in semiconductor manufacturing. *Expert Systems with Applications*, 51:85 – 106, 2016. ISSN 0957-4174. doi: <http://dx.doi.org/10.1016/j.eswa.2015.12.027>. URL <http://www.sciencedirect.com/science/article/pii/S0957417415008295>.
- Karam, M. A.; Fung, A. K.; Lang, R. H., and Chauhan, N. S. A microwave scattering model for layered vegetation. *IEEE Transactions on Geoscience and Remote Sensing*, 30(4):767–784, Jul 1992. ISSN 0196-2892. doi: 10.1109/36.158872.
- Kwok, Tin-Yau and Yeung, Dit-Yan. Constructive algorithms for structure learning in feedforward neural networks for regression problems. *IEEE Transactions on Neural Networks*, 8(3):630–645, May 1997. ISSN 1045-9227. doi: 10.1109/72.572102.

- Lin, Q.; Vesecky, J. F., and Zebker, H. A. New approaches in interferometric sar data processing. *IEEE Transactions on Geoscience and Remote Sensing*, 30(3):560–567, May 1992. ISSN 0196-2892. doi: 10.1109/36.142934.
- Madsen, S. N.; Zebker, H. A., and Martin, J. Topographic mapping using radar interferometry: processing techniques. *IEEE Transactions on Geoscience and Remote Sensing*, 31(1):246–256, Jan 1993. ISSN 0196-2892. doi: 10.1109/36.210464.
- Mardia, K.V.; Kent, J.T.; , and Bibby, J.M. *Multivariate Analysis*. Academic Press, 1982.
- Massonnet, D.; Feigl, K.; Rossi, M., and Adragna, F. Radar interferometric mapping of deformation in the year after the landers earthquake. *Nature*, MAY 1994.
- Mattia, F.; Toan, T. Le; Picard, G.; Posa, F. I.; D’Alessio, A.; Notarnicola, C.; Gatti, A. M.; Rinaldi, M.; Satalino, G., and Pasquariello, G. Multitemporal c-band radar measurements on wheat fields. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7):1551–1560, July 2003. ISSN 0196-2892. doi: 10.1109/TGRS.2003.813531.
- Meroni, M.; Colombo, R., and Panigada, C. Inversion of a radiative transfer model with hyperspectral observations for {LAI} mapping in poplar plantations. *Remote Sensing of Environment*, 92(2):195 – 206, 2004. ISSN 0034-4257. doi: <http://dx.doi.org/10.1016/j.rse.2004.06.005>. URL <http://www.sciencedirect.com/science/article/pii/S0034425704001695>.
- Mingmin, C. and Bruzzone, L. Semisupervised classification of hyperspectral images by svms optimized in the primal. *IEEE Transactions on Geoscience and Remote Sensing*, 45(6):1870–1880, 2007.
- Moreira, A.; Prats-Iraola, P.; Younis, M.; Krieger, G.; Hajnsek, I., and Papathanassiou, K. P. A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, 1(1):6–43, March 2013. ISSN 2473-2397. doi: 10.1109/MGRS.2013.2248301.
- Moser, G. and Serpico, S. B. Automatic parameter optimization for support vector regression for land and sea surface temperature estimation from remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3):909–921, March 2009. ISSN 0196-2892. doi: 10.1109/TGRS.2008.2005993.
- Moussessian, A.; Blankenship, D. D.; Plaut, J. J.; Patterson, G. W.; Gim, Y.; Schroeder, D. M.; Soderlund, K. M.; Grima, C.; Young, D. A., and Chapin, E. REASON for Europa. *AGU Fall Meeting Abstracts*, December 2015.
- Nielsen, U. and Dall, J. Direction-of-arrival estimation for radar ice sounding surface clutter suppression. *IEEE Transactions on Geoscience and Remote Sensing*, 53(9):5170–5179, Sept 2015. ISSN 0196-2892. doi: 10.1109/TGRS.2015.2418221.
- Nielsen, U.; Yan, J. B.; Gogineni, S., and Dall, J. Direction-of-arrival analysis of airborne ice depth sounder data. *IEEE Transactions on Geoscience and Remote Sensing*, 55(4):2239–2249, April 2017. ISSN 0196-2892. doi: 10.1109/TGRS.2016.2639510.
- Notarnicola, C.; Angiulli, M., and Posa, F. Use of radar and optical remotely sensed data for soil moisture retrieval over vegetated areas. *IEEE Transactions on Geoscience and Remote Sensing*, 44(4):925–935, April 2006. ISSN 0196-2892. doi: 10.1109/TGRS.2006.872287.
- Notarnicola, C.; Angiulli, M., and Posa, F. Soil moisture retrieval from remotely sensed data: Neural network approach versus bayesian method. *IEEE Transactions on Geoscience and Remote Sensing*, 46(2):547–557, Feb 2008. ISSN 0196-2892. doi: 10.1109/TGRS.2007.909951.

- Nouvel, J. F.; Herique, A.; Kofman, W., and Safaenili, A. Radar signal simulation: Surface modeling with the facet method. *Radio Science*, 39(1):1–17, Feb 2004. ISSN 1944-799X. doi: 10.1029/2003RS002903.
- Paloscia, S.; Pampaloni, P.; Pettinato, S., and Santi, E. A comparison of algorithms for retrieving soil moisture from envisat/asar images. *IEEE Transactions on Geoscience and Remote Sensing*, 46(10):3274–3284, Oct 2008. ISSN 0196-2892. doi: 10.1109/TGRS.2008.920370.
- Pasolli, L.; Notarnicola, C., and Bruzzone, L. Estimating soil moisture with the support vector regression technique. *IEEE Geoscience and Remote Sensing Letters*, 8(6):1080–1084, Nov 2011. ISSN 1545-598X. doi: 10.1109/LGRS.2011.2156759.
- Pasolli, L.; Notarnicola, C., and Bruzzone, L. Multi-objective parameter optimization in support vector regression: General formulation and application to the retrieval of soil moisture from remote sensing data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(5):1495–1508, Oct 2012. ISSN 1939-1404. doi: 10.1109/JSTARS.2012.2197178.
- Pasolli, L.; Notarnicola, C.; Bertoldi, G.; Bruzzone, L.; Remelgado, R.; Greifeneder, F.; Niedrist, G.; Chiesa, S. Della; Tappeiner, U., and Zebisch, M. Estimation of soil moisture in mountain areas using svr technique applied to multiscale active radar images at c-band. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(1):262–283, Jan 2015. ISSN 1939-1404. doi: 10.1109/JSTARS.2014.2378795.
- Peters, M. E.; Blankenship, D. D.; Carter, S. P.; Kempf, S. D.; Young, D. A., and Holt, J. W. Along-track focusing of airborne radar sounding data from west antarctica for improving basal reflection analysis and layer detection. *IEEE Transactions on Geoscience and Remote Sensing*, 45(9):2725–2736, Sept 2007. ISSN 0196-2892. doi: 10.1109/TGRS.2007.897416.
- Picardi, G.; Biccari, D.; Bazzoni, A.; Fois, F.; Iorio, M.; Seu, R.; Melacci, P.; Federico, C.; Frigeri, A.; Minelli, G.; Marinangeli, L.; Orosei, R.; Calabrese, D.; Zampolini, E.; Johnson, W. T. K.; Jordan, R. L.; Plaut, J., and Safaenili, A. Mars advanced radar for subsurface and ionosphere sounding (marsis): subsurface performances evaluation. In *2003 Proceedings of the International Conference on Radar (IEEE Cat. No.03EX695)*, pages 515–521, Sept 2003. doi: 10.1109/RADAR.2003.1278795.
- Raney, R. K. Radar sounder: Cross-track polarimetric selectivity. In *7th European Conference on Synthetic Aperture Radar*, pages 1–4, June 2008.
- Rasmussen, Carl Edward. Gaussian processes for machine learning. MIT Press, 2006.
- Rodriguez, E. and Martin, J. M. Theory and design of interferometric synthetic aperture radars. *IEE Proceedings F - Radar and Signal Processing*, 139(2):147–159, April 1992. ISSN 0956-375X. doi: 10.1049/ip-f-2.1992.0018.
- Rodriguez, E.; Morris, C.S.; Belz, J.E.; Chapin, E.; Martin, J.; Daffer, W., and Hensley, S. An assessment of the srtm topographic products. *JPL Technical Report*, Jan 2005.
- Rolland, P.; Chabot, M.; Priard, R.; Ledantec, P., and Decoust, C. Radarsat-2 mission operations status. In *2012 IEEE International Geoscience and Remote Sensing Symposium*, pages 4582–4585, July 2012. doi: 10.1109/IGARSS.2012.6350449.
- Rosen, P. A.; Hensley, S.; Joughin, I. R.; Li, F. K.; Madsen, S. N.; Rodriguez, E., and Goldstein, R. M. Synthetic aperture radar interferometry. *Proceedings of the IEEE*, 88(3):333–382, March 2000. ISSN 0018-9219. doi: 10.1109/5.838084.

- Russo, F.; Cutigni, M.; Orosei, R.; Taddei, C.; Seu, R.; Biccari, D.; Giacomoni, E.; Fuga, O., and Flamini, E. An incoherent simulator for the sharad experiment. In *2008 IEEE Radar Conference*, pages 1–4, May 2008. doi: 10.1109/RADAR.2008.4720761.
- Sandholt, Inge; Rasmussen, Kjeld, and Andersen, Jens. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of Environment*, 79(23):213 – 224, 2002. ISSN 0034-4257. doi: [http://dx.doi.org/10.1016/S0034-4257\(01\)00274-7](http://dx.doi.org/10.1016/S0034-4257(01)00274-7). URL <http://www.sciencedirect.com/science/article/pii/S0034425701002747>. Recent Advances in Remote Sensing of Biophysical Variables.
- Schroeder, D. M.; Blankenship, D. D.; Young, D. A.; Witus, A. E., and Anderson, J. B. Airborne radar sounding evidence for deformable sediments and outcropping bedrock beneath thwaites glacier, west antarctica. *IEEE Transactions on Geoscience and Remote Sensing*, 41(20):7200–7208, Oct 2014.
- Shi, L.; Allen, C. T.; Ledford, J. R.; Rodriguez-Morales, F.; Blake, W. A.; Panzer, B. G.; Prokopiack, S. C.; Leuschen, C. J., and Gogineni, S. Multichannel coherent radar depth sounder for nasa operation ice bridge. In *2010 IEEE International Geoscience and Remote Sensing Symposium*, pages 1729–1732, July 2010. doi: 10.1109/IGARSS.2010.5649518.
- Smola, Alex J. and Schölkopf, Bernhard. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, 2004. ISSN 1573-1375. doi: 10.1023/B:STCO.0000035301.49549.88. URL <http://dx.doi.org/10.1023/B:STCO.0000035301.49549.88>.
- Snoeij, P.; Attema, E.; Davidson, M.; Duesmann, B.; Floury, N.; Levrini, G.; Rommen, B., and Rosich, B. Sentinel-1 radar mission: Status and performance. *IEEE Aerospace and Electronic Systems Magazine*, 25(8): 32–39, Aug 2010. ISSN 0885-8985. doi: 10.1109/MAES.2010.5552610.
- Song, K.; Zhou, X., and Fan, Y. Empirically adopted iem for retrieval of soil moisture from radar backscattering coefficients. *IEEE Transactions on Geoscience and Remote Sensing*, 47(6):1662–1672, June 2009. ISSN 0196-2892. doi: 10.1109/TGRS.2008.2009061.
- Tuia, D.; Verrelst, J.; Alonso, L.; Perez-Cruz, F., and Camps-Valls, G. Multioutput support vector regression for remote sensing biophysical parameter estimation. *IEEE Geoscience and Remote Sensing Letters*, 8(4):804–808, July 2011. ISSN 1545-598X. doi: 10.1109/LGRS.2011.2109934.
- Turnbull, H.W. *Theory of Equations*. Oliver and Boyd, London, fourth ed. edition, 1947.
- Ulaby, Fawwaz T. and Long, David G. *Microwave radar and radiometric remote sensing*. University of Michigan Press, 1st edition, 2014.
- Vapnik, Vladimir N. *The Nature of Statistical Learning Theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995. ISBN 0-387-94559-8.
- Wang, X.; Ma, L., and Wang, X. Apply semi-supervised support vector regression for remote sensing water quality retrieving. In *2010 IEEE International Geoscience and Remote Sensing Symposium*, pages 2757–2760, July 2010. doi: 10.1109/IGARSS.2010.5653832.
- Wegmuller, U.; Matzler, C.; Huppi, R., and Schanda, E. Active and passive microwave signature catalog on bare soil (2-12 ghz). *IEEE Transactions on Geoscience and Remote Sensing*, 32(3):698–702, May 1994. ISSN 0196-2892. doi: 10.1109/36.297987.

- Wu, C.; Zhang, X., and Fang, G. A novel technique for radar depth sounding of fast outlet glaciers and ice-sheet margins. *IEEE Geoscience and Remote Sensing Letters*, 12(3):587–591, March 2015. ISSN 1545-598X. doi: 10.1109/LGRS.2014.2352272.
- Wu, X.; Jezek, K. C.; Rodriguez, E.; Gogineni, S.; Rodriguez-Morales, F., and Freeman, A. Ice sheet bed mapping with airborne sar tomography. *IEEE Transactions on Geoscience and Remote Sensing*, 49(10):3791–3802, Oct 2011. ISSN 0196-2892. doi: 10.1109/TGRS.2011.2132802.
- Yang, X.; Song, Q., and Cao, A. Weighted Support Vector Machine for Data Classification. *International Journal of Pattern Recognition and Artificial Intelligence*, 2:859 – 864, August 2007.
- Young, D. A.; Roberts, J. L.; Ritz, C.; Frezzotti, M.; Quartini, E.; Cavitte, M. G. P.; Tozer, C. R.; Steinhage, D.; Urbini, S.; Corr, H. F. J.; Van Ommen, T., and Blankenship, D. D. High resolution boundary conditions of an old ice target near dome c, antarctica. *The Cryosphere Discussions*, 2016:1–16, 2016. doi: 10.5194/tc-2016-169. URL <http://www.the-cryosphere-discuss.net/tc-2016-169/>.
- Zebker, H. A. and Villasenor, J. Decorrelation in interferometric radar echoes. *IEEE Transactions on Geoscience and Remote Sensing*, 30(5):950–959, Sep 1992. ISSN 0196-2892. doi: 10.1109/36.175330.
- Zhou, Z. H. and Li, M. Semisupervised regression with cotraining-style algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 19(11):1479–1493, Nov 2007. ISSN 1041-4347. doi: 10.1109/TKDE.2007.190644.
- Zhou, Zhi-Hua and Li, Ming. Semi-supervised regression with co-training. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI'05*, pages 908–913, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc. URL <http://dl.acm.org/citation.cfm?id=1642293.1642439>.
- Zhu, Xiaojin. *Semi-Supervised Learning*, pages 892–897. Springer US, Boston, MA, 2010. ISBN 978-0-387-30164-8. doi: 10.1007/978-0-387-30164-8_749. URL http://dx.doi.org/10.1007/978-0-387-30164-8_749.
- DigitalGlobe, . Worldview-1 and worldview-2. *Imagery*.